20 *for* TWENTY

Selected Papers from AQR Capital Management on Its 20th Anniversary

FOREWORD BY

John C. Bogle

Founder of the Vanguard Group

20 for TWENTY

Selected Papers from AQR Capital Management on Its 20th Anniversary

AQR Capital Management is a pioneer as a quantitative investor and as a publisher of influential academic research. This exclusive anthology commemorates the firm's 20th anniversary. It traces the practical research contributions of AQR's founders and its team of accomplished researchers to the field of financial economics and the investing world over the past two decades.

The 20 papers selected for this collection have formed the backbone of AQR's investment philosophy. They explore the innovative ideas that have given rise to an array of systematic global investment strategies and have made a lasting impact on investor portfolios. Some of the papers provide overarching perspectives on investment questions, practices, and strategies, while others focus on practical implementation.

The book also includes reflections on the history of AQR from its early days as a start-up hedge fund to becoming a global leader in asset management. Together these essays tell the story of AQR's philosophy and its approach to investment management, which is embodied by an unwavering commitment to transparent research and meaningful client solutions.

AQR

This page intentionally left blank

Praise for *20 for Twenty*

"Read every chapter, and you'll find acute analysis, fascinating insight, and some of the best writing on finance and investing that you'll ever see."
– John C. Bogle
Founder, Vanguard Group

"The top professionals at AQR are mostly former first-rate PhD students and faculty of the Booth School of Business at the University of Chicago. I know them all quite well and respect what they bring to the money management world. This book is an excellent compilation of their data-driven research, applied and academic."
– Eugene F. Fama
2013 Nobel Laureate in Economic Sciences and Robert R. McCormick Distinguished Service Professor of Finance, University of Chicago Booth School of Business

"There is likely no better endorsement than a foreword by legend, Jack Bogle. Still, only a few firms have had as great an impact on the study of investing and portfolio management as AQR."
– Frank J. Fabozzi
Editor, *Journal of Portfolio Management* and Senior Professor and Scientific Advisor, EDHEC-Risk Institute

"This is a valuable collection of papers at the forefront of investment science. AQR's academic engagement, through the Insight Award, the impressive publication record by AQR-affiliated academics, and simply the application of academic research to cost-efficient strategies is pushing the investment community to take a more analytic approach to investing, to the benefit of investors."
– Kent Daniel
William von Mueffling Professor of Business, Columbia Business School

"As a quant enthusiast with more than 20 years of history looking at the quantitative craft developed in investing, it is great when thoughtful organizations and their people turn their focus to high quality curation. We have information overload and yet a scarcity in clear-thinking narrative on investment goals and routes to their accomplishment. This is where *20 for Twenty* scores as a very interesting and accessible compendium, which does a good job of being both contemporary and forward looking."
– Roger Urwin
Global Head of Investment Content, Willis Towers Watson

Praise for *20 for Twenty*
(Continued)

"So many of 'the smartest guys in the room' are in the big room called AQR that all serious investors will be delighted to have, read, and ponder carefully this inside report on what they have been thinking about that we should all be thinking about if we want to keep up."
– Charles D. Ellis
Founder, Greenwich Associates and author of *Winning the Loser's Game*

"As I look over the amazing body of work that Cliff Asness and his associates have produced over the past 20 years, I recall with pride my (hard but vain) attempts to hire Cliff as a fresh PhD. And what shines through these papers is what Jack Bogle, in his foreword, describes as an all-too-rare 'passion for truth'—now *that* we should all admire!"
– Martin L. Leibowitz
Managing Director in Equity Research, Morgan Stanley

20 for Twenty

Selected Papers from AQR Capital Management on Its 20th Anniversary

Foreword by
John C. Bogle
Founder, Vanguard Group

Copyright © 2018 AQR Capital Management, LLC.

All rights reserved. This book or any portion thereof may not be reproduced or used in any manner whatsoever without the express written permission of the copyright owner except for the use of brief quotations in a book review.

Disclaimer: The views and opinions expressed herein are those of the authors and do not necessarily reflect the views of AQR Capital Management, LLC or its affiliates (collectively "AQR") nor do they constitute an offer or solicitation of an offer, or any advice or recommendation, to purchase any security, strategy or investment product. The information contained herein does not represent valuation judgments with respect to any financial instrument, issuer, security or sector that may be described or referenced herein and does not represent a formal or official view of AQR. While the publisher and AQR have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability of fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sale materials. The advice and strategies contained herein may not be suitable for a reader's situation. A reader should consult with a professional where appropriate. Any liability as a result of this book (including any direct, indirect, special or consequential loss or damage) is expressly disclaimed by the publisher and AQR.

Printed in the United States of America.

ISBN 978-1-5323-8538-4 (print)
ISBN 978-1-5323-8664-0 (ebook)

First printing, 2018.

www.aqr.com

Dedicated to AQR clients and employees.

This page intentionally left blank

Contents

Foreword XIII
John C. Bogle

Preface I XV
Cliff Asness

Preface II XXI
David Kabiller

Preface III XXV
John Liew

Overview: From Research to Implementation XXIX
Antti Ilmanen, Ronen Israel, Toby Moskowitz, and Lasse Pedersen

Part I CHALLENGING CONVENTIONAL THINKING

Chapter 1: The Great Divide 3
Clifford Asness and John Liew

Chapter 2: My Top 10 Peeves 17
Clifford Asness

Chapter 3: Buffett's Alpha 31
Andrea Frazzini, David Kabiller, and Lasse Pedersen

Part II ASSET ALLOCATION: DIVERSIFY, DIVERSIFY, DIVERSIFY

Chapter 4: Why *Not* 100% Equities 61
Clifford Asness

Chapter 5: The 5 Percent Solution 71
Clifford Asness and Antti Ilmanen

Part III TO TIME OR NOT TO TIME

Chapter 6: Bubble Logic 85
Clifford Asness

Chapter 7: Fight the Fed Model 131
Clifford Asness

Chapter 8: Style Timing: Value versus Growth 149
Clifford Asness, Jacques Friedman, Robert Krail, and John Liew

Part IV DEMYSTIFYING HEDGE FUND STRATEGIES

Chapter 9: Do Hedge Funds Hedge? 167
Clifford Asness, Robert Krail, and John Liew

Chapter 10: Characteristics of Risk and Return in Risk Arbitrage 191
Mark Mitchell and Todd Pulvino

Chapter 11: An Alternative Future: Part I 229
Clifford Asness

Part V PIONEERING FACTOR INVESTING

Chapter 12: Time Series Momentum 245
Tobias Moskowitz, Yao Hua Ooi, and Lasse Pedersen

Chapter 13: Value and Momentum Everywhere 285
Clifford Asness, Tobias Moskowitz, and Lasse Pedersen

Chapter 14: Betting Against Beta 339
Andrea Frazzini and Lasse Pedersen

Chapter 15: Common Factors in Corporate Bond Returns 383
Ronen Israel, Diogo Palhares, and Scott Richardson

Chapter 16: Size Matters If You Control Your Junk 415
Clifford Asness, Andrea Frazzini, Ronen Israel, Tobias Moskowitz, and Lasse Pedersen

Part VI TAKING RESEARCH TO THE REAL WORLD

Chapter 17: The Devil in HML's Details 467
Clifford Asness and Andrea Frazzini

Chapter 18: Fact, Fiction, and Momentum Investing 495
Clifford Asness, Andrea Frazzini, Ronen Israel, and Tobias Moskowitz

Chapter 19: Fact, Fiction, and Value Investing 519
Clifford Asness, Andrea Frazzini, Ronen Israel, and Tobias Moskowitz

Chapter 20: Craftsmanship Alpha 543
Ronen Israel, Sarah Jiang, and Adrienne Ross

Contributors 565

This page intentionally left blank

Foreword

John C. Bogle

To read this anthology is a learning experience in both the theory of investing and its practicalities. Yes, there are 20 papers, but I'd be hard pressed to single out even one that you should skip over.

AQR (Applied Quantitative Research), founded in 1998, has become one of the nation's largest managers of alternative investments. But the firm (I would argue) is an unchallenged #1 in its contribution to the academic community and to serious investment professionals. The firm continues to contribute to the dialogue of the day. Rare is the newest edition of the *Financial Analysts Journal* or the *Journal of Portfolio Management* not graced with an AQR article.

Of the 20 papers included in this collection, AQR co-founder Cliff Asness has authored or co-authored 14. My personal favorites are "My Top 10 Peeves" (Chapter 2—they are my top-ten, too), "The 5 Percent Solution" (Chapter 5), "Bubble Logic" (Chapter 6), and "Fact, Fiction, and Value Investing" (Chapter 19).

But don't let my choices be your choices. Read every chapter, and you'll find acute analysis, fascinating insight, and some of the best writing on finance and investing that you'll ever see. And, whatever you decide, don't skip over Cliff's footnotes,[1] some expository, some tinged with sarcasm, all revealing the nature of this remarkable man.

Oh, yes, Cliff and I have been friends for years. We've often met, we've done joint interviews, and we've commented on one another's journal articles. (I have a slight lead in the number of published journal articles: Cliff has had 15 papers in *JPM* and 9 more in *FAJ*, for a total 24—just shy of my 28. Alas, Cliff has more years left than I do to publish additional papers, so I'm sure he'll take the lead in the years ahead! Don't miss them either!) No, we don't always agree. For example, I'm skeptical that the value factor will provide consistent return superiority. Cliff generally disagrees and (of course!) offers much more scholarly evidence to support his point. Read these AQR papers and decide for yourself.

What explains our enduring relationship? Me, the creator of the first index mutual fund and founder of a $5 trillion empire based largely on passive investing. Cliff, the consummate active investor who, with his AQR co-founders, has built one of the most successful

---

[1]Altogether, the articles in the anthology have some 362 footnotes.

alternative investment money management firms in the world? The explanation, I think, is that both of us share a passion for the pursuit of the truth; a willingness to speak truth to power, even if what we say is unpopular; and a focus on the integrity of the data (Asness the algorithmic quant, Bogle the arithmetic quant). And we're staunch friends. That's all you need to know.

Preface I

Cliff Asness

First, I can't believe it has been 20 years. I feel like there's a Bob Seger song playing in the background except for the fact that I wasn't "like a rock" 20 years ago, either.

This book is about research, so perhaps this paragraph is out of place. But I have to start by thanking some folks. First, thank you to all my colleagues at AQR. I'm consistently awed by your ability and dedication to your craft. Occasionally, during my more self-aware moments, I reflect on the difficulty I might have in getting a job at AQR today. Thankfully, I got in under the wire.

Next, I must thank our clients for their partnership. Sharing our process and research in depth (which I think is still pretty rare for investment managers) naturally leads to many very deep relationships where we and our clients make each other better. Next, I can't leave out my professors, dissertation advisors, and employers (as a teaching assistant nearly 30 years ago): Eugene Fama and Ken French. I truly got to stand on the shoulders of giants (or, more accurately, had my dissertation ridiculed and picked apart by giants until it became adequate). Penultimately I have to thank Jack Bogle, my friend and hero, who generously wrote the foreword to this tome. Thank you, Jack. Finally, in particular, I have to thank my co-founders David Kabiller, Bob Krail, and John Liew. It's been a fantastic partnership. Complementary skills plus a shared work ethic and monomaniacal client focus have been a magical combination.[1]

In his preface, my co-founder, John Liew, writes about our decision to share our research and why we think that makes both our clients and our firm better-off. I won't repeat much of what he tells you, but I will share my favorite, related quote. The physicist Howard Aiken once said "Don't worry about people stealing an idea. If it's original, you will have to ram it down their throats." Well, we still do worry a little (as John explains, we think some things we do are pretty unique, and those we do not share publicly!). But being a part of the overall academic dialogue as consumers and producers of research is just in our DNA, and I agree with John that it's a big part of whatever success we've had.

---

[1]Unfortunately, Bob Krail retired about a decade ago for health reasons and continues to ail. We truly wouldn't be here without him, and he's so good we'd take him back in a heartbeat if he ever is up to it. Get well Bob!

I am now going to ask and answer my own question: "What about our research am I most proud of?" Well, I'm glad somebody asked. First and foremost, I'm proud of the quality. I hope that's a given. Next, I'm proud of the real-world relevance to our clients. I think there is a ton of information in just these 20 papers (a small subset of our whole) that has immediate applications to practical problems. I am also quite proud of the range. Although we consistently return to certain themes, we have not even come close to writing 20 of the same papers. Finally, and most importantly, I'm proud of the honesty and the real pursuit of truth. I hope you'll agree that our papers take great pains to include every variety we can imagine of "here's one way we might be wrong" type checks. In particular, on this front I've been told that the footnotes to my own papers tip over into the grey area between being diligent/honest and crazy. Well, this is a good thing to be a bit crazy about![2]

While our papers are indeed quite eclectic, at least one major theme runs through many of them. We often demystify investing strategies previously presented as a sort of "magic." We usually note that once you do this, investors should be getting a much better deal (magic is expensive!). Of course, we think this is great for our clients, but it's also great for our business (that's how it's supposed to work!). If you oversell, claiming that an investment strategy is indeed "magic" (your own secret sauce whose recipe is kept in a vault in the basement guarded by two Rottweilers named "value" and "momentum") and charge as if it is magic (when in reality, it's good but not magically good), then when the inevitable ups and downs of real-life investing hit and the illusion of magic is popped, you're really in for it. If you've been honest, proud of what you've created and the deal it represents for investors, avoided the overselling pitfall, set realistic expectations, and charged fair fees, it goes far in making your relationship with your clients a long-lived one.

Now, I'm going to do something I think will be fun (I hope you agree). What follows are my super-short summaries of all 20 articles in this anthology. My colleagues will do a fuller review of each paper and our research in the book's overview. Here, I'm simply trying to distill the message of each paper as much as possible. Obviously, these summaries are just to whet your appetite, not a replacement for reading the book. My belief, and why I'm doing this, is that seeing the key points all at once drives home the overall breadth and importance of the contributions. Of course, given my well-known predilections, I could not resist some of the summaries being mostly humor.

So, in order of appearance:

"The Great Divide"

Gene Fama and the efficient markets hypothesis are the MVPs of modern finance, but that doesn't mean we think markets are perfectly efficient. It's complicated.

"My Top 10 Peeves"

Many things seem to annoy me. There are no sidelines. Bond funds are not less safe than the bonds they own. It's not a stock picker's market. There are many more. I don't lack for peeves.

[2]Astute readers will note I sneaked in the words "a bit."

"Buffett's Alpha"

Buffett figured out things we think are genius many years before we all did. Darn.

"Why *Not* 100% Equities"

Two-fund separation really works. Where the decision of how much risk to take, and what's the best risk-adjusted return available, should be made separately. Diversification is consistently underrated.

"The 5 Percent Solution"

It's going to be very hard for traditional assets to make their approximate (for the USA) historical returns of about 5% annually over inflation going forward. We're quants, and we're here to help.

"Bubble Logic"

Real-time calling the 1999–2000 tech bubble a bubble, a word I explicitly try to use very little (see the "My Top 10 Peeves" piece) with lots of supporting evidence and a fair amount of sarcasm. First paper where, in snarky tone, I let the real me come through (I'm still not clear it was a good idea, but I never looked back).

"Fight the Fed Model"

Stocks are much closer to "real" than "nominal" assets and shouldn't be compared to nominal bond yields. This paper ushered in many years of analysts continuing to make this faulty comparison and ignoring me entirely.

"Style Timing: Value vs. Growth"

Another 1999–2000 favorite using, and in fact creating, what has come to be called the "value spread" to measure how a strategy is priced versus its own history. This method is still in widespread use and is the subject of much discussion these days.

"Do Hedge Funds Hedge?"

Yes, but less than they'd have you think, particularly when accounting for the less-than-perfect liquidity of their assets (which makes them look less correlated, or more hedged, than reality). This one got me yelled at by about 10 famous hedge fund managers early in AQR's existence. That was both cool and terrifying.

"Characteristics of Risk and Return in Risk Arbitrage"

Merger arbitrage is, historically, a good strategy, but it's not magic. Just betting on all the mergers also works well and is more diversified—something important in a strategy that is essentially selling insurance.

"An Alternative Future: Part I"

Part I focuses on the idea that if there is "alpha" out there (define it any way you choose) then index funds plus fully hedged funds pursuing that alpha may be a good way to harvest it. Part II focuses on all the things that could go wrong. Part II didn't make the cut for this book. Part II was a downer.

"Time Series Momentum"

Trends have existed in financial assets rather ubiquitously. Importantly, they've done particularly well when needed most.

"Value and Momentum Everywhere"

Value and momentum work everywhere. That's all I've got. OK, just a little more. Importantly they work (in the statistical sense[3]) not just for choosing stocks, where they're most well-known, but for many other investing decisions. This is both useful in running portfolios and a wonderful out-of-sample test of the original individual stock-based findings.

"Betting Against Beta"

Fischer Black was right about the security market line (among many other things).

"Common Factors in Corporate Bond Returns"

Classic factors like value, momentum, quality, and carry hold for credit selection (yes, I know we're a bit of a broken record, but an out-of-sample broken record!).

"Size Matters If You Control Your Junk"

There really isn't much of a size effect largely because small is "junky" and contains a lot of low-quality stocks that have, in other research, turned out to have low expected returns. Adjusting for this resurrects the moribund size effect. Oh, and we also learned that high-class journal editors are more permissive about titles than many would guess.

"The Devil in HML's Details"

A seemingly small difference, but when to measure price in a valuation ratio (at the same time as the fundamentals, which are often lagged, versus more up to date) matters a lot.

"Fact, Fiction, and Momentum Investing"

Lots of people seem to believe things about momentum investing that are not true. We set them straight. They never go on to say these things again.

"Fact, Fiction, and Value Investing"

Then we move on to value. We next do size, but that wasn't done in time for the book. All three papers show that we're generally right and many others often wrong. Curiously, some still disagree.

"Craftsmanship Alpha: An Application to Style Investing"

The decision of what general factors, if any, you believe in is just the start to making this stuff successfully investable. How to construct a portfolio? How to implement cheaply? What's the best way to measure the concept behind a factor? Yes, this paper is a thinly veiled "don't try this at home" statement, but we still think it is right.

---

[3]I must make my standard quip here that I use the word "work" like a statistician (a little more often than random chance for long enough that it's unlikely to be random chance) not like a regular person. If your car "worked" like this, you'd fire your mechanic.

So, there it is, simple! I hope I've indeed helped get you excited about reading the rest (there is just a tad more substance within than in my summaries here). Now, without further ado (except for the other intros), let's get to the research!

This page intentionally left blank

Preface II

David Kabiller

As I reflect on our 20th anniversary, it is clear that the world has changed in many ways over the past two decades, but our mission has remained the same—an unwavering commitment to meeting our clients' objectives through shared insights, education, and often, bespoke solutions.

To commemorate this milestone, we are pleased to share with you 20 of our favorite papers that we hope will provoke some thoughts and inform your investing. These papers, while not a complete set of our work, are one of the byproducts of our culture. Core to our firm's DNA is building a culture of applied research, which is driven by a singular trait—curiosity. This, in turn, helps us attract people who are compelled to understand the complexities that drive financial markets. Our firm has created a community not only where research can flourish but also that stimulates debate both internally and externally, resulting in impactful solutions for our clients.

Recruiting Outstanding Talent

As I reflect on our journey, I've learned that delivering on our mission means recruiting outstanding talent, investing in them, and building a culture that inspires curiosity and innovation. Of course, the first step in creating such an environment is a focus on people. Here at AQR, we apply rigorous standards to recruitment and assessment. As a pioneer in quantitative finance, our approach to hiring has always been anchored to our tight relationships with academia. It's where many of our partners came from, and it's where we source candidates eager for development. We seek outstanding talent across disciplines from the best schools around the world.

Our academic relationships, together with a deep commitment to investing in our people, have generated a growing pipeline of highly accomplished and intellectually curious talent who want to join AQR. Last year we had over 250,000 visits to our career site, received over 30,000 new applications, and conducted over 10,000 interviews.

Those who join our team have a unique combination of intellectual rigor, diverse perspectives, and a commitment to our shared values. These values include seeking and speaking the truth, challenging convention, delivering education and insight, cultivating a desire to build things that persist, and understanding that success is a result of innovation and excellence. I'm most proud that we've built a strong team of leaders who have complementary skill sets and these shared values. It's a diverse group of individuals with shared DNA working together toward a singular mission. My co-founders, Cliff Asness, John Liew, and

Bob Krail, all exemplify these values and have served as great leaders of the firm. It has been my honor to work with them.

Investing in Our People and Our Clients

I've always been touched by the notion that our people are not just in service of us. The relationship goes two ways, and we, as leaders, are equally in service of our people. We have a covenant with our employees: to help them realize their potential as professionals and to help inspire their pursuit of a life well-lived.

This philosophy inspired the formal launch of AQR's QUANTA Academy three years ago as our professional and personal growth program. This learning program puts our passion for lifelong education into action and helps our people, and more recently our clients, realize their potential through three lenses: technical skills and knowledge, leadership and management, and personal enrichment. Every year, we offer in excess of 350 classes, events, and speakers for people at all levels.

Also, core to our learning and advancement is staying at the forefront of innovative and potentially disruptive technologies that are at the heart of what we do at AQR. Our Tech Talk series delves into future forces like machine learning and artificial intelligence so that we can better apply them to our clients' needs.

We believe that when our employees reach their full potential, our clients will be best served. That's why, in developing managers, from our newest to our most senior, our QUANTA programs must go beyond technical skills and knowledge and also foster leadership and management excellence and experiences.

We then take it one step further by focusing on our employees' personal enrichment. Some recent memorable events include a workshop by Nicholas Epley, a professor of behavioral science at Chicago's Booth School of Business, who helped us better understand social cognition and foster empathy. Adam Grant, a preeminent organizational psychologist, discussed the notions of "Give and Take," encouraging a culture of generosity and collaboration.

Another widely embraced initiative here at AQR is the Insights Book Club, which invites our employees to come together and discuss literary works in sessions led by top professors. *To Live*, a contemporary classic of Chinese literature by Yu Hua, was a recent selection. Michael Puett from the Chinese History Department of Harvard University facilitated a classroom-style discussion that engendered a greater cross-cultural understanding. We hope that as we work hard each day to deliver on our promise to our clients, we can also take a moment to reflect on great thinkers and great writing that can help us all better understand ourselves and each other. I believe that employees who are more self-aware tend to have deeper insight into their purpose in life, which helps them to be more productive and ultimately benefits our clients.

The Next 20 Years

I am proud of the experience we've created for our employees and the recognition we've received. In 2016, we were named one of Connecticut's great places to work by *Connecticut Magazine,* and the following year, we were named one of *Pensions and Investments'*

Best Places to Work in Money Management.[1] As we grow, we are profoundly aware that constructive iterations and evolutions are necessary in order to attract the next generation of outstanding talent, inspire them, and perpetuate our culture of curiosity, intelligent risk taking and gratitude — all in the service of investors.

While the world around us will undoubtedly continue to change over the next 20 years, I expect that our dedication to our client-centric and employee-focused mission will remain unchanged. We will continue to embrace technological advancements, but we will also always remember the primary importance of our relationships with our people and our clients.

[1]Sources: Michael Lee-Murphy, Erik Ofgang, and Albie Yuravich, "Connecticut's Great Places to Work 2016," *Connecticut Magazine*, December 2, 2016. Christine Williamson, "Best Places to Work in Money Management: AQR Capital Management." *Pensions & Investments*, December 11, 2017.

This page intentionally left blank

Preface III

John Liew

If there is one golden rule for all quant money managers, it is that if you have a good idea, you keep it a secret. Well, we've never really followed that rule. We've always shared most of our ideas and research.[1] Admittedly, in the early days we were conflicted; we tried to be traditional secretive quants, but we just couldn't do it. A running joke in our group was that if any of us were in a client meeting where we were walking that line between telling them about our investment process and trying to be secretive, a sure-fire way to get us to spill the beans would be to give us a disappointed look and say something like, "oh, is that all you guys are doing? …" I'm not sure why we are wired this way. Part of it might be our academic training, which taught us that you're supposed to share research with the world. Part of it might be our nature. In any case, as a result, AQR has been an experiment in what happens when you do the exact opposite of what you are supposed to do as a quant money manager.

So, what's been the result? In my opinion, this openness and ethos of sharing ideas has had a profound effect on improving our research, our culture, and helped us to become an industry thought leader. It has affected us in ways that I think we never expected, and maybe it has even had some impact on the money management industry.

Research

Earlier in my career, I worked at a quant firm that took a more traditional approach to being secretive. Researchers were not allowed to publish any of their research and were discouraged from even talking to each other about what they were working on. One problem with this approach is that you were left trying to come up with ideas by sitting alone in your office reading and thinking. In my experience, that's not the most fertile environment for generating interesting ideas.

So what's a better way? Bob Zimmer, the president of the University of Chicago, says that when you're building a faculty you can't start small, you have to start with a critical mass of high quality people. I think he is 100% right. One reason is that smart people want to be around other smart people, and the more the better. Another is that, in general, good ideas don't come from one smart person sitting in his or her office alone thinking deep

[1]Ok, there are a few things we don't openly share with the world.

thoughts. Instead, ideas tend to come from smart people actively engaged and talking to each other. And, more people leads to more conversations. Of course, if you believe, as I do, that ideas come from conversations, then being secretive is a killer. Having 100 smart people isolated in their offices just reading and thinking alone is so much worse than having 100 smart people actively engaged with each other. It's that collaborative environment that we've tried to create at AQR.

In fact, we've taken that idea a step further. It doesn't end with AQR professionals talking to each other. We also actively engage with the broader financial academic community.

Culture and Academic Engagement

I'm bragging now, but if you walk the halls at AQR, you'll feel like you could be at a leading academic finance department. This academic culture is something that makes AQR special. Sure, a lot of firms will hire a few academics here and there, but the breadth of our engagement exceeds everyone else in our industry.

As some evidence of this statement, here is one of my favorite slides at AQR. It's the current ranking by institution of the number of academic working paper downloads.[2] It's published by the Social Sciences Research Network (SSRN), which is the website where most academics in the social sciences put their working papers when they feel like they're ready for public consumption.

1. New York University (NYU)
2. Harvard University
3. University of Chicago
4. Stanford University
5. University of Navarra
6. University of Pennsylvania
7. Columbia University
8. Yale University
9. University of Oxford
10. Duke University
11. Massachusetts Institute of Technology (MIT)
12. AQR Capital Management, LLC
13. Government of the United States
14. University of New South Wales
15. University of Toronto
16. London School of Economics and Political Science (LSE)
17. University of Southern California
18. University of Texas at Austin
19. Cornell University
20. London Business School

[2]SSRN Finance Economic Network ranked by total new downloads of papers in the last three years. The SSRN list is as of July 1, 2018.

Pretty cool, no? These rankings jump around, but as long as I've been looking at them, we've generally been in this range and often in the top 10 (and if you didn't notice, we're the sole private company). When I say that being at AQR is like being at a top academic finance department, this is what I mean.

So why is that? How are we able to attract so many top academics? Well, finance professors usually have a tough choice to make when it comes to working in industry. The typical trade-off is if you work or consult with a money manager, you can get some extra money, get exposure to what goes on in the real world, and perhaps, get access to data and resources that you wouldn't normally get. However, most money managers won't allow academics to publish any of the results. So that creates a conflict that forces the academic to decide between making some extra money and giving up some scholarly productivity. At AQR, this trade-off doesn't exist. In fact, we encourage academics to publish the work we do together.

I'll even take it one step further. I've argued to my academic colleagues that their academic productivity increases with their involvement with AQR. Why? Because academic finance is an applied field of study and, therefore, ideas don't come out of the ether. I learned this in grad school when I took Eugene Fama's class at the University of Chicago. I was a PhD student at the time, and the class was mostly meant for PhD students, but some brave MBA students took it too. One day, after looking through all the term paper ideas, Professor Fama said to the class, "I hate to say this, but I looked through all the proposed paper topics and the MBA students' ideas were more interesting than the PhD students' ideas." What??? We were the ones who were going to do this for a living! How is this possible? In hindsight, I'm not surprised. As a group, the MBAs were the ones who had more real world experience. As a PhD student, I struggled to come up with an interesting thesis idea. I think part of it was that I was limited to trying to find ideas by reading other people's research (and thinking about what they might have missed). When I started working in industry, through the natural course of the job, interesting research ideas seemed to come up left and right (of course, finding time to work on them was a different matter!).

As AQR has become more widely known for its research and academic culture, we have been able to attract more academic talent. Going back to Bob Zimmer's point, once you have critical mass you become a viable research department. And, as you grow and attract more talent, even more talent wants to be part of it. So it starts to take on a life of its own. Perhaps that's why the best research institutions, if they maintain their culture, stay the best for a long time. It takes a lot of time and effort to get there, but once you do, you get this very positive feedback effect that keeps it going and makes you better.

Thought Leadership

Ok, that's great, but we're not an academic institution, we're a money manager. Stop with all this publishing and get back to work! Well, I would argue that all this academic publishing is work and has had a profound effect on our business through enhancing our position as a thought leader.

How is thought leadership built? A lot of financial firms write "research papers." Well, quality matters. Occasionally, a piece of research gives the reader a real insight that he or she didn't have before. That's the moment when it happens. Do this over and over, and over

a long period of time you start to build something really valuable (and no one can download it on a thumb drive in the middle of the night and walk out the door with it).

So what does that do for us? One of the best things that can happen in building a business is to be at the right place at the right time. For instance, if you happened to be an internet geek in the late 1990s before the tech bubble took off, that probably would have been a really good thing (for a while). Being early to things that become industry trends is a great way to create business success. We've experienced some of this. When we look at our business today, it looks like we were early to a lot of investment strategies and ideas that grew in popularity. We were early movers in investment products like liquid alternatives, factor investing or smart beta, and risk parity. We were also early to investment ideas like fair fees for alternatives and creating hedge funds that actually hedge.

This might be a little arrogant (or maybe a lot arrogant), but I don't think that our success has been all lucky timing. It's great to be luckily early to an industry trend. However, it's even better if you can influence what becomes the next industry trend. I think that both the volume and quality of research that we have produced over the years has earned AQR a position of being a thought leader in our industry, and that position has helped us do exactly that kind of influencing.

Final Thoughts

Ok, I have to address one final critique. If you publish your good ideas, then won't your competitors get them too, and won't they then drive those inefficiencies out of the marketplace?

There is truth to this, and this is the cost of our choice of being open about what we do. In part, it's had an effect on where we've focused our research efforts because not all strategies are equally vulnerable to this threat. It has focused our research on long- and medium-horizon strategies in mature and liquid markets—markets where you get long histories of data that are best suited to these types of strategies.

Thus, we haven't become the black-box-never-lose-money-kick-out-your-clients-closed quant firm.[3] Instead, we've become an important provider of strategies for investors that want transparency and need scalable investment options. We've shed light on what is an otherwise secretive business, and we've built investment products that we think have been to some degree disruptive to our industry. We have also, in some cases, lifted the shroud of mystery behind some hedge fund investments, demystifying strategies and concepts to help investors understand them better.[4] If in that process a few competitors have learned something too, well that's a price we're willing to pay.

[3]Of course, many of those end up losing money, and sometimes a lot!

[4]I would also add that being able to stick with good strategies through tough times is an often underappreciated source of value for investors, and something that is not easy to do. I think our transparency has true value for clients as it makes it far easier to stick with these strategies through these inevitable tough times. In contrast, it's impossible for investors to stick with a black box that goes through similar tough times.

Overview
From Research to Implementation

Antti Ilmanen, Ronen Israel, Toby Moskowitz, and Lasse Pedersen

In August 2018, AQR turned 20 years old. To commemorate this special occasion and pay homage to our deep-seated roots in research, we put together this anthology of AQR papers. As alluded to in the opening prefaces by AQR's co-founders, we have always been, and will continue to be, avid consumers and producers of research—seeking to apply the very best ideas from academia to investment practice. During the course of our 20 years, we have produced over 300 scholarly research papers, publishing some 200 of them in peer-reviewed academic and practitioner journals.[1]

To provide a complete roster of our papers here, including their findings and relevance for investment practice, would be a daunting undertaking—not to mention making the book 5,000 pages long! Instead, we sought to distill the last 20 years of our research into a focused collection of papers that covers the breadth and depth of our work, research that we hope has positively impacted the practice and study of investment management. In the end, we chose 20 papers.[2]

In addition to illustrating our investment philosophy, we highlight several themes core to AQR that permeate this collection of articles: challenging conventional wisdom, demystifying the investment landscape, applying the best academic research to real-world investment problems, and implementing these ideas in practice. Research agendas progress through time. In some cases, we discuss how our views have evolved from an earlier paper. In other cases, we focus on papers that best exemplify the ideas most succinctly and clearly, which are not necessarily the first paper in a series of related articles.

Illustrating AQR's many thought leaders, papers included in this collection, and on our website, are written by a large number of authors. Not surprisingly, as our academic roots run deep, many at our firm have also contributed to important papers with academic co-authors and other investment professionals unaffiliated with AQR. However, for this book, we include only papers in which all co-authors are affiliated with AQR.

[1]A complete list of AQR's papers and publications can be found at www.aqr.com/research.

[2]While 20 is an arbitrary number, it is a reasonably small number that also not coincidentally matches the anniversary year and allows for the catchy book title of "20 for Twenty."

We also focus on the research that we have chosen to make public. There is, of course, plenty of internal research related to proprietary ideas, methods, and strategies that may not see public light. However, as discussed in the preface by John Liew, our philosophy has always been one of hyper-transparency and erring on the side of more openness—an approach we believe benefits the profession, our clients, and yes, our business. Hence, our bias is to publish, educate, and contribute to the public discourse on ideas. We have arguably published more than any other asset manager. Nevertheless, there is some research that we are quite proud of but that has not (yet) been published and remains absent from this book.

As testament to the impact of all of this research, we make a couple of observations. First, according to SSRN (the website where scholarly minded researchers post their papers for all to see), our paper downloads place AQR among the top 15 research organizations in financial economics, with all other organizations in the top 15 being top academic universities. Among all non-academic institutions, AQR ranks first.[3] Based on the number of cited papers in the top three academic finance journals over the last five years,[4] AQR is tied for second with Yale University and is just behind the University of Chicago.[5] Many of our papers have also won prestigious academic and practitioner awards, some 54 in all, including the *Journal of Finance*'s Smith-Breeden, *Journal of Financial Economics*' Fama–Jensen DFA, *Review of Financial Studies*' Michael Brennan, *Financial Analysts Journal*'s Graham and Dodd, *Journal of Portfolio Management*'s Bernstein/Fabozzi, and *Journal of Investment Management*'s Markowitz prize.

We have organized the papers into six broad topics, each demonstrating our research approach to addressing some of the biggest challenges facing investors. Here, we discuss why these topics are so central to investing, and in the process, we provide an overview of the 20 papers included in the book (highlighted in **boldface**). In addition, we briefly touch on some of the papers that did not make it into this anthology but even so importantly contributed to our research agenda over the years.

In this brief overview, we have not tried to write a history of thought on each topic. Hence, many of our own papers will be omitted from this discussion, and more importantly, many worthy papers by others are also not mentioned. We cannot even do justice to our own lengthy research list, so attempting to do so for the exhaustive research done by others would be futile. That said, we acknowledge here that many of the ideas we have had the privilege and pleasure to work on, and hopefully contribute to, have had many parents. Research is collaborative and ongoing. The biggest ideas that influence science are a fusion of facts, theories, and ideas from an array of contributors. So, with apologies to the many colleagues inside and outside of AQR whose work deserves mentioning as predecessors, contemporaries, or followers of our work—including some of the heroes Cliff Asness mentions in his preface—we do not attempt to cover others' contributions.

---

[3]Social Science Research Network (SSRN) Finance Economic Network ranked by total new downloads of papers in the last three years. The SSRN list is as of July 1, 2018.

[4]The *Journal of Finance*, *Journal of Financial Economics*, and *Review of Financial Studies*. Google Scholar list as of August 10, 2017.

[5]We note that some of our authors are cross-listed and so contribute to AQR's high totals but also to those of Yale University, New York University, and the University of Chicago.

I. Challenging Conventional Thinking

The first section of this book deals with broader investment questions including challenging conventional wisdom. One of the most basic questions in finance is whether (and to what extent) markets are informationally efficient. This is a crucial yet daunting issue for investors, because if markets are perfectly efficient, then no one can beat the market as prices are always "right" (meaning they reflect all available information and are not affected by non-information). In this world, alpha does not exist, so investors should pursue purely passive investing. The alternative view is that markets are somewhat inefficient (maybe significantly so) because prices are set by humans who share common behavioral biases. For instance, people can become exuberant, follow the herd toward rising or falling prices, and can panic in a rush for the exit when the market is turbulent. These actions may push prices away from their fundamental values, potentially causing bubbles and crashes. An active investor may therefore beat the market by profiting from these price inefficiencies (at the expense of the investors who create the inefficiencies).

The world, of course, is not so black and white. Markets cannot be perfectly efficient all of the time, nor are they always completely inefficient. Complicating matters is the difficulty in knowing whether a profitable trading strategy or return-enhancing opportunity is due to inefficient mispricing or is compensation for non-diversifiable risk (the "joint hypothesis problem" noted by Fama 1970). In addition, as a practical matter, someone must *make* markets efficient by collecting information and trading on it. That activity requires a lot of work, and investors must be compensated for it with higher returns or alpha, which cannot go to zero (a point famously made by Grossman and Stiglitz 1980).

Through much of our research and writings, AQR has tried to apply these concepts to investment practice. In "**The Great Divide," Asness and Liew (2014)** examine the market efficiency debate after the Nobel prize was awarded to Gene Fama (the father of market efficiency), Robert Shiller (a "founding father" of behavioral finance), and Lars Hansen (the creator of econometric methods to test asset pricing theories)**.** Through the lens of an investment manager who grapples with these same issues, they argue that while market efficiency is an extremely useful practical benchmark, some active investment opportunities do exist. As the authors put it, market efficiency is "our North Star even if we often, or always, veer 15 degrees left or right of it." These opportunities may be exploiting inefficiencies in markets or harvesting risk premia that other investors may find difficult or costly to pursue. Capturing these profit opportunities is not easy because they involve risk, require skill to implement, and are subject to other smart investors competing to take advantage of them. Echoing these same themes, Pedersen (2015), Garleanu and Pedersen (2016), and Pedersen (2018) argue why alpha must exist and also why it may be difficult to find in what Pedersen (2015) calls an "efficiently inefficient" market.

AQR has also taken a micro approach to studying these issues. For example, the extraordinary performance of Warren Buffett has often been taken as a sign that markets cannot be fully efficient. How else could Buffett have been so successful? Others have argued that Buffett may simply have been lucky, the random winner among the legions of money managers who have tried to beat the market. Buffett has contended that his success is not the result of luck as he is not alone—other investors using similar methods have also performed well coming from the same intellectual philosophy pioneered by Benjamin Graham and David Dodd. If pure luck, why would these successful investors have anything

in common? Like AQR, Buffett attributes his success to his academic roots and the lessons learned from Professors Graham and Dodd about the benefits of value and quality investing—buying cheap stocks with sound fundamentals.

Buffett's performance and its link to value and quality investing concepts are analyzed in **"Buffett's Alpha," Frazzini, Kabiller, and Pedersen (2018)**. The paper finds that Buffett's success can be attributed to his investment style and that his returns can be explained by exposure to cheap, high-quality, and low-risk securities. These investment styles, or factors in academic parlance, explain much of his performance. In other words, Buffett's success arises from a deep understanding of the benefit of these investment themes and their long-term success rather than individual stock selection. To his credit, Buffett recognized these thematic advantages early on, and his very skillful implementation of them, including careful risk management and clever use of insurance products to leverage the strategy, produced his exceptional track record. Brooks, Tsuji, and Villalon (2018) look at other superstar investors, including Bill Gross, George Soros, and Peter Lynch, and find that similar analysis reveals that their track records can also be explained by a set of common investment styles. These papers also demystify hedge fund strategies, a topic we discuss further in Section IV.

Finally, AQR has always tried to bring insights to investment practice, especially when challenging conventional wisdom. In **"My Top 10 Peeves," Asness (2014b)** debunks many commonly held but misguided investment concepts and conventional wisdoms. As one example, Asness discusses how the term "bubble" is way overused and that it should be reserved for only those cases in which prices cannot be justified by any reasonable future outcome (e.g., stocks in early 2000). It should not be used in cases when prices are merely expensive. He then delves into the best ways to evaluate investments, looking at investment horizons (the fallacy of focusing on too-short evaluation periods), and benchmarking (most hedge fund strategies are not fully hedged). He also makes clear that "active" refers to any strategy that deviates from a market-cap-weighted index and that investors should worry about companies that do not expense stock options (building on Asness 2004c). In addition to **"My Top 10 Peeves,"** Asness (2009) and Goyal, Ilmanen, and Kabiller (2015) seek to improve portfolios by offering investors practical advice, avoiding bad habits, and debunking investing myths.

II. Asset Allocation: Diversify, Diversify, Diversify

Over the past half century, asset allocation practices have been guided by modern portfolio theory, belief in the equity premium, and evolving institutional constraints and conventions. At least in the Anglo-American markets, many portfolios converged in the late 1900s toward a 60/40 stock/bond allocation. More recently, many endowments, foundations, and public pension plans seeking diversifying return premia outside traditional markets have increased their allocations to alternative assets.

Although the equity premium has been the most popular source of long-run return, we have argued that many portfolios rely too heavily on it. In typical portfolios, such as the 60/40 stock/bond allocation, 90% of total portfolio risk often comes from equity markets, implying poor risk diversification. In **"Why *Not* 100% Equities," Asness (1996)** highlights the problem of concentrated risk in many portfolios and a key insight into what would come to be called "risk parity" investing almost a decade later. The article emphasizes a core lesson in modern portfolio theory: selecting the optimal portfolio among risky assets and deciding how much risk an investor takes are distinct decisions, as long as cash

can be borrowed and lent without constraint. Creating a portfolio with comparable risks coming from stocks and bonds improves risk diversification, and with the use of leverage, such a portfolio can reach any portfolio risk target where the expected return exceeds that of a traditional equity risk–dominated portfolio, such as a 60/40.

Although risk parity portfolios later added other asset classes beyond stocks and bonds and other such features as dynamic risk targeting, the core idea of better portfolio risk diversification combined with leverage was originally advocated in "**Why Not 100% Equities.**" Later research by Asness, Frazzini, and Pedersen (2012) linked the risk parity strategy's long-run outperformance explicitly to better diversification, which is often ignored due to investor leverage aversion. With risk parity, leverage and diversification go hand in hand, and while leverage increases risk, the article argues that it is more manageable and reliable than the risk associated with a more concentrated portfolio.

Whatever the chosen asset allocation, investors have in recent years faced the challenge of low expected returns (if not yet the pain of low realized returns). Low starting yields and high valuations in virtually all long-only assets make it seemingly hard to achieve the 5% long-run real return that typical stock/bond allocations promised and delivered during the 20th century. In **"The 5 Percent Solution," Asness and Ilmanen (2012)** first show that prevailing market yields point to lower prospective real returns near 2.5%, about half of the historical real average return. Annual updates in the AQR *Alternative Thinking* reports show that the outlook for 60/40 portfolios is broadly similar in 2018 to that in 2012 and that the low expected return challenge is not specific to U.S. markets.

"The 5 Percent Solution" offers prescriptions for investors who want to achieve a 5% long-run real expected return in this more challenging environment. Unlike others, who propose raising expected returns through larger equity allocations or illiquid alternative allocations, the paper emphasizes the benefits of harvesting multiple return sources from not only traditional equity markets but also alternative risk premia and proprietary alpha. Market risk premia can be harvested efficiently through risk parity portfolios as discussed earlier and by diversifying across international markets as discussed in Asness, Israelov, and Liew (2011) and Ilmanen, Maloney, and Ross (2014). Alternative risk premia are liquid long/short strategies with well-known rewards and diversification abilities. Two forms of alternative risk premia are discussed later, hedge fund premia (Section IV) and style or factor premia (Section V). Finally, the paper points out that while alpha is scarce, it can be generated not only by identifying new proprietary strategies but also through effective implementation, as we discuss in Section VI.

Besides identifying good long-run return sources, **"The 5 Percent Solution"** argues that investors need to accept some degree of unconventionality, including the use of such financial tools as leverage, shorting, and derivatives (which Asness has dubbed "the three dirty words in finance") to achieve these performance goals. The article summarizes AQR's distinctive investment philosophy for relatively unconstrained investors.

III. To Time or Not to Time

If timing is a sin, then our short answer is to sin a little.

The longer answer is that we find strategic diversification, as discussed in the previous section specifically for traditional markets, more effective than tactical timing. Hence, timing should be used sparingly and cautiously. We do not completely admonish timing, but

rather endorse humility and risk allocations that reflect the relatively low Sharpe ratios and concentrated nature of timing bets.

Our early research held some hope for contrarian timing strategies. For example, **"Bubble Logic: Or, How to Learn to Stop Worrying and Love the Bull," Asness (2000),** published near the tech bubble peak, made a prescient timing call against the dot-com era valuations. The paper discusses many arguments being made to justify the high valuations of tech stocks at the time, challenging each of them along the way. Looking back and recognizing how difficult timing is in practice, this particular episode with its extreme valuations may be a rare illustrative case in which some contrarian timing may have been warranted.

In **"Fight the Fed Model," Asness (2003)** evaluates one contrarian market timing model that focuses on the yield spread between equities and Treasuries. Although the Fed model is logically inconsistent (as it mixes real and nominal indicators) and has limited ability to predict long-horizon equity market returns, it does have some predictive ability over short horizons. Applying similar ideas to timing long/short factors, **"Style Timing: Value vs. Growth," Asness, Friedman, Krail, and Liew (2000)** reports some ability of value spreads (the relative valuations of value stocks versus growth stocks) to predict subsequent performance of the value strategy when selecting U.S. stocks. These early, encouraging results are consistent with the classic framework of Campbell and Shiller (1998) and match the literature on timing reviewed in Ilmanen (2011).

However, our latest research has made us more humble about the promise of timing strategies, because of both weak out-of-sample evidence and various econometric problems associated with long-horizon predictability. Asness, Ilmanen, and Maloney (2017) find relatively disappointing empirical evidence when contrarian strategies are applied in practical out-of-sample contexts to timing U.S. equities and duration timing in U.S. Treasuries. Asness (2016a), Asness, Chandra, Ilmanen, and Israel (2017), and Ilmanen, Israel, Moskowitz, Thapar, and Wang (2018) find weak evidence for contrarian timing of long/short style premia over out-of-sample periods from the original studies on these style premia, including going back a full century. In addition, Boudoukh, Israel, and Richardson (2018) discuss the statistical challenges of finding reliable evidence for timing strategies.

However, our latest research is not entirely negative on timing. For instance, Asness, Liew, Pedersen, and Thapar (2017) find some timing benefit to deep value opportunities (much like the extreme valuations of the tech era) broadly within industries, across countries, bond markets, and currencies. Conversely, the use of procyclical strategies—based on trend following or momentum—are shown by Asness, Ilmanen, and Maloney (2017) and Gupta and Kelly (2018) to enhance performance somewhat.

Based on our research, we find it hard, economically or statistically, for tactical timing to greatly improve upon strategic long-term diversification. Some dynamic market and factor timing may be considered ("sin a little"), especially in extreme events, but even then should be used cautiously and with multiple signals (contrarian and trend) that complement each other.

IV. Demystifying Hedge Fund Strategies

Our research has often attempted to challenge conventional wisdom and to take on tough topics, even at the risk of being unpopular with many of our peers. Perhaps the best example of this attribute is evident in our work on hedge fund strategies, where we have tried

to both extol the virtues and point out the failures of hedge funds. Our hope is that through this research, we have made an impact on how investors think about their hedge fund investments, how they allocate to them, and what they ultimately pay in fees for what they get from these investments.

One of the main criticisms we have hurled against hedge funds is that they don't "hedge" enough. We have argued that more often than not hedge fund strategies have a meaningful, long equity market bias. In **"Do Hedge Funds Hedge?" Asness, Krail, and Liew (2001)** take on this topic directly by examining to what extent hedge funds hedge. Importantly, the authors show that the presence of equity market exposure in hedge fund returns is even more pronounced after accounting for the fact that hedge funds often own illiquid and difficult-to-price securities. An investor should be concerned with this market exposure for two reasons. First, it means that hedge fund strategies are not as diversifying as they could be to a traditional investor's portfolio, which is already typically highly exposed to equities. At a minimum, investors should account for this market exposure when allocating to hedge funds instead of assuming that what is being provided is a market-neutral source of return. Second, passive equity market exposure is something that can be obtained easily and at very low cost, so investors should not pay high hedge fund fees for it.[6] Obviously, this paper didn't make us any friends in the hedge fund community.

While we have criticized hedge funds for running with a higher market beta than they should (and charging large fees for that beta), we have also been vocal in discussing that they actually do hedge to some degree (i.e., they don't run at a beta of one to the market). Thus, comparing their returns with a passive long-only equity market index is also wrong. This issue is particularly important during bull markets, such as the decade since the Global Financial Crisis. It is not surprising that hedge funds would lag the equity market index over this period if they have a beta less than one and may also lag a typical long-only equity mutual fund whose beta is greater than one. While a simple, but important, point, it is one that requires repeated reminders, as addressed in Asness (2014a), Asness (2016b), and Asness (2018).

Beyond market exposure, our research has further decomposed hedge fund returns. We believe hedge fund strategies can be broken down into three distinct parts: exposure to traditional, common risk (e.g., equity market exposure), exposure to well-known, dynamic sources of alternative risks and returns ("hedge fund betas" or "hedge fund risk premia"), and true, unique alpha. The recognition that you can decompose hedge fund returns in this way was an important insight with far-reaching implications. For example, hedge fund betas, while significantly more complex to implement than traditional betas, can offer positive long-term, risk-adjusted returns that are uncorrelated to traditional assets. Because these hedge fund betas are more common and not unique, however, they should be obtained at lower fees than pure alpha returns or what typical hedge funds were charging. This research brought to focus the discussion of what an investor is getting for the fees they pay.

Our earliest work on this concept of decomposing hedge fund returns can be found in **"Characteristics of Risk and Return in Risk Arbitrage"** by **Mitchell and Pulvino**

[6]There is a direct parallel in active fixed income investing, where active managers often load heavily on credit beta, leading to similar questions about reduced diversification and fair fees; see AQR's *Alternative Thinking* "The Illusion of Active Fixed Income Diversification" (2017).

(2001). In that paper, we examine a classic, well-known hedge fund strategy, merger arbitrage. A common argument at the time was that the value creation in merger arbitrage came from a manager's unique ability to predict the deals that would go through and avoid the deals that would fail (i.e., unique alpha). The paper shows, however, that there is a systematic component to merger arbitrage returns that comes from arbitrageurs providing liquidity to fundamental holders after deals are announced and from taking on systematic deal failure risk (in essence, hedge funds were providing deal insurance to the marketplace.) As a result, value can be derived without taking a view on individual deals. Instead, by simply diversifying across all merger deals, one can capture the systematic return component while diversifying away idiosyncratic deal risk. In this way, a merger arbitrageur is compensated for providing a service to the marketplace, not for unique insights into the probability of deal success or failure. This result means that an investor could gain exposure to the "hedge fund beta" element of merger arbitrage without having to identify and pay high fees to a star manager. Or stated differently, it raises the bar on the definition of alpha and provides a way to benchmark merger arbitrage managers to see if they were truly adding unique value.

The implications of this research were dramatic, giving rise to the notion of hedge fund beta or hedge fund risk premia and changing how investors invest in hedge funds and think about their hedge fund investments. In **"An Alternative Future: Part I," Asness (2004a)** anticipates this trend and discusses how the role of hedge funds would change given this new notion of hedge fund betas. Asness then follows up with Part II in Asness (2004b) by extending the concept of hedge fund betas to other markets and strategies. This idea is then developed further by Berger, Crowell, and Kabiller (2008).

As another example, Hurst, Ooi, and Pedersen (2013) demonstrate how the returns of managed futures funds and CTAs (commodity trading advisors) can be explained by a simple time-series momentum factor, demystifying the returns generated by these funds and, as a byproduct, helping to the lower fees investors pay for accessing these strategies.

Our research also demonstrates how convertible arbitrage can be implemented systematically by buying a diversified portfolio of convertible bonds and hedging their equity risk using standard tools (Mitchell, Pedersen, and Pulvino 2007). More broadly, Pedersen (2015) explains the principles behind the major hedge fund strategies and shows how each core strategy can be captured systematically.

When discussing hedge fund betas, it is important to recognize that evolution is at play. As discussed in Berger, Crowell, and Kabiller (2008), strategies that once were idiosyncratic and unique often become more common and better understood over time. That doesn't take away the value they offered in the past when few understood or pursued them, and it doesn't mean that they are not valuable strategies going forward. It simply implies that the way investors access these strategies will change, often including paying less for them.

Frankly, a lot of the "is it alpha or is it beta?" debate comes down to semantics. Practically speaking, if you're an investor who has no exposure to hedge fund betas, it is alpha to you!

V. Pioneering Factor Investing

Factor investing is the concept of grouping securities together by commonly shared characteristics that are related to expected returns and risks. The concept was first introduced in academia, backed by economic theory (Merton 1973 and Ross 1976). Guided by this idea,

an explosion of empirical work, perhaps most famously led by Gene Fama and Kenneth French, over the next several decades produced a host of "factors" or common characteristics that predicted returns. In fact, the proliferation of academic research generated several hundred candidate factors, leading to recent criticisms of many factors being non-robust and data-mined (overfit to the tested sample) and not surviving real-world implementation costs.

AQR has contributed to this area of research through the discovery of new factors, extending known factors to previously unexplored asset classes and markets, studying how factors interact with one another and across markets and asset classes. In particular, addressing the above-mentioned issue of overfitting, AQR has analyzed the robustness of factors, including questions of implementation (discussed in Section VI) and documenting and testing various economic theories for the existence of these factors.

It is our firm belief that a truly reliable and important factor is one that has strong economic intuition, pervasive and robust in-sample empirical evidence, strong out-of-sample evidence supporting it in a variety of contexts (including across different time periods and asset classes), and contributes significantly to live portfolios after considering implementation issues. Together, these criteria constitute a pretty high bar that only a few factors are able to clear. Factors that appear robust, pervasive, real, and applicable are value, momentum, trend, defensive/quality, carry, variance risk, and liquidity, all of which we have written about extensively. However, we have also written about factors that do not pass muster, identifying the criteria on which they fail, such as the size premium.

AQR research has made several important contributions to the discovery of new factors. For example, in **"Time Series Momentum," Moskowitz, Ooi, and Pedersen (2012)** present a new factor called time series momentum (TSMOM), which is distinct from the cross-sectional momentum factor discovered by Jegadeesh and Titman (1993) and Asness (1994). While cross-sectional momentum is a relative strength strategy (go long securities that did better relative to other securities and short those that did relatively worse), TSMOM is an absolute strength or trend-following strategy that is long all securities with positive past returns and short those with negative past returns. In **"Time Series Momentum,"** the authors implement their strategy among nearly 60 futures contracts that include equity indices, government bonds, currencies, and commodities and find large average returns across all contracts. Although trend-following strategies had been deployed and discussed in practice, this pioneering paper in academia helped clarify, in part, the debate over market efficiency, as discussed in Section I, by challenging even the weakest notion of market efficiency. The paper also attempts to identify where the influences come from, finding that behavioral theories of initial underreaction and delayed overreaction are most likely driving the excess returns to the factor. Importantly, the time series momentum factor introduced in this paper has performed particularly well during extreme stock market booms and busts, the latter feature providing unique hedging benefits to equity markets (see Hurst, Ooi, and Pedersen 2017 for a century's worth of evidence on these hedging properties).

Expanding upon new momentum factors, Cohen and Frazzini (2008) uncovered predictable return patterns between customer and suppliers, which they dubbed "customer momentum," that launched a series of studies looking at economic links between firms. Brooks (2017) examines "macro momentum" over a half century, studying the return predictability of momentum in such macroeconomic indicators as production growth and inflation.

Another factor discovery is presented in **"Betting Against Beta," Frazzini and Pedersen (2014)**, which took an old fact—the relationship between market beta and average

returns being too flat[7]—and formed a factor to exploit it by going long low-beta and short high-beta stocks, levered to the same risk, which they termed "betting against beta," or BAB. The BAB factor produces positive risk-adjusted returns over the last century in the United States, in 20 other international equity markets, and in government bonds, credit, and futures. Subsequent work by Asness, Frazzini, and Pedersen (2014) shows that this effect is not driven by industry bets or missing factors. **"Betting Against Beta"** argues that the BAB factor is driven by leverage aversion and borrowing constraints faced by some investors, who then overweight high-beta stocks in an effort to achieve desired risk levels. This added demand for risky (i.e., high-beta) securities make them expensive, while safer (low-beta) assets become cheap, explaining why the BAB factor delivers high average risk-adjusted returns. Frazzini and Pedersen (2012) also show that assets with embedded leverage, such as index and stock options and levered ETFs (exchange-traded funds), generate similar effects. Asness, Frazzini, Gormsen, and Pedersen (2017) find that decomposing beta into its correlation and volatility components further helps substantiate the leverage story.

More broadly, Asness, Frazzini, and Pedersen (2018) examine a whole host of factors related to firm quality (e.g., profitability, growth, and low risk) and find that they share similarities. Using a novel equity valuation model, they connect a variety of quality and defensive characteristics that populate the accounting and finance literatures. Forming one composite factor they call "quality minus junk," or QMJ, encapsulates and condenses these characteristics into a single unifying theme, giving structure to what was previously a disparate set of signals.

In addition to new factor discoveries, AQR was among the first to apply well-known factors to other asset classes. Asness, Liew, and Stevens (1997) use factors commonly used to select individual stocks, such as value and momentum, to select country equity indices. In **"Value and Momentum Everywhere," Asness, Moskowitz, and Pedersen (2013)** find strong evidence that value and (cross-sectional) momentum, which predominantly had only been applied to U.S. equities, explain returns well not only in international equities and country index futures but also in government bonds, currencies, and commodities. While others had applied these factors to some of the asset classes (applying value to government bonds and commodities and momentum to government bonds was novel), no previous study had examined these effects simultaneously across all markets. Using value and momentum to connect asset classes in a unifying framework, the paper found stronger and more reliable factor return premia and more precise estimates of factor risk exposure to, for example, macroeconomic and liquidity risks, than would be detectable looking only at one asset class.

Another impactful result from **"Value and Momentum Everywhere"** was the discovery that a value strategy in one asset class is positively correlated to value strategies in other asset classes and a momentum strategy in one asset class is positively correlated to momentum strategies in other asset classes, yet value and momentum strategies are negatively correlated to each other, both within and across asset classes. These results suggest commonality among value strategies across diverse assets and commonality among momentum strategies, yet strong diversification benefits between value and momentum.

[7]Black, Jensen, and Scholes (1972) show this empirically. **"Betting Against Beta"** updated their work using a full century of stock return data and many other markets and asset classes, also extending the theory of Black (1972), who first showed that leverage constraints could cause the low-beta effect.

While both value and momentum each exhibit positive correlation across asset classes, those correlations are modest, suggesting there remain diversification benefits from applying these factors across asset classes. Moreover, neither value nor momentum is correlated to the underlying asset class. Assembling these results, there appears to be tremendous diversification benefits from applying value and momentum factors simultaneously to the same asset class as well as across asset classes.

The novel application of value and momentum to many asset classes besides equities, and the discovery of strong correlation structure across diverse asset classes for a given factor, has made a significant contribution to the literature. As testament to its impact, since its publication in 2013, **"Value and Momentum Everywhere"** is the most cited among all papers published in the *Journal of Finance*.[8]

Continuing with the "everywhere" theme—that is, pioneering, ubiquitous factors—Koijen, Moskowitz, Pedersen, and Vrugt (2018) apply the carry factor, previously only examined in currencies, to nine other asset classes and find that a multi-asset carry factor is significantly stronger than within any single asset class. Furthermore, the crashes associated with currency carry (Brunnermeier, Nagel, and Pedersen 2008) are ameliorated with a diversified carry factor. Similarly to **"Value and Momentum Everywhere,"** they show that carry strategies are slightly positively correlated across asset classes but still exhibit diversification benefits from being applied across diverse assets. Similarly, Israelov and Nielsen (2015), Israelov and Klein (2016), and Ang, Israelov, Sullivan, and Tummala (2018) show evidence of and diversification benefits from applying variance risk premia to many asset classes. In **"Common Factors in Corporate Bond and Bond Fund Returns," Israel, Palhares, and Richardson (2018)** provide an original application of (value, momentum, carry, and defensive/BAB) factors to corporate bonds, where factor premia also reside. Brooks and Moskowitz (2017) uniquely show that the same factors explain returns across the yield curves of government bonds globally. These papers also demonstrate the diversification benefits of using multiple factors within the same asset class.

Taking this theme a step further, Asness, Ilmanen, Israel, and Moskowitz (2015) build the case for a portfolio that diversifies across *both* asset classes and factors (value, momentum, carry, and defensive), where the diversification benefits across factors are substantially larger than those across asset classes. Ilmanen, Israel, Moskowitz, Thapar, and Wang (2018) show similar results going back a century with unique out-of-sample data, indicating that these factors were a significant part of markets well before they were studied as a science.

AQR's research has also provided clarity by ruling out and challenging the conventional wisdom about some factors. For example, Israel and Moskowitz (2012b) and Alquist, Israel, and Moskowitz (2018) show that the size effect—which many believe to be an important factor—does not appear to have solid evidence behind it, in or out of sample, is not robust, and does not have a good economic story. Taking it a step further, in **"Size Matters, If You Control Your Junk," Asness, Frazzini, Israel, Moskowitz, and Pedersen (2018)** show that once you control for a quality factor, or its inverse "junk," many of these challenges hurled at the size effect disappear and a significant size premium emerges, which appears to be related to compensation for liquidity and liquidity risk (e.g., Acharya and Pedersen 2005).

[8]From Google Scholar, July 17, 2018.

While the concept of factors began decades earlier, AQR has contributed to some of the most cited and innovative research in this area and is a pioneer of factor-based investing in practice. This research has led to the discovery of new factors, provided the first applications of factors to new asset classes, demystified real-world investments, and improved our understanding of investing. We have also made sole contributions to real-world implementation issues associated with these factors that were absent from the literature and which we discuss next.

VI. Taking Research to the Real World

Earlier sections of this book deal with AQR's research efforts in identifying and documenting various sources of returns, from asset classes to classic hedge fund strategies to factors or styles. This research is obviously important, but too often missing from typical academic studies is a discussion of real world application and implementation. As practitioners, these are key issues and ones about which we care tremendously, devoting considerable energy to researching and understanding them. Seemingly small differences in how you build and implement strategies can lead to strikingly different results for investors. Delivering a world-class investment process requires a focus on all the little details. The sum total of many small, individual value-added steps can result in considerable long-term benefit. In many ways, we believe that efficient and effective implementation is one of the truest forms of "alpha." While perhaps brazen, we often refer to this notion, for lack of a better term, as the "craftsmanship" of investing.

In addition to conducting implementation research for practical purposes, we also wrote about it and published much of it in an effort to educate and fill a void in the literature. Also, as discussed earlier, it is just our nature to do so.

One form of craftsmanship is to think more deeply about the very definitions of the investment themes themselves to see if more efficient harvesting of the underlying return sources can be achieved. For example, in a couple of papers we argue that the traditional, basic definitions of certain factors can be improved upon. In **"The Devil in HML's Details," Asness and Frazzini (2013)** argue that replacing the traditional academic value measure of book-to-price ratio (which lags the market price by 6- to 18-months) with a measure that uses the current market price provides more up-to-date information. By lagging the market price, the traditional academic value measure introduces incidental (and likely unintended) exposure to momentum. Using a more updated market price provides a better proxy for true "value." This paper is a great example of how precision matters, even in something as simple as price lags, and there are many more enhancements that we, as practitioners, use to build better signals and themes. Indeed, the devil is in the details, as the paper's title reminds us.

Beyond definitions, how you combine and capture multiple themes can have a profound impact on the risk and return prospects of portfolios. For example, **"Craftsmanship Alpha: An Application to Style Investing," Israel, Jiang, and Ross (2018),** and **"Fact, Fiction, and Value Investing," Asness, Frazzini, Israel, and Moskowitz (2015)** discuss the importance of using multiple definitions of signals to enhance robustness, and combining multiple themes, geographies and asset classes to enhance diversification when building portfolios. When combining multiple themes, "**Craftsmanship Alpha,"** as well as Fitzgibbons, Friedman, Pomorski, and Serban (2017) and Frazzini, Israel, Moskowitz,

and Novy-Marx (2013) emphasize the importance of integrating (i.e., endogenously combining signals) versus simply combining separate, standalone themes.

When building and implementing portfolios in practice, it is also critical to take into account implementation costs. There are two related aspects. One is the ability to minimize costs during the implementation process. The other is to fully understand what exactly those costs are and how they vary across assets and over time to ensure that the right trade-offs are being made when constructing effective portfolios. Transaction costs and taxes are two main costs associated with implementing portfolios. **"Craftsmanship Alpha,"** Frazzini, Israel, and Moskowitz (2012), Ross, Moskowitz, Israel, and Serban (2017), and Garleanu and Pedersen (2016) show that the costs of implementation can be dramatically reduced by rebalancing and trading in a smart and patient way. Our research shows that the costs assumed in most academic studies are much higher than in reality, that many of the main factors easily survive transaction costs, and that one can build better net-of-cost portfolios by optimizing and allowing for some deviation from the theoretical portfolio to capture the bulk of the factor exposure while minimizing costs. Similar arguments can be made about taxes. Israel and Moskowitz (2012a) and Sialm and Sosner (2018) show that what drives tax liabilities is not simply trading or turnover but the characteristics of the turnover and that certain strategies are more tax efficient than many would guess based on just turnover, as they tend to realize both long-term gains and short-term losses (e.g., momentum). Once again, optimized portfolios that allow for slight deviation from the theoretical portfolio can minimize the impact of taxes.

In thinking about real-world applications, often an important step is to challenge conventional thinking and take on common misconceptions. In an effort to separate fact from fiction, we dissect in a series of papers some of the main academic factors. In **"Fact, Fiction, and Momentum Investing," Asness, Frazzini, Israel, and Moskowitz (2014)** deal with the topic of momentum, in **"Fact, Fiction, and Value Investing"** the same authors deal with value, and Alquist, Israel, and Moskowitz (2018) deal with the size factor. These papers tackle some common myths around each factor and factor investing in general. The myths emanate from academia, practitioners, and investors. Using economic arguments, widely circulated academic papers, and publicly available data to refute these myths, we attempt to further the understanding of factor investing. We confront questions of implementation, as discussed earlier, including how best to capture these factors in practice. We also focus on other misunderstood aspects of factor investing, such as whether these factors are at risk of going away and how best to allocate to them.

As a final point, when it comes to craftsmanship, the key is to capture a lot of little edges that collectively add up over time. Over shorter time periods, certain choices can work for or against you for random reasons. For example, there can be times when one specific definition of a factor performs better than a robust collection of definitions intended to capture the same theme, or times when transactions costs are higher than anticipated. But, if these decisions are based on sound economic intuition, long-term evidence, and are robust to different geographies, asset classes, and sample periods, you can reasonably expect that they will add value over the long run.

This collection of articles emphasizes how, over the course of the past 20 years, AQR has been at the forefront of capturing path-breaking investment concepts in real-world portfolios, our steadfast commitment to being transparent with our strategies and research, and of course, educating the broader investment community on investing best practices.

The 20 articles selected for this anthology provide a sample of the research we hope has made a meaningful impact on the profession and the direction of research. We can say with certainty, though, that they have made a lasting impression on our firm, its philosophy, and our approach to investment management.

References

Acharya, Viral V., and Lasse H. Pedersen. 2005. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics* 77 (2) (August): 375–410.

Alquist, Ron, Ronen Israel, and Tobias J. Moskowitz. 2018. "Fact, Fiction and the Size Effect." *Journal of Portfolio Management*, forthcoming.

Ang, Ing-Chea, Roni Israelov, Rodney Sullivan, and Harsha Tummala. 2018. "Understanding the Volatility Risk Premium." White paper, AQR Capital Management.

Asness, Clifford S. 1994. "The Power of Past Stock Returns to Explain Future Stock Returns." Working paper, AQR Capital Management.

_____. 1996. "Why Not 100% Equities." *Journal of Portfolio Management* 22 (2): 29–34.

_____. 2000. "Bubble Logic: Or, How to Learn to Stop Worrying and Love the Bull." Working paper, AQR Capital Management.

_____. 2003. "Fight the Fed Model." *Journal of Portfolio Management* 30 (1): 11–24.

_____. 2004a. "An Alternative Future: Part 1." *Journal of Portfolio Management* 30 (5), 30th Anniversary Issue: 94–103.

_____. 2004b. "An Alternative Future: Part II." *Journal of Portfolio Management* 31 (1): 8–23.

_____. 2004c. "Stock Options and the Lying Liars Who Don't Want to Expense Them." *Financial Analysts Journal* 60 (4): 9–14.

_____. 2009. "Seven Thoughts on Running Big Money for the Long-Term." Working paper, AQR Capital Management.

_____. 2014a. "Hedge Funds: The (Somewhat Tepid) Defense." *Cliff's Perspective*, AQR Capital Management.

_____. 2014b. "My Top 10 Peeves." *Financial Analysts Journal* 70 (1): 22–30.

_____. 2016a. "The Siren Song of Factor Timing, aka 'Smart Beta Timing' aka 'Style Timing'." *Journal of Portfolio Management* 42 (5), Quantitative Equity Strategies Special Issue: 1–6.

_____. 2016b. "Hedging on the Case Against Hedge Funds." *Bloomberg View*, May 12, 2016.

_____. 2018. "The Hedgie in Winter." *Cliff's Perspective*, AQR Capital Management.

Asness, Cliff, Swati Chandra, Antti Ilmanen, and Ronen Israel. 2017. "Contrarian Factor Timing Is Deceptively Difficult." *Journal of Portfolio Management* 43 (5), Quantitative Strategies: Factor Investing Special Issue: 72–87.

Asness, Cliff, and Andrea Frazzini. 2013. "The Devil in HML's Details." *Journal of Portfolio Management* 39 (4): 49–68.

Asness, Clifford S., Andrea Frazzini, Ronen Israel, and Tobias J. Moskowitz. 2014. "Fact, Fiction, and Momentum Investing." *Journal of Portfolio Management* 40 (5), 40th Anniversary Issue: 75–92.

_____. 2015. "Fact, Fiction, and Value Investing." *Journal of Portfolio Management* 42 (1): 34–52.

Asness, Cliff, Andrea Frazzini, Ronen Israel, Tobias Moskowitz, and Lasse Pedersen. 2018. "Size Matters If You Control Your Junk." *Journal of Financial Economics* 129 (3): 479–509.

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen. 2012. "Leverage Aversion and Risk Parity." *Financial Analysts Journal* 68 (1): 47–59.

_____. 2014. "Low-Risk Investing without Industry Bets." *Financial Analysts Journal* 70 (4): 24–41.

_____. 2018. "Quality Minus Junk." *Review of Accounting Studies*, forthcoming.

Asness, Clifford, Andrea Frazzini, Niels Joachim Gormsen, and Lasse Pedersen. 2017. "Betting Against Correlation." Working paper, AQR Capital Management.

Asness, Cliff, Jacques Friedman, Robert J. Krail, and John M. Liew. 2000. "Style Timing: Value vs. Growth." *Journal of Portfolio Management* 26 (3): 50–60.

Asness, Clifford, and Antti Ilmanen. 2012. "The 5 Percent Solution." *Institutional Investor*, May.

Asness, Clifford, Antti Ilmanen, Ronen Israel, and Tobias Moskowitz. 2015. "Investing with Style." *Journal of Investment Management* 13 (1): 27–63.

Asness, Cliff, Antti Ilmanen, and Thomas Maloney. 2017. "Market Timing: Sin a Little, Resolving the Valuation Timing Puzzle." *Journal of Investment Management* 15 (3): 23–40.

Asness, Clifford, Antti Ilmanen, Tobias Moskowitz, Lukasz Pomorski, and Lei Xie. 2018. "Who Is on the Other Side?" Working paper, AQR Capital Management.

Asness, Clifford, Roni Israelov, and John Liew. 2011. "International Diversification Works (Eventually)."*Financial Analysts Journal* 67 (3): 24–38.

Asness, Cliff, Robert J. Krail, and John M. Liew. 2001. "Do Hedge Funds Hedge?" *Journal of Portfolio Management* 28 (1): 6–19.

Asness, Clifford, and John Liew. 2014. "The Great Divide." *Institutional Investor*, March 2014.

Asness, Clifford, John M. Liew, Lasse H. Pedersen, and Ashwin Thapar. 2017. "Deep Value." Working paper, AQR Capital Management.

Asness, Cliff, John M. Liew, and Ross L. Stevens. 1997. "Parallels between the Cross-Sectional Predictability of Stock and Country Returns." *Journal of Portfolio Management* 23 (3): 79–87.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen. 2013. "Value and Momentum Everywhere." *Journal of Finance* 68 (3): 929–985.

Berger, Adam, Brian Crowell, and David G. Kabiller. 2008. "Is Alpha Just Beta Waiting to Be Discovered?" White paper, AQR Capital Management (revised 2012).

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business* 45 (3): 444–455.

Black, Fischer, Michael C. Jensen, and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*, edited by M.C. Jensen. New York: Praeger.

Boudoukh, Jacob, Ronen Israel, and Matthew Richardson. 2018. "Long Horizon Predictability: A Cautionary Note." *Financial Analysts Journal*, forthcoming.

Brooks, Jordan. 2017. "A Half Century of Macro Momentum." White paper, AQR Capital Management.

Brooks, Jordan, and Tobias Moskowitz. 2017. "Yield Curve Premia." Working paper, AQR Capital Management.

Brooks, Jordan, Severin Tsuji, and Daniel Villalon. 2018. "Superstar Investors." *Journal of Investing*, forthcoming.

Brunnermeier, Markus, Stefan Nagel, and Lasse Pedersen. 2008. "Carry Trades and Currency Crashes." *NBER Macroeconomics Annual* 23: 313–347.

Campbell, John Y., and Robert J. Shiller. 1998. *Journal of Portfolio Management* 24 (2): 11–26

Cohen, Lauren, and Andrea Frazzini. 2008. "Economic Links and Predictable Returns." *Journal of Finance* 63 (4): 1977–2011.

Fama, Eugene. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (2), Papers and Proceedings of the Twenty-Eighth Annual Meeting of the American Finance Association: 383–417.

Fitzgibbons, Shaun, Jacques Friedman, Lukasz Pomorski, and Laura Serban. 2017. "Long-Only Style Investing: Don't Just Mix, Integrate." *Journal of Investing* 26 (4): 153–164.

Frazzini, Andrea, and Lasse H. Pedersen. 2012. "Embedded Leverage." Working paper, AQR Capital Management.

_____. 2014. "Betting Against Beta." *Journal of Financial Economics* 111 (1): 1–15.

Frazzini, Andrea, David Kabiller, and Lasse Pedersen. 2018. "Buffett's Alpha." *Financial Analysts Journal*, forthcoming.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz. 2012. "Trading Costs of Asset Pricing Anomalies." Working paper, AQR Capital Management.

Frazzini, Andrea, Ronen Israel, Tobias J. Moskowitz, and Robert Novy-Marx. 2013. "A New Core Equity Paradigm." White paper, AQR Capital Management, March.

Garleanu, Nicolae, and Lasse Pedersen. 2016. "Efficiently Inefficient Markets for Assets and Asset Management." Working paper, AQR Capital Management.

Goyal, Amit, Antti Ilmanen, and David Kabiller. 2015. "Bad Habits and Good Practices." *Journal of Portfolio Management* 41 (4): 97–107.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70 (3): 393–408.

Gupta, Tarun, and Bryan Kelly. 2018. "Factor Momentum Everywhere." Working paper, AQR Capital Management.

Hurst, Brian, Yao Hua Ooi, and Lasse Pedersen. 2013. "Demystifying Managed Futures." *Journal of Investment Management* 11 (3): 42–58.

_____. 2017. "A Century of Evidence on Trend-Following Investing." *Journal of Portfolio Management* 44 (1): 15–29.

"The Illusion of Active Fixed Income Diversification." 2017. *Alternative Thinking,* AQR Capital Management, December 7.

Ilmanen, Antti. 2011. *Expected Returns: An Investor's Guide to Harvesting Market Rewards*. John Wiley and Sons: West Sussex, U.K.

Ilmanen, Antti, Thomas Maloney, and Adrienne Ross. 2014. "Exploring Macroeconomic Sensitivities: How Investments Respond to Different Economic Environments." *Journal of Portfolio Management* 40 (3): 87–99.

Ilmanen, Antti, Ronen Israel, Tobias Moskowitz, Ashwin Thapar, and Franklin Wang. 2018. "A Century of Evidence: Value, Momentum, Carry, and Defensive." Working paper, AQR Capital Management.

Israel, Ronen, and Tobias Moskowitz. 2012a. "How Tax Efficient Are Equity Styles?" Working paper, AQR Capital Management.

Israel, Ronen, and Tobias Moskowitz. 2012b. "The Role of Shorting, Firm Size, and Time on Market Anomalies." Working paper, AQR Capital Management.

Israel, Ronen, Sarah Jiang, and Adrienne Ross. 2018. "Craftsmanship Alpha: An Application to Style Investing." *Journal of Portfolio Management* 44 (2), Multi-Asset Special Issue: 23–39.

Israel, Ronen, Diogo Palhares, and Scott Richardson. 2018. "Common Factors in Corporate Bond Returns." *Journal of Investment Management* 16 (2): 17–46.

Israelov, Roni, and Lars N. Nielsen. 2015. "Covered Calls Uncovered." *Financial Analysts Journal* 71 (6): 44–57.

Israelov, Roni, and Matthew Klein. 2016. "Risk and Return of Equity Index Collar Strategies." *Journal of Alternative Investments* 19 (1): 41–54.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1): 65–91.

Koijen, Ralph S.J., Tobias J. Moskowitz, Lasse Heje Pedersen, and Evert B. Vrugt. 2018. "Carry." *Journal of Financial Economics* 127 (2): 197–225.

Merton, Robert. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5) (September): 867–887.

Mitchell, Mark L., and Todd C. Pulvino. 2001. "Characteristics of Risk and Return in Risk Arbitrage." *Journal of Finance* 56 (6): 2135–2175.

Mitchell, Mark L., Lasse Heje Pedersen, and Todd Pulvino. 2007. "Slow Moving Capital." *American Economic Review* 97: 215–220.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse H. Pedersen. 2012. "Time Series Momentum." *Journal of Financial Economics* 104 (2): 228–250.

Pedersen, Lasse Heje. 2015. *Efficiently Inefficient: How Smart Money Invests and Market Prices Are Determined.* Princeton University Press.

_____. 2018. "Sharpening the Arithmetic of Active Management." *Financial Analysts Journal* 74 (1): 21–36.

Ross, Stephen A. 1976. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13 (3): 341–360.

Ross, Adrienne, Tobias J. Moskowitz, Ronen Israel, and Laura Serban. 2017. "Implementing Momentum: What Have We Learned?" Working paper, AQR Capital Management.

Sialm, Clemens, and Nathan Sosner. 2018. "Taxes, Shorting, and Active Management." *Financial Analysts Journal* 74 (1): 88–93.

This page intentionally left blank

PART I

Challenging Conventional Thinking

The Great Divide
My Top 10 Peeves
Buffett's Alpha

This page intentionally left blank

The Great Divide

The Nobel committee decided to split the economics prize this time—and that's okay

Clifford Asness and John Liew

Every December the Royal Swedish Academy of Sciences concludes a 16-month nomination and selection process by awarding the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, founder of the Nobel Prize. The Nobel committee recently recognized work on the Efficient Market Hypothesis with a dramatic splitting of the prestigious prize between EMH pioneer Eugene Fama and EMH critic Robert Shiller. (University of Chicago economist Lars Hansen also shares the $1.2 million prize, but we only briefly had the math chops to understand his work back in the late 1980s; we're told he is very deserving!) This makes now a great time to review EMH, its history, its controversies, where things stand today—and perhaps make our own small contribution to the discussion.

By way of background, we both got our Ph.D.s at the University of Chicago under Gene Fama and consider him one of the great mentors of our lives and an extraordinary man. This might reasonably worry a reader that we are very biased. But for the past 20 years, we've also pursued investment strategies we think are at least partly explained by market inefficiencies. We pursued these through the Asian crisis in 1997, the liquidity crisis of 1998, the tech bubble of 1999–2000, the quant crisis of August 2007, the real estate bubble and ensuing financial crisis culminating in 2008 and (for Cliff) the New York Rangers' not making the National Hockey League playoffs for seven years in a row, starting in 1997.Throughout this experience we have more than once come face-to-face with John Maynard Keynes's old adage that "markets can remain irrational longer than you can remain solvent," a decidedly folksier and earlier version of what has come to be known as the limits of arbitrage—a concept we will return to in this article. We could arrogantly describe our investment strategies as a balanced and open-minded fusion of Fama and

---

Reprinted from *Institutional Investor*, March 2014, with permission.

Shiller's views but admit they could also be described uncharitably as "risk versus behavioral schizophrenia."

All of this has put us somewhere between Fama and Shiller on EMH. We usually end up thinking the market is more efficient than do Shiller and most practitioners—especially, active stock pickers, whose livelihoods depend on a strong belief in inefficiency. As novelist Upton Sinclair, presumably not a fan of efficient markets, said, "It is difficult to get a man to understand something, when his salary depends upon his not understanding it!" However, we also likely think the market is less efficient than does Fama. Our background and how we've come to our current view make us, we hope, qualified—but perhaps, at the least, interesting—chroniclers of this debate.

Last, we seek to make a small contribution to the EMH conversation by offering what we think is a useful and very modest refinement of Fama's thoughts on how to test whether markets are in fact efficient. We hope this refinement can help clarify and sharpen the debate around this important topic. Essentially, we strategically add the word "reasonable" and don't allow a market to be declared efficient if it's just efficiently reflecting totally irrational investor desires. If you thought that last line was confusing, good. Keep reading.

The concept of market efficiency has been confused with everything from the reason that you should hold stocks for the long run (and its mutated cousins, arguments like the tech bubble's "Dow 36,000") to predictions that stock returns should be normally distributed to even simply a belief in free enterprise. This last idea is the closest to reasonable. It is true that there is a strong correlation between those who believe in efficient markets and those who believe in a laissez-faire or free-market system; however, they are not the same thing. In fact, you do not have to believe markets are perfectly efficient or even particularly close to believe in a mostly laissez-faire system. Though it may have implications for many of these things, market efficiency is not directly about any of these ideas.

So what does it really mean for markets to be efficient? As Fama says, it's "the simple statement that security prices fully reflect all available information." Unfortunately, while intuitively meaningful, that statement doesn't say what it means to reflect this information. If the information at hand is that a company just crushed its earnings target, how is the market supposed to reflect that? Are prices supposed to double? Triple? To be able to make any statement about market efficiency, you need to make some assertion of how the market should reflect information. In other words, you need what's called an equilibrium model of how security prices are set. With such a model you can make predictions that you can actually observe and test. But it's always a joint hypothesis. This is famously, in the narrow circles that care about such things, referred to as the joint hypothesis problem. You cannot say anything about market efficiency by itself. You can only say something about the coupling of market efficiency and some security pricing model.

For example, suppose your joint hypothesis is that EMH holds and the Capital Asset Pricing Model is how prices are set. CAPM says the expected return on any security is proportional to the risk of that security as measured by its market beta. Nothing else should matter. EMH says the market will get this right. Say you then turn to the data and find evidence against this pairing (as has been found). The problem is, you don't know which of the two (or both) ideas you are rejecting. EMH may be true, but CAPM may be a poor model of how investors set prices. Perhaps prices indeed reflect all information, but there are other risk factors besides market risk that investors are getting compensated for

bearing. Conversely, CAPM may precisely be how investors are trying to set prices, but they may be failing at it because of investors' behavioral biases or errors. A third explanation could be that both EMH and CAPM are wrong. We will argue later that although the joint hypothesis is a serious impediment to making strong statements about market efficiency, this problem does not have to make us nihilistic. Within reason, we believe we can still make useful judgments about market efficiency.

This framework has served as the foundation for much of the empirical work that has gone on within academic finance for the past 40 years. The early tests of market efficiency coupled efficiency with simple security pricing models like CAPM. The joint hypothesis initially held up well, especially in so-called event studies that showed information was rapidly incorporated into security prices in a way consistent with intuition (if not always with such a formal equilibrium model). However, over time some serious challenges have come up. These can be broadly grouped into two categories: microchallenges and macrochallenges.

The microchallenges center on what are called return anomalies. Of course, even the term "anomaly" is loaded, as it means an anomaly with respect to the joint hypothesis of EMH and some asset pricing model (like, but certainly not limited to, CAPM). Within this category of challenges, researchers have identified other factors that seem to explain differences in expected returns across securities in addition to a security's market beta. Two of the biggest challenges to the joint hypothesis of EMH and CAPM are value and momentum strategies.

Starting in the mid-1980s, researchers began investigating simple value strategies. That's not to say value investing was invented at that time. We fear the ghosts of Benjamin Graham and David Dodd too much to ever imply that. This was when researchers began formal, modern academic studies of these ideas. What they found was that Graham and Dodd had been on to something. Stocks with lower price multiples tended to produce higher average returns than stocks with higher price multiples. As a result, the simplest diversified value strategies seemed to work. Importantly, they worked after accounting for the effects of CAPM (that is, for the same beta, cheaper stocks still seemed to have higher expected returns than more expensive stocks). The statistical evidence was strong and clearly rejected the joint hypothesis of market efficiency and CAPM.

The reaction? Academics have split into two camps: risk versus behavior. The risk camp says the reason we are rejecting the joint hypothesis of market efficiency and CAPM is that CAPM is the wrong model of how prices are set. Market beta is not the only source of risk, and these price multiples are related to another dimension of risk for which investors must be compensated. In this case the higher expected return of cheaper stocks is rational, as it reflects higher risk.

The behaviorists don't buy that. They say the reason we're rejecting the joint hypothesis of market efficiency and CAPM is that markets aren't efficient; behavioral biases exist, causing price multiples to represent not risk but mispricing. Prices don't reflect all available information because these behavioral biases cause prices to get too high or too low. For instance, investors may overextrapolate both good and bad news and thus pay too much or too little for some stocks, and simple price multiples may capture these discrepancies. Another way to say this: The market is trying to price securities according to some rational model like CAPM but falling short because of human frailty. Thus the market is not efficient.

Very much along the same lines as the value research, in the late 1980s researchers such as Narasimhan Jegadeesh, Sheridan Titman and Cliff Asness (yes, the dissertation of one of the authors—bias alert) began empirical studies of diversified momentum strategies. The studies found that stocks with good momentum, as measured quite simply by returns over the previous six months to a year, tended to have higher average returns going forward than stocks with poor momentum, again fully adjusting for any return differences implied by CAPM or any other rational equilibrium model known at the time—more evidence against the joint hypothesis.

In contrast to value, this finding has been considerably harder to deal with for efficient-market proponents. Cheap stocks tend to stay cheap for a long time. They are usually crappy companies (we apologize for the technical term). Thus it is not a stretch to believe there is something risky about these stocks for which the willing holder gets compensated. As a result, we find it inherently plausible, even if hard to precisely define, that these may be riskier stocks to a rational investor. But price momentum changes radically from year to year. What kind of risk changes so quickly? Can a stock be risky one year and then safe the next? You can't dismiss such a thing. Extreme performance in either direction may inherently change risk characteristics. But most researchers, including EMH fans, still find it quite hard to devise a story that reconciles the success (net of CAPM and value) of momentum with a risk factor story.

Last, value and momentum are negatively correlated factors. This observation adds to the challenge. Negative correlation means that a portfolio of the two reduces risk because when one is hurting your portfolio, the other tends to be helping. Because both value and momentum average positive long-term returns, this risk reduction creates a higher risk-adjusted return to the portfolio. Furthermore, value and momentum are not just useful for U.S. stock picking. Both of these effects are incredibly robust within stock markets around the world, as well as for a broad array of other asset classes, such as bonds, currencies and commodities. The larger the total risk-adjusted return generated by a market-neutral (no exposure to CAPM) strategy, the bigger the challenge. In this sense, value and momentum are more than the sum of their parts.

The bottom line is that there are some factors, like momentum, that at this point seem to pose a considerable challenge to EMH. The verdict is more mixed for value, but most would agree that it still presents an additional challenge. Add to that the power of combining value and momentum, and at the very least it is fair to say there are important micro-challenges to EMH.

On the macro side—meaning, dealing with the whole markets, not relative value—one of the 2013 Nobel winners, longtime Yale University economist Bob Shiller, points out a puzzling observation in his now-famous 1981 paper, "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" Stock prices should be the present value of future dividends. So Shiller determines what he calls an ex post rational price for the stock market as a whole by computing the present value of actual future dividends. This is obviously cheating, because in real life you don't know the value of future dividends. The market has to make forecasts. However, the cheating is for an interesting and honest purpose. It turns out that dividends are not very volatile and change smoothly through time; Shiller asserts that reasonable forecasts should reflect this characteristic. The striking observation is how much actual market prices swing around Shiller's "cheating" rational price. Can reasonable forecast uncertainty justify

such wildly changing market prices? Shiller, one of the central protagonists of our story, says no.

Proponents of efficient markets will point out that Shiller's methodology uses a constant discount rate. Yet there can be times when people require a higher rate of return (or discount rate) to bear the risk of owning stocks, and there can be times when people require a lower rate. If discount rates vary over time, even without any change in expected future dividends, prices should change, and that can have a big impact on the level of market prices. Thus on first principles EMH fans say we would expect market prices to vary more than Shiller's version of a "rational" price. Again we run into the joint hypothesis problem. Can reasonable equilibrium models produce such time-varying required rates of return on the stock market?

It's clear that while insightful, original and thought-provoking, Shiller's observation is not quite as damning as his original interpretation asserts. Rest assured, soon we will also level polite criticism of those supporters of EMH who put too much stock in this ability of changing discount rates to save their story. We aim to be relatively equal-opportunity offenders.

So where do we stand? Spoiler alert: after a lot of discussion and 20 years of implementing much of what we have discussed, and a lot more than just value and momentum, we're still confused. Putting on a more positive spin, perhaps this is why finance is such a live and interesting field.

We started our careers in the early 1990s, when as a young team in the asset management group at Goldman, Sachs & Co. we were asked to develop a set of quantitative trading models. Why they let a small group of 20-somethings trade these things we'll never know, but we're thankful that they did. Being newly minted University of Chicago Ph.D.s and students of Gene Fama and Ken French, the natural thing for us to do was develop models in which one of the key inputs was value. We also used momentum from the get-go (as Cliff had written his dissertation on it), but here we'll focus on the simple value story, as it explains most of what happened in the early days.

(As an aside, one of Cliff's favorite stories is asking Fama, no natural fan of momentum investing, if he could write his thesis on momentum, and Fama responding, "If it's in

Taking the High Road
Buying cheap stocks and shorting expensive ones has done well over time.
Source: AQR Capital Management.

the data, write the paper" and then fully supporting it. That kind of intellectual honesty doesn't come along too often.)

Above is a graph of the cumulative returns to something called HML (a creation of Fama and French's). HML stands for "high minus low." It's a trading strategy that goes long a diversified portfolio of cheap U.S. stocks (as measured by their high book-to-price ratios) and goes short a portfolio of expensive U.S. stocks (measured by their low book-to-price ratios).The work of Fama and French shows that cheap stocks tend to outperform expensive stocks and therefore that HML produces positive returns over time (again, completely unexplained by the venerable CAPM). The graph above shows this over about 85 years.

If you notice the circled part, that's when we started our careers. Standing at that time (before the big dip you see rather prominently), we found both the intuition and the 65 years of data behind this strategy pretty convincing. Obviously, it wasn't perfect, but if you were a long-term investor, here was a simple strategy that produced positive average returns that weren't correlated to the stock market. Who wouldn't want some of this in their portfolio?

Fortunately for us, the first few years of our live experience with HML's performance were decent, and that helped us establish a nice track record managing both Goldman's proprietary capital, which we began with, and the capital of some of our early outside investors. This start also laid the groundwork for us to team up with a fellow Goldman colleague, David Kabiller, and set up our firm, AQR Capital Management.

As fate would have it, we launched our first AQR fund in August 1998. You may remember that as an uneventful little month containing the Russian debt crisis, a huge stock market drop and the beginning of the rapid end of hedge fund firm Long-Term Capital Management. It turned out that those really weren't problems for us (that month we did fine; we truly were fully hedged long-short, which saved our bacon), but when this scary episode was over, the tech bubble began to inflate.

We were long cheap stocks and short expensive stocks, right in front of the worst period for value strategies since the Great Depression. Imagine a brand-new business getting that kind of result right from the get-go. Not long cheap stocks alone, which simply languished, but long cheap and short expensive! We remember a lot of long-only value managers whining at the time that they weren't making money while all the crazy stocks soared. They didn't know how easy they had it. At the nadir of our performance, a typical comment from our clients after hearing our case was something along the lines of "I hear what you guys are saying, and I agree: These prices seem crazy. But you guys have to understand, I report to a board, and if this keeps going on, it doesn't matter what I think, I'm going to have to fire you." Fortunately for us, value strategies turned around, but few know the limits of arbitrage like we do (there are some who are probably tied with us).

With this experience in mind, let's go back to the debate over whether the value premium is the result of a value-related risk premium or behavioral biases. What does it feel like sitting in our seats as practitioners who have traded on value for the past 20 years? To us, it feels like some of both at work.

The risk story is actually quite compelling. One prerequisite for this story is that for risks to command a risk premium, they must not be diversifiable. What we saw in the tech bubble was an extreme version of exactly that. Cheap stocks would get cheaper across the board at the same time. It didn't matter if the stock was an automaker or an insurance

company. When value was losing, it was losing everywhere. We saw the same phenomenon on the expensive side. Furthermore, though very pronounced in the tech bubble, this seems to be the norm. There is a strong factor structure to value. In other words, cheap assets and expensive assets tend to covary, or move together. (This is true not only for value in stocks but for value within most asset classes we've looked at.) This doesn't prove that value is a risk factor—you could imagine it occurring in a model based on irrationality—but it is a very direct implication of a rational risk-based model.

However, if you're looking for us to make a final decision, we, as promised, offer you disappointment. There are reasons to believe some or even a lot of the efficacy of value strategies (at times) is behavioral. In addition to the long list of reasons that behaviorists put forth, we'll offer a couple of thoughts.

Throughout our experience managing money, we've seen that a lot of individuals and groups (particularly committees) have a strong tendency to rely on three- to five-year performance evaluation horizons. Of course, looking at the data, this is exactly the horizon over which securities most commonly become cheap and expensive. Put these two observations together and you get a large set of investors acting anticontrarian. One of our favorite sayings is that these investors act like momentum traders over a value time horizon. To the extent the real world is subject to price pressure, and of course it is, you'd expect this behavior to lead to at least some mispricing (inefficiency) in the direction of value.

Also, many practitioners offer value-tilted products and long-short products that go long value stocks and short growth stocks. But if value works because of risk, there should be a market for people who want the opposite. That is, real risk has to hurt. People should want insurance against things like that. Some should desire to give up return to lower their exposure to this risk. However, we know of nobody offering the systematic opposite product (long expensive, short cheap). Although this is far from a proof, we find the complete lack of such products a bit vexing for the pure rational risk-based story.

Last, one thing often ignored in the EMH-versus-behaviorist debate is that there is not necessarily a clear winner in reality. Life, and the large subset of our lives (perhaps sadly) revolving around security prices, can be driven by a mix of rational and behavioral forces. Researchers looking for a clean answer don't tend to love this fact. They all seem to want to be the declarers of a clear winner (and possibly the next Nobel laureate to come out of these studies). But the real world does not exist to make financial researchers happy, and both rational and irrational forces may be at work.

Furthermore, if value works because of a mix of rational and irrational forces, there is absolutely no reason to believe this mix is constant through time (in fact, that would be very odd). In our view, it's likely that at most times risk plays a significant role in value's effectiveness as a strategy (the EMH story). However, there are times when value's expected return advantage seems like it is driven more by irrational behavioral reasons. We believe that even the most ardent EMH supporters will admit, if only when they are alone at night, that in February 2000 they thought the world had gone at least somewhat mad. (We are tempted to say there are no pure EMH believers in foxholes.)

The tech bubble wasn't just a cross-sectional "micro" phenomenon (value versus growth within the stock market), but the whole market itself was priced to extremely high levels (versus any measure of fundamentals). This brings us to Shiller's macro critique of EMH. How is it possible that prices rationally vary so much given the relative stability of

dividends? EMH supporters' argument about time-varying discount rates is plausible, at least in direction. However, periods like 1999–2000 present a challenge for these explanations. Take a look at the chart, which we called "The Scariest Chart Ever!" in our first-quarter 2000 letter to investors (in which we also pleaded with them not to fire us). It's a graph of the Shiller P/E from 1881 to the end of March 2000.

Is it possible that a rational market could ever be priced so high that it simply could not deliver an acceptable long-term risk premium without making absolutely incredible assumptions about future dividends? We think not. In other words, we think the discount rate would have to be implausibly low to save EMH from Shiller this time. We think this one was a bubble.

Efficient marketers often point to the fact that it seems to be very difficult for active managers to consistently beat the market. But does this mean the market is efficient? Not necessarily. You can have an inefficient market that is hard to beat because of the limits of arbitrage. Even if markets are wrong, taking advantage of them is still risky. Further, given human biases—of money managers, their clients and whomever their clients report to—additional effective limits to arbitrage can be imposed, making even an inefficient market difficult to beat. We, of course, have firsthand experience with this. John says that before and after the tech bubble Cliff aged like Lincoln before and after the Civil War. (No, we are not elevating sticking with a value strategy to ending slavery and preserving the union—though perhaps it's on a par with the Battle of Vicksburg.)

Along these lines, as much deserved recognition as Shiller has gotten for calling the stock market bubble, remember that he was saying very similar things at least as far back as 1996. In fact, the famous term "irrational exuberance" was Federal Reserve chairman Alan Greenspan's statement, inspired by the analysis of Shiller and his colleague John Campbell. Note: Unlike near the peak of 2000, in 1996 we did not assert we were in a bubble and wouldn't change that view now even with hindsight. Only near the peak in early 2000 did we think the word "bubble" could be applied. Other times, like 1996, or today, seem to us to be periods where the stock market offers lower expected returns than average, but is still perhaps rational.

Thankfully for us, our value strategies, when combined with all else we did, only began hurting a year or so before the bubble burst. We doubt we could have survived losing for significantly longer than that. Someone listening to Shiller starting in 1996 likely would have lost money without much recovery, as few if any investors could have stuck with this recommendation to reap the ultimate reward so far down the road.

Although failure to beat the market doesn't mean markets are efficient, the opposite would have clear implications. If we found the market was easy to beat with great regularity, it would be a blow to efficiency as well as to most equilibrium models. It's asymmetric. Nobody said this was fair.

Along these lines, some critics of EMH get a lot of joy pointing to the handful of long-term successful money managers, like Warren Buffett, and, less well known outside the hedge fund world, the amazing returns of James Simons' Renaissance Technologies. Taking billions of dollars out of the market at low risk for a handful of people is a big deal to the manager in question (call that a mastery of the blindingly obvious). But, as perhaps the exceptions that prove the rule, even this is not much of a blow against EMH in general. The idea that markets are perfectly efficient was always an extreme and unlikely hypothesis. (Fama told us this in class in the 1980s.) The amazing success of a relatively few is,

The Scariest Chart Ever
The P/E on the S&P 500 quintupled from 1982 to 2000.
Source: AQR Capital Management.

of course, very interesting. However, as rich as these few have become, they are still very small versus the size of markets and much easier to identify after the fact than before. Thus it's less of a blow to EMH than some behaviorists would make it. We told you we'd be equal-opportunity offenders!

Let's go back to the joint hypothesis. It says you can only test the combination of some equilibrium model and market efficiency together. Does that mean you can propose any model of market equilibrium? And if that model's predictions are consistent with the data, can you declare success? We think not—at least, not with just any equilibrium model.

Suppose you imagine some investors get joy from owning particular stocks (for example, being able to brag at a cocktail party about the growth stocks they own that have done well over the past five years). One way to describe this: Some investors have a "taste" for growth stocks beyond simply their effect on their portfolios. It certainly can be rational to them to accept somewhat lower returns for this pleasure. But even if rational to the individuals who have this taste, if some investors are willing to give up return to others because they care about cocktail party bragging, can we really call that a rational market and feel this statement is useful? If so, what would we call irrational? One clear critique we have for EMH fans is that it seems some at times take the joint hypothesis too far and allow for unreasonable equilibrium models. In our view, this simply shouldn't count.

In constructing law (yes, we are borrowing from lawyers, arguably a more suspicious lot than economists), often you need to interject the word "reasonable" to make it work. We think this applies to the joint hypothesis problem as well. That is, for the purposes of making statements about market efficiency, we should examine only combinations of a reasonable model of market equilibrium and EMH. Reasonable in this case should mean a model based on clearly rational behavior, as that is the point. Suppose the only models that save EMH are unreasonable (like a model that just asserts people don't mind losing on growth versus value, as it's fun!). In that case—though you can never prove a theory but only fail to reject it (and this includes evolution, relativity and the theory that

eventually, if they are successful enough, quants will finally get the girls)—we would have to say EMH has been dealt a serious blow. To save EMH from any particular attack, in our view, you must produce not just any model of market equilibrium that bails it out, but a reasonable one.

Under the reasonable joint hypothesis, to make statements about market efficiency we should consider only the combination of market efficiency and reasonable models of how equilibrium prices are set, with "reasonable" meaning passing some intuitive tests. You cannot endogenize irrationality into the model itself.

The price of our critique is subjectivity about what is reasonable, but we believe that has always been unavoidable, if unstated. In other words, we believe our modification is just making de jure what always has been de facto. For instance, in event studies, though simple risk adjustment is often undertaken using some equilibrium model, it rarely if ever is focused on or seems to matter. Implicitly, researchers believe that no reasonable equilibrium model could explain consistent short-term profits if such are found.

Without the above modification, codified as we do by adding and interpreting the extra word "reasonable," there is always a potential way to save efficient markets, by contending that irrational models might drive equilibrium but markets still are efficient (that is, reflect all information). That might literally fit the classic definition, but it is not what has come to be meant by efficient markets and in our view violates much of the spirit and much of the point of the debate. EMH has come to mean some type of generally rational market. If all we mean by efficiency is that the market is bat-sh*t nuts but that bat-sh*t nuts is being accurately reflected in prices, we find that empty.

As concrete examples, we do believe some EMH proponents have proposed explaining things like the 1999–2000 bubble with tastes, as we described earlier, or discount rates that vary beyond a plausible amount. Can a market that efficiently reflects these irrational things in prices save EMH? Our mildly stronger version of the joint hypothesis above would rule out these defenses. To us they miss the point and create an untestable hypothesis. Again we ask: If irrational tastes are allowed in EMH, what can we ever find that we'd call inefficient? If that is the empty set, then what is EMH really saying? We believe it is saying a lot, but only if such defenses are out of bounds.

Does the above make us behaviorists? Maybe, but we still think most declared behaviorists go too far. Reading the behaviorist literature, you might get the impression that anomalies are everywhere and easily profited from. We've spent many years both studying and trading on these anomalies. Our experience, though certainly a net positive, is that many of these are out-of-sample failures. That is, it's relatively easy to find something that looks like it predicts return on paper, and it's also relatively easy to come up with a seemingly plausible behavioral rationale for why markets might be missing something. But when you actually try to trade on the anomaly (the best kind of out-of-sample test if done for long enough in a consistent manner), in our experience most of these things don't work. (Value, momentum and a few other strategies have in fact stood the test of time, but many others have not.) Taken too far, the behaviorist literature may be potentially harmful in that it encourages the idea that beating the market is easy, and its stories are readily adaptable to almost any empirical finding. Obviously, the flexibility of behavioral finance is both its strength and its weakness.

So, going the other way, are we proponents of efficient markets? Generally, yes, at least as the base case. We believe the concept of efficient markets is a healthier and more

correct beginning point for thinking about markets and investing. But having said that, we don't fully buy some of the arguments that the defenders of efficient markets sometimes trot out, as we've detailed above. And, as Fama himself says, we don't believe markets are perfectly efficient, and there's room for some factors (for example, part of the value return and probably much of the momentum return) to survive and thrive in the limited amount of inefficiency out there.

As we stated early on, risk versus behavioral schizophrenia describes us well. It's fair to say that some major bubbles have pushed us down the spectrum toward the behaviorist view. It's also fair to say that some of the micro anomalies have pushed us the same way, but maybe less so (momentum more than value). But although it's not a necessary condition of inefficient markets that markets be easily beatable, we still believe that if markets were gigantically, obviously and often inefficient, people could come in and take advantage of all these inefficiencies in a far easier manner than seems to play out in real life. Our experience suggests you can do it (over the long haul), but it ages you rapidly. (Cliff has been told he has the spleen and Golgi apparatus of a 75-year-old coal miner.) If we're schizophrenic on this issue, we are at least consciously so, and it's because we believe the middle is the closest to the right answer.

So if markets are not perfectly efficient but not grossly inefficient either—though occasionally pretty darn wacky—what should investors do? We believe the vast majority would be better off acting like the market was perfectly efficient than acting like it was easily beatable. Active management is hard.

That's not to say we think it's impossible. Take, for instance, our favorite example, briefly mentioned earlier, of people who seem to be able to consistently beat the market: Renaissance Technologies. It's really hard to reconcile their results long-term with market efficiency (and any reasonable equilibrium model). But here's how it's still pretty efficient to us: We're not allowed to invest with them (don't gloat; you're not either). They invest only their own money. In fact, in our years of managing money, it seems like whenever we have found instances of individuals or firms that seem to have something so special (you never really know for sure, of course), the more certain we are that they are on to something, the more likely it is that either they are not taking money or they take out so much in either compensation or fees that investors are left with what seems like a pretty normal expected rate of return. (Any abnormally wonderful rate of return for risk can be rendered normal or worse with a sufficiently high fee.)

Does this mean we should all go to Vanguard Group, buy their index funds and be done forever? While not at all a bad idea, we wouldn't go quite that far. For instance—another self-serving alert—we vote with our feet (and wallets) on this every day. Many of our own investments are based on strategies rooted in the academic work of testing EMH. Again, these strategies, like value and momentum (and others), can be interpreted as working over time because they are taking advantage of behavioral biases or they are compensation for bearing different types of risk. If an investor starts with a portfolio that is dominated by equity market risk, as most are, we believe that adding these strategies makes sense. You don't have to take a stand on whether markets are efficient. If you believe markets are inefficient, obviously you want to take advantage of these. If you believe markets are efficient and these strategies work because they are compensating you for taking risk, you still should want to own some of them (unless you fear that risk more than the average).

In our experience, actually running these strategies can be a bit trickier than what you see in the academic literature. Implementation details matter. Take value as an example. Does the measurement of value end at book-to-price ratios? In our research we find that there are many things you can do to (mildly) improve on a sole reliance on the academic version of book-to-price ratios. Does that mean we are moving away from efficient markets to being inefficiency guys trying to come up with some secret sauce to add value without risk? Not necessarily. It might simply be that in real life there is a value risk factor, but simple academic book-to-price isn't the best or only way to measure it. (We know of no theory that argues that book-to-price is perfect.) By improving on your signals, you may get a cleaner read on the underlying risk factor.

Also, it is most certainly the case that with sloppy trading you can easily throw away any expected return premium—whatever its source—that might exist around these strategies by paying too much to execute them (and sloppy can include overpaying in a slavish, high-turnover attempt to own precisely the portfolios from the academic papers). Clearly, the line between active and passive management starts to blur with these types of investment strategies.

What does this mean the government, including quasigovernment and self-regulatory institutions, should do? If we accept that markets are not perfect, then let's help them be as good as they can be. If they are perfectly efficient, then things like good versus bad accounting rules, or any rules for that matter, aren't important, as the market will always figure it out. But again, perfect efficiency is a chimera nobody believes in. However, if they are mostly close to efficient but not perfectly efficient (and occasionally perhaps even crazy), then everything matters to some degree. So here's an admittedly incomplete list of suggestions:

- **The government should recognize that bubbles can happen.** However, there are two important issues to consider. One, officials should recognize the difficulty in identifying bubbles, and, two, they should recognize the potential harm in acting on them wrongly or way too early (remember, Shiller was about half a decade too early). Unfortunately, most of us have fairly weak powers to identify bubbles as they are going on—identifying them after they have popped is a lot easier—and it is our belief that even the existence of bubbles does not for one second mean that a government panel will have any success in identifying them and, more important, acting at the right time. Central planning still runs face-first into Austrian economist Friedrich Hayek's fatal conceit. In addition, fostering a belief that someone is out there diligently preventing all bubbles can have the paradoxical effect of making bubbles they don't catch and expertly prick far more dangerous.
- **The government should not subsidize or penalize some activities over others.** These actions classically induce all kinds of unintended consequences and distortions. The most glaring example is government subsidies' contributing to the recent housing bubble (though we think this is a different question from whether government or business helped convert the housing bubble to a financial crisis).
- **The government should not promise to eliminate the downside.** "Too big to fail" is an efficient market's enemy. Admittedly, this advice is far easier given than taken, but recognizing this fact is quite important. Markets may be close to efficient if left alone, but markets with the downside banned are hamstrung and have little hope of being efficient.

- **The government should encourage disciplining mechanisms like short-selling (and conversely, it shouldn't ban or penalize them).** Markets should have the chance to reflect all information, not just positive or optimistic information. Short-selling is rarely popular, but its free and unfettered activity makes us safer. Discouraging, penalizing or banning short-selling is "too big to fail" applied at the micro level (too micro to fail?).
- **The government should encourage, not tax, liquidity provision.** The way prices get "fixed" generally involves someone trading. Poor liquidity makes this difficult. Obviously, more liquidity means lower costs to reflecting information in market prices. This is simply better for everyone. Some attribute bubbles to too much liquidity (we refer to trading liquidity, not the money supply, here) and too much trading. That is hard to believe. Bubbles—to the extent that we are right, and they are rare but real—come from people believing they are going to make ten times their money, not trade the next share cheaply. On the other hand, systematic diversified traders who may be willing to trade against bubbles are in dire need of reasonably priced liquidity, as they, if they aren't crazy, run very diversified portfolios with real but narrow spreads (like the returns to value investing) and transactions costs—the cost of liquidity—matter a lot. We want them running these portfolios.
- **The government should punish true fraud harshly.** However, we should also recognize that regulating to create an all-but-fraud-free world is too costly and getting all the way there is impossible.
- **The government should have consistent laws consistently applied (for example, when it comes to bankruptcy).** If markets are not perfect, we must help them, and arbitrary rules, and ill-defined property rights that change through time, are among the easiest ways to hinder rather than help.
- **The government and self-regulatory bodies should encourage consistent and reasonable accounting.** Once we give up on perfect efficiency, we recognize we're at some point on the market efficiency spectrum. That is, markets may be mostly but not entirely efficient. Giving up perfect efficiency, you can no longer argue, "Who cares, the market will see through it" as an excuse not to have reasonable accounting rules (as some EMH proponents did after the tech bubble in regard to expensive executive stock options).
- **The government should encourage that financial institutions mark more things to market.** Some argue that "if you marked to market, no bank would survive." In that case, change the capital rules around survival, but don't disseminate false information by using prices you know are wrong or stale. If we had to prioritize, "too big to fail" and not marking things to market are our personal two choices for the original sins behind the financial contagion in 2007–'08 that followed the real estate bubble.

Last, there is a really bad notion that we have heard people talk about that drives us crazy. The notion is that the 2007–08 credit and real estate bubble and ensuing financial crisis, and perhaps other bubbles, were caused by a belief in market efficiency, or "market fundamentalism." If there are bubbles—and we believe there are but that they are rare—they are likely caused by people who think they are getting an extraordinarily impossibly good deal, not a fair deal in an efficient market. That is, we think bubbles are driven by believers in highly inefficient markets. No speculator ever created a bubble by buying something he or she thought was simply a fair deal in an efficient market. This was certainly what it felt like during the dot-com bubble and the recent housing bubble. Many

people were thinking their dot-com stocks or three vacation homes would continue to soar, not a little but a lot. They were not thinking about Fama's research and investing with Vanguard founder John Bogle. At their core bubbles seemingly are caused by an intense belief in the hypothesis that markets are ridiculously inefficient, not the opposite. To say it's believers in efficient markets that cause bubbles is simply a political slur—and a backward one at that.

The broad point is that we believe markets are wonderful. They're the best system for allocating resources and the spread of freedom and prosperity that the world has ever seen. But they are not magic. As we, and again even Gene Fama, have said many times, they are not perfectly efficient. How efficient they are is partly a function of the care and thought we put into designing them and the rules around them. Many of the actions we collectively take actually hamstring markets, making them less efficient, and then the cry invariably goes out: See, blame the believers in markets! That needs to change.

At the end of the day, we think the Nobel committee did fine splitting the baby that is the prize in economic sciences. EMH has contributed more to our understanding of finance and even general economics than any other single idea we can think of in the past 50 years. One way to assess the impact of this idea is to ask whether we know more as a result of the introduction and testing of, and the debate about, the Efficient Market Hypothesis. Most certainly, the answer is an ear-splitting yes. As such, Fama's introduction of this hypothesis and his active (incredibly active) study of it all this time make him our clear pick as the MVP of modern finance and perhaps economics as a whole for the past almost half century. Shiller, as a major EMH gadfly, has also earned his place on this shared podium, as his is a significant and important case against EMH and Shiller has led that charge admirably. The study of EMH has made our thinking far more precise. (Again, we do not mean to understate Hansen's contribution to the analysis of asset prices. More than its mathematical nature, it's simply on a different spectrum from the EMH debate we focus on here.)

Moreover, the impact of the Efficient Market Hypothesis has gone well beyond academia. It's hard to remember what finance was like before EMH, but it was not a science; it was barely even abstract art. Markets might not be perfect, but before EMH they were thought to be wildly inefficient. It was assumed that a smart corporate treasurer added lots of value by carefully choosing among debt and equity for his capital structure. It was assumed that a diligent, hardworking portfolio manager could beat the market. Anything else was un-American! At a minimum, index funds and the general focus on cost and diversification are perhaps the most direct practical result of EMH thinking, and we'd argue the most investor-welfare-enhancing financial innovation of the past 50 years. Not bad.

So where does that leave us as students of Professor Fama and practitioners for the past 20 years of much of what he taught us? Simply put, we'd have nothing without EMH. It is our North Star even if we often or always veer 15 degrees left or right of it. But despite this incredible importance, the idea that markets are literally perfect is extreme and silly, and thankfully (at least for us), there's plenty of room to prosper in the middle. Apparently, the Nobel committee agrees.

My Top 10 Peeves

Clifford S. Asness

The author discusses a list of peeves that share three characteristics: (1) They are about investing or finance in general, (2) they are about beliefs that are very commonly held and often repeated, and (3) they are wrong or misleading and they hurt investors.

Saying I have a pet peeve, or some pet peeves, just doesn't do it. I have a menagerie of peeves, a veritable zoo of them. Luckily for readers, I will restrict this editorial to only those related to investing (you do not want to see the more inclusive list) and to only a mere 10 at that. The following are things said or done in our industry or said about our industry that have bugged me for years. Because of the machine-gun nature of this piece, these are mostly teasers. I don't go into all the arguments for my points, and I blatantly ignore counterpoints (to which I assert without evidence that I have counter-counterpoints). Some of these are simple, so perhaps the teaser suffices. But some deserve a more thorough treatment that hopefully I, or someone else, will undertake. Some are minor, truly deserving the title "peeve," and some, more weighty. In each case, as befits an opinion piece, it's not just my discussion of the peeve but the very prevalence of the peeve itself that is my opinion. I do not extensively cite sources for them. I contend that they are rather widespread throughout the land of financial media, pundits, advisers, and managers. Thus, citing one or two sources would be unfair, and citing them all, impossible. Therefore, please feel free to disagree not just with my discussion of the peeves but also about their very existence! Without further ado, here is a list of things held together by only three characteristics: (1) They are about investing or finance in general, (2) I believe they are commonly held and often repeated beliefs, and (3) I think they are wrong or misleading and they hurt investors.

Reprinted from *Financial Analysts Journal*, January/February 2014, Vol. 70 (1), 22–30, with permission.

1. "Volatility" Is for Misguided Geeks; Risk Is Really the Chance of a "Permanent Loss of Capital"

There are many who say that such "quant" measures as volatility are flawed and that the real definition of risk is the chance of losing money that you won't get back (a permanent loss of capital). This comment bugs me.

Now, although it causes me grief, the people who say it are often quite smart and successful, and I respect many of them. Furthermore, they are not directly wrong. One fair way to think of risk is indeed the chance of a permanent loss of capital. But there are other fair methods too, and the volatility measures being impugned are often misunderstood, with those attacking them setting up an easy-to-knock-down "straw geek."

The critics are usually envisioning an overvalued security (which, of course, they assume they know is overvalued with certainty) that possesses a low volatility. They think quants are naive for calling a security like this "low risk" because it's likely to fall over time. And how can something that is expected to fall over time—and not bounce back—be low risk?

What we have here is a failure to communicate. A quant calling something "low risk" is very different from a quant saying, "You can't lose much money owning this thing." Even the simplest quant framework allows for not just volatility but also expected return.[1] And volatility isn't how much the security is likely to move; it's how much it's likely to move versus the forecast of expected return. In other words, after making a forecast, it's a reflection of the amount you can be wrong on the upside or downside around that forecast. Assuming the quant and non-quant agree that the security is overvalued (if they don't agree, then that is an issue separate from the definition of risk), the quant has likely assigned it a negative expected return. In other words, both the quant and the non-quant dislike this security. The quant just expresses his dislike with the words "negative expected return" and not the words "very risky."

A clean example is how both types of analysts would respond to a rise in price unaccompanied by any change in fundamentals now or in the future. On the one hand, those who view risk as "the chance of permanent loss" think this stock just got riskier. Viewed in their framework, they are right. On the other hand, quants tend to say this stock's long-term expected return just got lower (same future cash flows, higher price today) rather than its risk/volatility went up, and they too are right!

It is also edifying to go the other way: Think about a super-cheap security, with a low risk of permanent loss of capital to a long-term holder, that gets a lot cheaper after being purchased. I—and everyone else who has invested for a living for long enough—have experienced this fun event. If the fundamentals have not changed and you believe risk is just the chance of a permanent loss of capital, all that happened was your super-cheap security got superduper cheap, and if you just hold it long enough, you will be fine. Perhaps this is true. However, I do not think you are allowed to report "unchanged" to your clients in this situation. For one thing, even if you are right, someone else now has the opportunity to buy it at an even lower price than you did. In a very real sense, you lost money; you just expect to make it back, as can anyone who buys the same stock now without suffering your losses to date.

If you can hold the position, you may be correct (a chance that can approach a certainty in some instances if not ruined by those pesky "limits of arbitrage").[2] For example, when my firm lost money in 1999 by shorting tech stocks about a year too early (don't worry;

it turned out OK), we didn't get to report to our clients, "We have not lost any of your money. It's in a bank we call 'short NASDAQ.'" Rather, we said something like, "Here are the losses, and here's why it's a great bet going forward." This admission and reasoning is more in the spirit of "risk as volatility" than "risk as the chance of a permanent loss of capital," and I argue it is more accurate. Putting it yet one more way, risk is the chance you are wrong. Saying that your risk control is to buy cheap stocks and hold them, as many who make the original criticism do, is another way of saying that your risk control is not being wrong. That's nice work if you can get it. Trying not to be wrong is great and something we all strive for, but it's not risk control. Risk control is limiting how bad it could be if you are wrong. In other words, it's about how widely reality may differ from your forecast. That sounds a lot like the quants' "volatility" to me.

Although I clearly favor the quant approach of considering expected return and risk separately, I still think this argument is mostly a case of smart people talking in different languages and not disagreeing as much as it sometimes seems.[3]

2. Bubbles, Bubbles, Everywhere, but Not All Pop or Sink

The word "bubble," even if you are not an efficient market fan (if you are, it should never be uttered outside the tub), is very overused. I stake out a middle ground between pure efficient markets, where the word is verboten, and the common overuse of the word that is my peeve. Whether a particular instance is a bubble will never be objective; we will always have disagreement *ex ante* and even *ex post*. But to have content, the term *bubble* should indicate a price that no reasonable future outcome can justify. I believe that tech stocks in early 2000 fit this description. I don't think there were assumptions—short of them owning the GDP of the Earth—that justified their valuations. However, in the wake of 1999–2000 and 2007–2008 and with the prevalence of the use of the word "bubble" to describe these two instances, we have dumbed the word down and now use it too much. An asset or a security is often declared to be in a bubble when it is more accurate to describe it as "expensive" or possessing a "lower than normal expected return." The descriptions "lower than normal expected return" and "bubble" are not the same thing.

As a current example, take US government bonds. They are without a doubt priced to offer a lower prospective real return now than at most times in the past (as, in my view, are equities). But could it work out? With an unchanged yield curve, which is certainly possible, you would make a very comfortable 4%+ nominal (call it 1%–2% real) a year now on a 10-year US bond, and to find a case where bonds worked out from similar levels, one only has to utter the word "Japan." Does this make bonds a particularly good investment right now? No. Does it show that they do not satisfy the criteria for the word *bubble*, thereby demonstrating how the word is overused? Yes.

3. Had We but World Enough, and Time, Using Three- to Five-Year Evaluation Periods Would Still Be a Crime

Nobody, including me in this essay, wants to deal with the very big problem that we often do not have enough applicable data for the investing decisions we make. We evaluate

strategies, asset classes, managers, and potential risk events using histories the statisticians tell us are too short or too picked over. These histories are generally insufficient and very vulnerable to such things as data mining, *ex post* selection of winners who don't repeat (though it's generally churlish to be horribly disappointed when your monkey who typed *Hamlet* produces only *Coriolanus* next time), and simple randomness.

Too often, we default to thinking like, "We have to make decisions, and even if historical data are inadequate, you have nothing better to offer, so we'll use what we have." I think there is something better. Investors should elevate judgment (not minute-by-minute judgment but judgment in portfolio and strategy selection) and a consistent philosophy to be more equal partners with data.[4]

But all of these issues are subject for a much longer piece. Here, I set my sights (and peeve) on easier game. Not only are insufficient data often driving our decisions, but the data we have are often used with the wrong sign. I refer to the three- to five-year periods most common in making asset class, strategy, and manager selection decisions. One of the few things we do know is that over three to five years, pretty much everything has shown some systematic, if certainly not dramatic, tendency to mean revert (especially when one accounts for and avoids the powerful effect of momentum at shorter horizons). This means that when we rely on three- to five-year periods to make decisions—favoring things that have done well over this time period and shunning things that have done poorly (note the past tense)—we aren't just using data meaninglessly; rather, we are using data backwards. Essentially, with a disciplined approach, value and momentum are both good long-term strategies, but you don't want to be a momentum investor at a value time horizon. That's precisely what many of us who use three- to five-year periods end up being.

4. Whodunit?

The war continues over whether the 2007–08 financial crisis was caused by government or by the big banks. This debate is another clear candidate for a far longer exploration, but a few peeves jump out. Those who have read some of my other work will, I hope, be pleasantly surprised at the nonpartisan nature of my ire here. My first peeve is with the idea that we will ever find, or should find, one real culprit. When any bubble bursts (yes, I remember Peeve 2 and still choose the word "bubble") and a global financial meltdown follows, nearly everyone shares some blame. You can't get a good bubble off the ground without government, the financial industry (including not just bankers, who are often the named party, but the whole real estate industry, the entire mortgage finance industry, rating agencies, and others), and regular individuals (nobody wants to lay any blame on Main Street; where is the political hay in that?) acting stupid and short-sighted (e.g., quitting real jobs to flip houses).

My second, and more important, peeve regarding this issue is that the typical narratives and debates conflate two events. We had (1) a real estate/credit bubble in prices that, upon bursting, precipitated (2) a massive financial crisis. The collapse of the real estate/credit bubble did not have to lead directly to the financial crisis. See the tech bubble's deflation for a clear counterexample.[5] The questions of who should shoulder what share of the blame for the real estate bubble and who should shoulder what share of the blame for the financial crisis do not necessarily lead to the same answer. There are basically two sides here: What

you might call the progressive "left" wants to blame Wall Street, and what you might call the libertarian "right" wants to blame government. Both sides act like they're arguing with each other even though the blame-the-government side is mostly talking about the creation of the real estate bubble and the blame-the-banks side is mostly talking about the financial crisis. They are both being, perhaps, rational—if not always honest—debaters, and without picking a winner, I'd say they are both focusing on what's best for their argument. But much of the time, they are not really debating each other. Although perhaps it is ultimately hopeless because of inherent difficulty and all our personal biases, until we treat these as two separate, deeply related,[6] but sequential events to which all major parties contributed to some degree, we haven't a prayer of really understanding what happened and making serious headway on reducing the risk of it happening again (which, by the way, I don't think we've done much of—a peeve for another day).

5. I Would Politely Request People Stop Saying These Things

"It's a stock picker's market." I don't know what it means to say, "It's a stock picker's market." It may mean the whole market isn't going straight up now so you have to make your money picking the right stocks, but I don't understand why active managers would suddenly get better at stock picking at those times. Note that I do think a valid use of this concept may occur when, after adjusting for market moves, there is not a lot of dispersion in stock returns, meaning that individual stocks tend to move in lockstep, leaving little idiosyncratic volatility—a necessary (but not sufficient!) ingredient to generate outperformance (assuming one is unwilling to lever up smaller differences at these times). But that's a quant measure, and I don't think that's what many people mean by this comment. I think they mean, "We will have to pick stocks now because the market isn't making us money the easy way." To the extent I'm wrong, I withdraw the peeve (is there a specific form I need to file for that?).

Similarly, you often hear financial professionals say such things as, "Forecasting market direction from here is exceptionally difficult" in a tone conveying, "Gee, this is really strange." Well, I think forecasting the market over short-term horizons is always exceptionally difficult. If they said, "Our market-timing forecasts are mostly useless most of the time, but right now, they are completely useless," I suppose I'd be OK with it, but I'm not holding my breath that they will.

"Arbitrage." The word "arbitrage" in academia means "certain profits," whereas in practical investing, arbitrage often means "a trade we kind of like." Some in the industry adhere to a perhaps reasonable middle ground: that arbitrage is not riskless, but unlike much of investing, it involves going long and short very similar securities and betting on a price difference. I can live with that. But it is clear that many use it in the loosest sense and, therefore, strip it of its meaning.

"There is a lot of cash on the sidelines." Every time someone says, "There is a lot of cash on the sidelines," a tiny part of my soul dies. There are no sidelines. Those saying this seem to envision a seller of stocks moving her money to cash and awaiting a chance to return. But they always ignore that this seller sold to somebody, who presumably moved a precisely equal amount of cash off the sidelines.

If you want to save those who say this, I can think of two ways. First, they really just mean that sentiment is negative but people are waiting to buy. If sentiment turns, it won't move any cash off the sidelines because, again, that just can't happen, but it can mean prices will rise because more people will be trying to get off the nonexistent sidelines than on. Second, over the long term, there really are sidelines in the sense that new shares can be created or destroyed (net issuance), and that may well be a function of investor sentiment.

But even though I've thrown people who use this phrase a lifeline, I believe that they really do think there are sidelines. There aren't. Like any equilibrium concept (a powerful way of thinking that is amazingly underused), there can be a sideline for any subset of investors, but someone else has to be doing the opposite. Add us all up and there are no sidelines.[7]

6. The First Step Is Admitting It

To me, if you deviate markedly from capitalization weights, you are, by definition, an active manager making bets.[8] Many fight this label. They call their deviations from market capitalization—among other labels—smart beta, scientific investing, fundamental indexing, or risk parity.[9] Furthermore, sometimes they make distinctions about active versus passive based on why they believe in their strategies. You can believe your strategy works because you're taking extra risk or because others make mistakes, but if it deviates from cap weighting, you don't get to call it "passive" and, in turn, disparage "active" investing. This peeve may be about form over substance—marketing versus reality—but these things count.

In particular, some of the discussion these days about "smart beta" refers to it as "a better way to get market exposure." It's not. It may indeed offer added return, and many who offer smart beta do it with value tilts (which I like). But whether it is profitable and a good idea is an issue separate from whether it is active or passive. Calling it "a better way to get market exposure" frames it incorrectly. A smart beta portfolio, SB, is equal to a cap-weighted index, CW, plus the deviation of SB from CW. For those who like really simple equations: SB = CW + (SB – CW). The expression (SB – CW) indicates a kind of simple long–short portfolio (for smart beta, it's usually designed so that the net SB is not really short anything) representing a bet. Note that this is not unique to today's smart beta discussion. For instance, the justly famous Fama–French HML portfolio is simply a long–short portfolio.[10] If you add it to a cap-weighted index, you don't get "better market exposure"; you get market exposure plus a separate bet on value investing. Like HML, the tilt from CW to SB also may be a good bet if offered at a fair fee (again, I love some of these strategies as if they were my own).[11] And, relating this to the topic of fees, of course they matter a lot. By my definition, an active portfolio does not have to have high fees (the fair fee depends on the size of the deviation from cap weighting, the expected risk-adjusted performance per unit of deviation, and the uniqueness of the active tilt). But again, a tilt, even if fairly priced, is still an active bet.[12]

I think people should call a bet a bet. If you own something very different from the market, you're making a bet and someone else is making the opposite bet. You might

believe in your bet because you are being compensated for taking a risk, because the market has behavioral biases, or because your research is just that good. Your bet might be low or high turnover. But, regardless, you aren't passive.[13]

7. To Hedge or Not to Hedge?

There has been a flurry of discussion regarding hedge fund performance these days, with the highlight (or low light) being a rather priapic cover of *Bloomberg Businessweek* on the topic. But much of the discussion of hedge fund returns is just not cogent. Here, I feel a little like a brother to other hedge fund managers (not in the true fraternal sense but in the sense of "Nobody picks on my brother but me").[14] When other people overstate the virtues of hedge funds, I become a critic, as my firm has been quite a few times in the past. But when other people attack them unfairly, I defend them.

The big disconnect here is that hedge funds are not fully hedged vehicles (I have long lobbied for hedge funds that fully hedge—see Asness, Krail, and Liew 2001—but it's still not the reality on the ground). But they are also not fully long the equity market. Many hedge funds have averaged about 40%–50% equity exposure over the long term, although that number certainly has varied through time.[15] So, in some years, like the current one, the press runs story after story about how hedge funds are being trounced by long-only stock indices (as an aside, out of simple familiarity, the press also focuses on the US stock market, which is trouncing the world this year, even though hedge funds tend to be more global, making the comparison seem more dire). In other years, when markets are down, commentators can oddly be overly generous to or overly critical of hedge funds. Sometimes, they ignore all evidence that hedge funds are net long stock markets and can get shocked and dismayed that "so-called uncorrelated" hedge funds are down at all with the market, even if they are down substantially less. In contrast, they may laud hedge funds as heroes because they are down less than the market (because, although not fully hedged, the funds are rarely fully long the equity market either). On this most basic issue of market exposure, popular reports are almost always breathlessly, excitedly misleading in one direction or another—be it scorn or praise. It's admittedly a hard thing to deal with because partial exposure is a more complicated story than having zero exposure or being fully invested (furthermore, as a group, hedge funds' exposure can vary over time, and this exposure can vary tremendously across individual funds). But the solution cannot be to ignore the problem and always purport to have found something extreme.

More generally, in my opinion, there are great problems and great promise in the hedge fund world. Hedge funds carry out strategies that are quite valuable (producing return with low correlation with traditional markets), such as merger and convertible arbitrage, derivatives-based trend following, macro trading based on value and carry that are not available in other formats, and even a level of bespoke stock picking I claim no expertise in evaluating but that certainly may be valuable (particularly in a stock picker's market). The problem is that as a group, they do these things without hedging enough and, worse, they charge fees—especially performance fees—as if they were providing purely uncorrelated returns (this could be fixed by charging lower fees or performance fees based on a comparison with their expected passive net long position) on strategies that they often claim are unique to them (many are quite well known).

But I digress. My peeve is that hedge fund reporting, by both the media and industry, is almost always wrong, but in a fascinatingly varied kind of way depending on market direction and the inclination of the commentator. My hope is that this reporting and the whole hedge fund industry can improve and become a better value proposition for investors.[16]

8. I Know Why the Sage Nerd Pings

This peeve is about the rolling brouhaha over "high-frequency trading" (HFT); I believe it's massively overwrought. HFT is mostly a good thing, not an evil conspiracy to crush Main Street. I should mention that my firm is quantitative and algorithmic but is not even close to being a high-frequency trader; we mostly have long-term views and tend to hold our positions for quite a while. But being quantitative and algorithmic is frequently, and incorrectly, mistaken for being high frequency.

I believe that most asset managers trade cheaper, and their investors are thus better off, because highly efficient high-frequency traders have largely replaced the traditional high-priced market making of the past. HFT is how modern market making in a technologically advanced world is done, and there is no going back (which doesn't preclude continued discussions about regulatory tweaks; I'm not vouching for every individual high-frequency trader or every practice of their trading).[17] It's cheaper because it smashed the old dealer and exchange cartel (providing a much lower barrier to entry for competing market makers), democratized flow information, and replaced very expensive humans trading a handful of securities with very cheap machines trading a great many.

To be more specific, I believe small investors, whose trade quantities tend to be fully satisfied by the size of the inside market quote, are obvious winners because the bid–ask spread is now much tighter. But traders who try to trade large amounts very quickly—and often foolishly, in my view—still benefit from lower costs because of HFT, in spite of the protestations of some of them. This assertion is, admittedly, hard to prove or disprove, but I suspect that before HFT, big traders didn't see prices move as immediately (as they should in a more efficient market, which HFT has created). Although this initial lack of price movement may seem like a boon, you must recall that they were still facing a much larger bid–ask spread in those days. It is quite possible that they have mistaken that initially lower market impact with their all-in costs being lower in the old days (of course, in the old days, prices still usually moved substantially during the course of a large trade).

So, why the inflamed opposition? Well, some market participants who, before HFT, used to provide liquidity in more traditional ways are simply annoyed that their business has been taken by those who provide it more cheaply. These more traditional types generate many of the negative comments about HFT (some have based entire business models around rants about HFT coupled with more traditional and more expensive services).

Still other participants, such as some large asset managers, have long resisted adapting to the new market, not only refusing to "go electronic" but also ignoring the important fact that nearly everyone else already did. They are starting to change now, but for a long time, quite a few seemed to find it easier to complain than to modernize. I suspect that many speak negatively about HFT, even though it makes them better off, for the simple reason

that they don't understand it, and they are speaking to, and for, those who also don't understand it. Good anti-progress rants have always been popular. After all, until the 1700s, they still burned quant geeks at the stake.

A new and advanced technology always creates critics and predictors of all kinds of doom. I've heard HFT blamed for some bizarre things. I've heard it blamed for bubbles. How high-frequency traders who go home flat (close to no positions held) every day create bubbles is beyond me. I've heard it blamed for why some markets and strategies seem more correlated today than in the past (again, how this can be true is beyond me). I've heard some say HFT causes volatility (on net, you would expect the opposite from market makers), even though volatility in this age of HFT has mostly been very low except for the bursting of the decidedly low-frequency housing bubble (thankfully, I've not yet found anyone who believes high-frequency stock trading made people pay crazy amounts for housing). I've heard it blamed for hurting investor "confidence." That may be true but only because people are telling investors such silly things about it. It's rather circular logic to blame HFT for this fact.

Although it's not all they do, market making—that is, taking the other side of whoever is trading and making some fraction of a bid–ask spread for it—is indeed the core economic activity of HFT.[18] Once we see that market making is the core activity of HFT, we can also see that it is being compared with a mythical gentle giant of the past, the old-school market makers who allegedly often stopped crashes in their tracks by buying securities at prices they knew were way above current market prices. These heroic figures risked their own bankruptcy to save the financial world. Why do I call this a myth? Perhaps because it has happened zero times in financial history. It is not now and has never been the business of a market maker to go broke buying securities at the wrong price in a crashing market. Furthermore, no regulations or exchange rules are going to make them do so. In fact, it is in chaotic times that market makers have tended to make their biggest profits, hardly indicative of any noble "take one for the team" ethos during tough times.

By the way, of course there have been glitches, and some were quite scary. That will happen with any new technology, and it will happen more when the system is complicated and organically grown in separate places, which it most certainly is now. But these glitches have actually been more about electronic trading than HFT (again, a longer explanation is needed, but these are far from the same), and you would have to turn the clock much further back to eliminate electronic trading in addition to HFT. But you usually don't hear that from the critics and those in the press who find HFT a convenient villain for nearly all investing scares but who won't blame electronic trading in general because doing so would be a little too Luddite even for them. Of course, as with any new technology, a continued industrywide effort to reduce technological problems, make disparate systems work and play nicer together, and modify and streamline regulation (including market structure design) is appropriate and quite possibly overdue.

To add to my general peeve that HFT is extremely misunderstood and maligned, there is a misperception that HFT firms are making money hand over fist and are a serious drain on the investing economy. HFT firms, as a group, make money, but they make far less money in aggregate than most would guess (which befits the hypothesis that they offer the customer a better deal on liquidity) and certainly less than old-school market makers made in the past. So, this is a tempest in a teapot. And if the histrionic complainers (who are

protecting their own interests or trying to benefit from the hysteria they're creating) win, it can have some very bad consequences for markets and investors (for example, see the recent imposition by some European countries of financial transaction taxes, which have already proven to be failures).

Finally, two specific elements of HFT raise concerns among some: the frequent use of cancel-and-correct orders and the speed at which transactions occur (it's the only aspect of finance I know of where the speed of light matters; our field's physics envy is finally bearing fruit!). But if you view high-frequency traders as mainly (not entirely) market makers, it is easy to see the reason for both of these aspects of HFT. If you are putting your willingness to buy and sell out there, as any market maker does, you are the one in danger from market moves if you do not cancel and correct your orders fast enough. That is, you can be caught trading on the wrong side of an off-market price. Ironically, being able to use speedy cancel-and-correct orders to protect themselves against having to do such off-market trades allows high-frequency traders to, on average, provide investors with tighter bid–ask spreads than would otherwise be possible. Of course, unrelated to market making, when news comes out, being able to move slightly faster than others is a slight advantage (only slight because your trade size is still limited) over customers and other, slower high-frequency traders. Although I believe this issue is small compared with the larger issue of market making, it should be noted that somebody has to be first, and investing in being the first to be able to act on news and profiting from doing so is at least as old as the Rothschilds allegedly profiting from learning of the outcome of the Battle of Waterloo first and almost certainly far older. Over time, slower investors should learn to avoid trading at these precise junctures because they are at an informational disadvantage (again, this applies to whoever was slower throughout history; it is not unique to HFT) and perhaps should pull their orders out ahead of major news announcements. Another possible application of speed is to simply be the first to win the customer trades at the current best bid or offer without any tightening of the bid–ask spread. Ironically, the failure to tighten bid–ask spreads further is often the result of overly large tick sizes (the allowable increments of price change), which are more a function of regulation and futures exchange monopolies than any failing of HFT. "Trying to be the one to serve the customer" sounds much less nefarious than normal descriptions of HFT speed but is more accurate.

But if being on the other side of a scary high-frequency trader really freaks investors out, they can just own an index fund and sidestep quite nearly the whole thing ("quite nearly" because even index funds trade a little). If you are trading up a storm on your own, this is almost definitely good advice completely separate from the HFT issue.

9. Antediluvian Dilution Deception and the Still-Lying Liars[19]

Companies with executives who execute stock options still carry out buybacks to "prevent dilution." This is still idiocy. It may be time-honored idiocy, but it is idiocy nonetheless. The only rational reason for a company to carry out a buyback is that management believes its shares are undervalued.[20] To do so just to be able to tell shareholders, "See, you own as much of the company as you did before we handed out those

stock options" is just nonsensical. It's wrong not on the math but on the relevance. Given the right amount of buyback, the shareholders can indeed end up owning as much of the company as before. So what? They now own as much of a company that happens to have a bunch less cash (i.e., the cash just used to buy back shares). Having a bunch less cash is worse! If the shares were overvalued, in fact, this act of financial camouflage directly hurts the very shareholders management is trying to assuage. Please don't take this too far: Issuing executive options as part of compensation may still be in the company's and other shareholders' best interest. But buying back shares to prevent dilution when they are exercised is a cosmetic silliness designed, in my view, to obscure the fact that option issuance is costly.

On a related note, the forces of good won the battle to expense executive stock options about a decade ago, yet many firms—abetted by some Wall Street analysts who apparently remember 1999–2000 with fondness instead of horror or perhaps remember it only as the year their braces came off—still report pro forma earnings before the necessary and legally mandated act of expensing them and somehow persuade people to use these silly numbers. It's amazing how hard it is to kill a scam even after you make it illegal to use it on the front page.

10. Bonds Have Prices Too (How Do You Think We Price Those Bond Funds?)

This misconception is perhaps the least harmful— in fact, it may even be helpful to investors—but is perhaps the most annoying to me (I'm not sure why; maybe it's because I've been hearing it the longest, given that I started in fixed income). Many advisers and investors say things like, "You should own bonds directly, not bond funds, because bond funds can fall in value but you can always hold a bond to maturity and get your money back." Let me try to be polite: Those who say this belong in one of Dante's circles at about three and a half (between gluttony and greed).[21]

Bond funds are just portfolios of bonds marked to market every day. How can they be worse than the sum of what they own? The option to hold a bond to maturity and "get your money back" (let's assume no default risk, you know, like we used to assume for US government bonds) is, apparently, greatly valued by many but is in reality valueless. The day interest rates go up, individual bonds fall in value just like the bond fund. By holding the bonds to maturity, you will indeed get your principal back, but in an environment with higher interest rates and inflation, those same nominal dollars will be worth less. The excitement about getting your nominal dollars back eludes me.

But getting your dollars back at maturity isn't even the real issue. Individual bond prices are published in the same newspapers that publish bond fund prices, although many don't seem to know that. If you own the bond fund that fell in value, you can sell it right after the fall and still buy the portfolio of individual bonds some say you should have owned to begin with (which, again, also fell in value!). Then, if you really want, you can still hold these individual bonds to maturity and get your irrelevant nominal dollars back. It's just the same thing.

Those believing in the subject fallacy often also assert that another negative feature of bond funds is that "they never mature" whereas individual bonds do. That's true. I'm

not sure why anyone would care, but it's true. But the real irony is that it's only true for individual bonds—not the actual individual bond portfolio these same investors usually own. Investors in individual bonds typically reinvest the proceeds of maturing bonds in new long-term bonds (often through the use of a "laddered portfolio"). In other words, their portfolio of individual bonds, each of which individually has the wonderful property of eventually maturing, never itself matures. Again, this is precisely like the bond funds that they believe they must avoid at all costs.

I'm sorry if I've destroyed the peace of mind of individual bondholders everywhere by informing them that owning only individual bonds does not solve the problem that bonds are risky. I'm also sorry if the irrelevant idea that you'll eventually get your nominal money back on a bond was comforting to many. It is actually quite possible that I have made some readers worse off by destroying these illusions. It's possible that the false belief that individual bonds don't change in price each day like a bond fund's net asset value does led to better, more patient investor behavior. I admit that listening to me is not always a pleasant or even wealth-enhancing experience.

Conclusion

I am not so arrogant that I dismiss the idea that I also have some crazy notions that might make another's list of peeves and that I could benefit from reading it. Also, although I will admit to nothing, at times, I have certainly made some of the mistakes I have discussed, and whatever lessons I've learned have often come from experience—not before-the-fact superior reasoning.

Finally, I certainly have some peeves I haven't shared, but I must stop here lest anyone think I'm a curmudgeon.

I'd like to thank Jonathan Beinner, Aaron Brown, Tom Dunn, Antti Ilmanen, Ronen Israel, John Liew, Michael Mendelson, and Hitesh Mittal for very helpful comments.

Endnotes

[1]This really is the very simplest framework. Although this section is, to a large extent, a defense of the concept and use of volatility, simple well-behaved symmetric volatility measures are not the only weapon in the quantitative risk management toolkit by a long shot. For instance, another common criticism of measuring volatility is that it treats the upside and the downside similarly, but there are quantitative measures that deal with this distinction. For instance, volatility would be particularly inappropriate for option-like securities with by-design asymmetric payoffs.

[2]See Shleifer and Vishny (1997) for the situation where you are metaphysically certain you are right but you can still be practically wrong if you are forced out of a position, if you panic, if your client panics, and so on.

[3]Studying this issue further is definitely worthwhile because many other issues arise when examining it more deeply, including the following: (1) risk being about individual investments versus being about portfolios (preview: although it can still be expressed with expected return, I have more sympathy for obsessing over the left tail of each security in a very concentrated portfolio versus a very diverse one); (2) rebalancing (preview: unless things jump to zero, how you act along the way determines much of how harmful a particular left tail can be); (3) the idea that some risk may

be very short term and mean reverting, which inflates "return volatility," but perhaps can be traded against (by quants or non-quants); (4) geometric versus arithmetic returns; and (5) dynamic versus static portfolio choice.

[4]For a related editorial, see Sullivan (2010).

[5]Some argue that debt bubbles are more dangerous than equity bubbles or that the tech bubble was a smaller segment of the economy, but these are some of the counter-points I mentioned in the beginning that I'm leaving for another day!

[6]For instance, the regulatory and rating agency blind spot in AAA securities probably contributed to the bubble's growth and the ensuing financial crisis, so I'm not saying some of the same explanations do not apply to both. But these two events were still separate, and many other much-argued-over potential explanations apply to only one or the other.

[7]See Sharpe (1991, 2010) for similar points regarding this issue and investment outperformance.

[8]There are weaker definitions of active versus passive than mine. For instance, one definition is that anyone who follows an index is passive (so any definable strategy, no matter what the fee or turnover, is passive as long as you can write it in the form of an index, which really amounts to calling any quantitative strategy passive, which I think is quite odd). Another is that passive just means low turnover. If people are using this definition, I partially withdraw my peeve. But I would replace this section with a peeve about weak straw-man definitions that don't mean much because I think low turnover is a very weak definition of passive. For instance, to me, possessing low turnover but maintaining a huge difference from cap weighting, such as owning a very concentrated portfolio, is a very active—not passive—strategy that simply doesn't call for changes very often. In other words, to me, passive is about being cap weighted or close to it, not about how many calories you burn while investing.

[9]In the interest of full disclosure, my own firm runs investment products like some of these and occasionally uses these labels.

[10]HML, originally from Fama and French (1993), stands for "high minus low" and refers to going long a diversified portfolio of high-book-to-price (low-price-to-book) stocks and shorting a diversified portfolio of the opposite. In other words, it's a diversified factor that represents the return spread of "cheap" or "risky" (depending on your views on market efficiency) value stocks over "expensive" or "low-risk" growth stocks.

[11]Disclosure: Some indeed are mine or, more precisely, AQR Capital Management's.

[12]Some who criticize "active management" are really criticizing those who pick individual stocks who often use nonsystematic means (traditional stock pickers). Without commenting on this criticism, I would call such a style a particular type of active management, with active still being a far more general category that, again, includes any nontrivial deviation from the cap-weighted market.

[13]And, as a side point, if that bet does the same thing others have been doing for years, you don't get to call it "new" because you have a new name for it and it's in a new package.

[14]Disclosure: Most of AQR's assets under management are not hedge funds, but the firm does run a significant amount of hedge fund assets.

[15]Consider rolling seven-year versions of the betas of equity long–short hedge funds, as calculated in Asness et al. (2001), who accounted for both contemporaneous and lagged market exposure: The median seven-year period has a beta with the S&P 500 Index of 0.47, and the most recent seven-year period is at 0.44.

[16]Of course, these things are connected because if hedge funds ran their portfolios fully hedged, the commentators would find it far easier to get it right!

[17]See Harris (2013) for an article that, in my opinion, agrees with my view in general but that delves much deeper into, among other things, this set of activities I'm not vouching for (or condemning) here. Although I believe more discussion is needed, Harris examined some interesting possibilities regarding market design to mitigate what he sees as some of the negative aspects of

HFT (while still agreeing that on net, HFT greatly lowers investors' costs). He also discussed the fact that many potential problems that are laid at the feet of HFT, such as order anticipation, are far from new even though they are often treated as problems unique to HFT.

[18]This activity includes trying to avoid market making when on the wrong end of an "informed" trader; admittedly, this is sometimes hard to distinguish from an information-based trade by the high-frequency traders themselves.

[19]See "Stock Options and the Lying Liars Who Don't Want to Expense Them" (Asness 2004).

[20]One other potential purpose of a buyback is to raise the financial leverage of the common stock. If more aggressive common stock is desired and investors cannot, as simple theory often assumes, lever on their own, doing so may provide added value to investors. However, I believe I'm now creating reasonable explanations that are decidedly not the reasons being claimed in the real world.

[21]Of course, there can be other reasons to choose between a portfolio of directly owned bonds and a bond fund that I don't address here. Taxes can be different. The commission and bid–ask costs of individual bonds can differ from the fee the fund charges and the trading costs incurred. If the fund is actively managed, you should invest largely on the basis of your belief in the net-of-fee skill of the manager. Diversification is generally greater in a fund—particularly, of course, for bonds with default risk. If the bonds are particularly illiquid, hard to mark, or expensive to trade, being in a fund might subject you to costs imposed by others getting in or out. Those can all matter. Someone recommending bonds over a bond fund or vice versa on these grounds is not subject to this peeve.

References

Asness, Clifford S. 2004. "Stock Options and the Lying Liars Who Don't Want to Expense Them." *Financial Analysts Journal*, vol. 60, no. 4 (July/August):9–14.

Asness, Clifford S., Robert J. Krail, and John M. Liew. 2001. "Do Hedge Funds Hedge?" *Journal of Portfolio Management*, vol. 28, no. 1 (Fall):6–19.

Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, vol. 33, no. 1 (February):3–56.

Harris, Larry. 2013. "What to Do about High-Frequency Trading." *Financial Analysts Journal*, vol. 69, no. 2 (March/April):6–9.

Sharpe, William. 1991. "The Arithmetic of Active Management." *Financial Analysts Journal*, vol. 47, no. 1 (January/February):7–9.

_____. 2010. "Adaptive Asset Allocation Policies." *Financial Analysts Journal*, vol. 66, no. 3 (May/June):45–59.

Shleifer, Andrei, and Robert W. Vishny. 1997. "The Limits of Arbitrage." *Journal of Finance*, vol. 52, no. 1 (March):35–55.

Sullivan, Rodney N. 2010. "The Art of the Better Forecast." *Financial Analysts Journal*, vol. 66, no. 6 (November/December):6–9.

Buffett's Alpha

Andrea Frazzini, David Kabiller, and Lasse Heje Pedersen

Warren Buffett's Berkshire Hathaway has realized a Sharpe ratio of 0.79 with significant alpha to traditional risk factors. The alpha became insignificant, however, when we controlled for exposure to the factors "betting against beta" and "quality minus junk." Furthermore, we estimate that Buffett's leverage is about 1.7 to 1, on average. Therefore, Buffett's returns appear to be neither luck nor magic but, rather, a reward for leveraging cheap, safe, high-quality stocks. Decomposing Berkshire's portfolio into publicly traded stocks and wholly owned private companies, we found that the public stocks have performed the best, which suggests that Buffett's returns are more the result of stock selection than of his effect on management.

Much has been said and written about Warren Buffett and his investment style, but little rigorous empirical analysis has been conducted to explain his performance. Every investor has a view on how Buffett has done it and the practical implications of his success, but we sought the answer via a thorough empirical analysis of Buffett's results in light of some of the latest research on the drivers of returns.[1]

Buffett's success has become the focal point of the debate on market efficiency that continues to be at the heart of financial economics. Efficient market supporters suggest that his success may simply be luck; Buffett is the happy winner of a coin-flipping contest, as articulated by Michael Jensen at a famous 1984 conference at Columbia Business School celebrating the 50th anniversary of the classic text by Graham and Dodd (1934).[2] Tests of this argument via a statistical analysis of Buffett's performance cannot fully resolve the issue. Instead, Buffett countered at the conference that it is no coincidence that many of the winners in the stock market come from the same intellectual village—that is, "Graham-and-Doddsville" (Buffett 1984). How can Buffett's counterargument be tested? Selecting successful investors who are informally classified as belonging to Graham-and-Doddsville *ex post* is subject to biases.

In our study, we used a different strategy to rigorously examine this issue. The standard academic factors that capture the market, size, value, and momentum premiums cannot explain Buffett's performance, so his success has to date been a mystery (Martin and Puthenpurackal 2008). We show that accounting for the general tendency of high-quality,

Reprinted from the *Financial Analysts Journal*, 2018, with permission.

safe, and cheap stocks to outperform can explain much of Buffett's performance.[3] This finding is consistent with the idea that investors from Graham-and-Doddsville follow similar strategies to achieve similar results and inconsistent with stocks being chosen based on coin flips. Hence, Buffett's success appears not to be luck, but, rather, Buffett personalizes the success of value and quality investment, providing real-world out-of-sample evidence on the ideas of Graham and Dodd (1934).

Buffett's record is remarkable in many ways, but we also examined just how spectacular the performance of Berkshire Hathaway has been when compared with that of other stocks or mutual funds.

To illustrate the practical relevance of our findings, we created a portfolio that tracked Buffett's market exposure and active stock-selection themes, leveraged to the same active risk as that of Berkshire Hathaway. We found that this systematic Buffett-style portfolio performed comparably to Berkshire.

Of course, explaining Buffett's performance with the benefit of hindsight does not diminish his outstanding accomplishment. He decided half a century ago to base his investments on the Graham and Dodd principles, and he found a way to apply leverage. Finally, he managed to stick to his principles and continue operating at high risk even after experiencing some ups and downs that have caused many other investors to rethink and retreat from their original strategies.

Finally, we consider whether Buffett's skill is the result of his ability to buy the right stocks or his ability as a CEO.

Data Sources

Our data come from several sources. We used stock return data from the CRSP database, balance sheet data from the Compustat North America database as well as hand-collected annual reports, holdings data for Berkshire Hathaway from the Thomson Reuters Institutional (13F) Holdings Database (based on Berkshire's US SEC filings), the size and cost of the insurance float from hand-collected comments in Berkshire Hathaway's annual reports, and mutual fund data from the CRSP Mutual Fund Database. We also used factor returns from Kenneth French's website and from Frazzini and Pedersen (2014) and Asness, Frazzini, and Pedersen (2013).[4] We describe our data sources and data filters in more detail in Appendix A.

Buffett's Track Record

Warren Buffett's track record is clearly outstanding. A dollar invested in Berkshire Hathaway in October 1976 (when our data sample starts) would have been worth more than $3,685 in March 2017 (when our data sample ends). Over this time period, Berkshire realized an average annual return of 18.6% in excess of the US T-bill rate, significantly outperforming the general stock market's average excess return of 7.5%.

Berkshire Hathaway stock also entailed more risk than the market; it realized a volatility of 23.5%, higher than the market volatility of 15.3%. Berkshire's excess return was high

even relative to its risk, however; it earned a Sharpe ratio of 18.6%/23.5% = 0.79, 1.6 times higher than the market's Sharpe ratio of 0.49. Berkshire realized a market beta of only 0.69, an important point that we discuss in more detail when we analyze the types of stocks that Buffett buys. Adjusting Berkshire's performance for market exposure, we computed its information ratio to be 0.64.

These performance measures reflect Buffett's impressive returns but also the fact that Berkshire Hathaway has been associated with some risk. Berkshire has had a number of down years and drawdown periods. For example, from 30 June 1998 to 29 February 2000, Berkshire lost 44% of its market value while the overall stock market was gaining 32%. Many fund managers might have had trouble surviving a shortfall of 76%, but Buffett's impeccable reputation and unique structure as a corporation allowed him to stay the course and rebound as the internet bubble burst.

To put Buffett's performance in perspective, we compared Berkshire's Sharpe and information ratios with those of all other US common stocks. If Buffett is more of a stock picker than a manager, then an even better reference group than other stocks might be the universe of actively managed mutual funds. Table 1 shows a comparison of Berkshire with both of these groups.

Buffett is in the top 3% among all mutual funds and the top 7% among all stocks. The stocks or mutual funds with the highest Sharpe ratios, however, are often ones that have existed only for short time periods and had a good run, which is associated with a large degree of randomness.

To minimize the effect of randomness, Table 1 also provides a comparison of Berkshire Hathaway with all stocks or mutual funds with at least a 10-year, 30-year, and 40-year history. Buffett's performance is truly outstanding from this perspective. Among all stocks with at least a 40-year history from 1976 to 2017, Berkshire realized the highest Sharpe ratio and information ratio. If you could travel back in time and pick one stock in 1976, Berkshire would be your pick. Figure 1 and Figure 2 also illustrate how Buffett lies in the very best tail of the performance distribution of mutual funds and stocks that have survived at least 40 years.

If an investment in Berkshire Hathaway were combined with an investment in the market, the optimal combination would put about 72% of the money in Berkshire, giving rise to a Sharpe ratio of 0.81. Hence, putting 100% of the money in Berkshire (rather than 72%) gives nearly the same as the optimal Sharpe ratio.[5]

Buffett's Leverage: Magnitude and Cost

Warren Buffett's large returns come from both his high Sharpe ratio and his ability to leverage his performance to achieve large returns at high risk. Buffett uses leverage to magnify returns, but how much leverage does he use? Furthermore, what are Buffett's sources of leverage, their terms, and their costs? To answer these questions, we studied Berkshire Hathaway's balance sheet, which can be summarized as in Exhibit 1.

We can compute Buffett's leverage, L, as follows:

$$L_t = \frac{\text{TA}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}},$$

Table 1. Buffett's Performance Relative to All Other Stocks and Mutual Funds, 1976–2017

A. Sample Distribution of Sharpe Ratios

						Buffett Performance	
Stock/Fund Measure	**Number of Stocks/ Funds**	**Median**	**95th Percentile**	**99th Percentile**	**Maximum**	**Rank**	**Percentile**
Sharpe ratio of equity mutual funds							
All funds in CRSP data	4,585	0.356	0.69	1.10	3.20	137	97.0%
All funds alive in 1976 and 2017	133	0.36	0.54	0.63	0.79	1	100.0%
All funds alive in 1976 with at least 10-year history	304	0.30	0.49	0.61	0.79	1	100.0%
All funds with at least 10-year history	2,872	0.39	0.62	0.74	0.99	11	99.7%
All funds with at least 30-year history	432	0.38	0.59	0.73	0.93	3	99.5%
All funds with at least 40-year history	186	0.33	0.52	0.63	0.79	1	100.0%
Sharpe ratio of common stocks							
All stocks in CRSP data	23,257	0.211	0.88	1.47	2.68	1,454	93.8%
All stocks alive in 1976 and 2017	504	0.36	0.51	0.57	0.79	1	100.0%
All stocks alive in 1976 with at least 10-year history	3,774	0.28	0.51	0.62	0.89	8	99.8%
All stocks with at least 10-year history	9,523	0.28	0.57	0.75	1.12	57	99.4%
All stocks with at least 30-year history	2,021	0.32	0.52	0.61	0.81	2	100.0%
All stocks with at least 40-year history	1,111	0.34	0.50	0.55	0.79	1	100.0%

(Continued)

Table 1. ***(Continued)***

B. Sample Distribution of Information Ratios							
Information ratio of equity mutual funds							
All funds in CRSP data	4,585	0.356	0.69	1.10	3.20	137	97.0%
All funds alive in 1976 and 2017	133	0.358	0.54	0.63	0.79	1	100.0%
All funds alive in 1976 with at least 10-year history	304	0.301	0.49	0.61	0.79	1	100.0%
All funds with at least 10-year history	2,872	0.390	0.62	0.74	0.99	11	99.7%
All funds with at least 30-year history	432	0.382	0.59	0.73	0.93	3	99.5%
All funds with at least 40-year history	186	0.331	0.52	0.63	0.79	1	100.0%
Information ratio of common stocks							
All stocks in CRSP data	23,257	0.211	0.88	1.47	2.68	1,454	93.8%
All stocks alive in 1976 and 2017	504	0.357	0.51	0.57	0.79	1	100.0%
All stocks alive in 1976 with at least 10-year history	3,774	0.276	0.51	0.62	0.89	8	99.8%
All stocks with at least 10-year history	9,523	0.277	0.57	0.75	1.12	57	99.4%
All stocks with at least 30-year history	2,021	0.323	0.52	0.61	0.81	2	100.0%
All stocks with at least 40-year history	1,111	0.335	0.50	0.55	0.79	1	100.0%

Notes: The information ratio is defined as the intercept in a regression of monthly excess returns on the excess return of the value-weighted market portfolio, divided by the standard deviation of the residuals. Sharpe ratios and information ratios are annualized.

Figure 1. How Berkshire Stacks Up in the Mutual Fund Universe. This figure shows the distribution of annualized information ratios of all actively managed equity funds in the CRSP mutual fund database with at least 40 years of return history. See also definitions in the notes to Table 1.

Figure 2. How Berkshire Stacks Up in the Common Stocks Universe. This figure shows the distribution of annualized information ratios of all common stock in the CRSP database with at least 40 years of return history. See also definitions in the notes to Table 1.

Exhibit 1. Stylized Balance Sheet of Berkshire Hathaway

Assets	**Liabilities and Shareholders' Equity**
Publicly traded equities	Liabilities
Privately held companies	Equity
Cash	
Total assets	Total liabilities

where TA_t is total assets, Cash_t is cash that Berkshire Hathaway owns, and Equity_t is Berkshire's equity value. The superscript *MV* is market value. We computed this measure of leverage for each month. We wanted to compute the leverage always using market values, but for some variables, we could observe only book values (indicated with superscript *BV*). We considered the market value of Berkshire's equity to be the stock price multiplied by the shares outstanding. Cash holdings are from Berkshire's consolidated balance sheet (see Appendix A). The balance sheet also provides the book value of the total assets, TA_t^{BV}, and the book value of equity, Equity_t^{BV}, which allowed us to estimate the market value of the total asset as

$$\text{TA}_t^{MV} = \text{TA}_t^{BV} + \text{Equity}_t^{MV} - \text{Equity}_t^{BV}.$$

Using this method, we estimated Buffett's average leverage to be 1.7 to 1. This amount is a nontrivial use of leverage and helps explain why Berkshire Hathaway experiences high volatility despite investing in a number of relatively stable businesses.

By focusing on total assets to equity, we captured all kinds of liabilities. We found, as we discuss later, that Berkshire Hathaway's financing arises from various types of liability. The two main liabilities are debt and insurance float. If we computed leverage, instead, as $\left(\text{Equity}_t^{MV} + \text{Debt}_t + \text{Float}_t\right) \Big/ \text{Equity}_t^{MV}$, then we found an average leverage of 1.4 to 1. In any event, note that our measure of leverage is subject to measurement noise.

Another expression of Buffett's use of leverage is shown in Berkshire Hathaway's stock price, which is significantly more volatile than the portfolio of publicly traded stocks that it owns, as we depict in Table 2. In fact, Berkshire's 23.5% stock volatility is 1.4 times higher than the 16.2% volatility of the portfolio of public stocks, which corresponds to a leverage of 1.4 if Berkshire's private assets are assumed to have similar volatility and if diversification effects are ignored. This leverage is similar to the leverage computed on the basis of balance sheet variables.

The magnitude of Buffett's leverage partly explains how he outperforms the market—but only partly. For example, if one applies 1.7-to-1 leverage to the market, it magnifies the market's average excess return to about 12.7%. Such a leveraged market return still falls far short, however, of Berkshire's 18.6% average excess return (and would result in a riskier and higher-beta portfolio than Buffett's).

In addition to the magnitude of Buffett's leverage, his sources of leverage, including their terms and costs, are interesting. Berkshire Hathaway's debt, enjoying a AAA rating from 1989 to 2009, has benefited from being highly rated. An illustration of the low financing rates enjoyed by Buffett is that Berkshire issued the first ever negative-coupon security in 2002, a senior note with a warrant.[6]

Berkshire Hathaway's anomalous cost of leverage arises, however, from its insurance float. Collecting insurance premiums up front and later paying a diversified set of claims is like taking a "loan." Table 3 shows that the estimated average annual cost of Berkshire's insurance float is only 1.72%, about 3 percentage points below the average T-bill rate. Hence, Buffett's low-cost insurance and reinsurance businesses have given him a significant advantage in terms of unique access to cheap, term leverage. After hand collecting the float data from Berkshire's annual reports, we estimated that 35% of Berkshire's liabilities, on average, consist of insurance float.[7]

Table 2. Berkshire Hathaway Return Decomposed into Leverage, Public Stocks, and Private Companies and Systematic Buffett-Style Strategies

	Performance				Buffett-Style Portfolio			Buffett-Style Long-Only Portfolio		
	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings	Overall Stock Market Performance	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings
Sample	1976–2017	1980–2017	1984–2017	1976–2017	1976–2017	1980–2017	1984–2017	1976–2017	1980–2017	1984–2017
Beta	0.69	0.77	0.29	1.00	0.69	0.77	0.29	0.85	0.86	0.87
Average excess return	18.6%	12.0%	9.3%	7.5%	28.5%	19.0%	18.1%	8.4%	8.6%	8.3%
Total volatility	23.5%	16.2%	20.6%	15.3%	23.5%	16.2%	20.6%	13.7%	13.8%	14.3%
Idiosyncratic volatility	21.1%	11.2%	20.1%	0.0%	21.1%	11.2%	20.1%	4.4%	4.5%	5.1%
Sharpe ratio	0.79	0.74	0.45	0.49	1.21	1.17	0.88	0.61	0.62	0.58
Information ratio	0.64	0.51	0.35	0.00	1.11	1.13	0.78	0.45	0.38	0.25
Leverage	1.71	1.00	1.00	1.00	5.09	2.99	4.14	1.00	1.00	1.00
Subperiod excess returns:										
1976–1980	41.2%	31.5%		7.3%	5.8%	27.3%		7.9%	30.5%	
1981–1985	28.6	21.3	18.4%	4.8	52.5	28.1	38.2%	7.0	6.3	9.7%
1986–1990	17.3	12.6	9.6	5.7	23.2	14.0	18.1	9.5	10.7	8.3
1991–1995	29.7	18.8	22.9	12.6	38.1	23.2	25.9	11.5	11.5	12.0

(Continued)

Table 2. ***(Continued)***

	Performance				Buffett-Style Portfolio			Buffett-Style Long-Only Portfolio		
	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings	Overall Stock Market Performance	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings
1996–2000	14.9	12.1	8.7	12.1	37.1	23.1	26.0	16.2	16.2	14.1
2001–2005	3.2	2.2	1.8	0.9	28.0	14.2	13.8	–0.8	–1.6	1.1
2006–2010	6.1	4.1	4.0	2.5	1.9	3.6	–5.7	2.7	1.1	1.0
2011–2015	10.8	9.9	5.0	12.1	33.8	21.8	21.8	11.2	11.5	11.0
2016–2017	19.3	13.6	11.1	15.0	45.3	32.9	20.3	15.0	14.6	14.5

Notes: Excess return is in excess to the US T-bill rate. To construct the mimicking portfolio of Berkshire's publicly traded stocks, at the end of each calendar quarter, under the assumption that the firm did not change holdings between reports, we collected Berkshire's common stock holdings from its 13F filings and computed portfolio monthly returns, weighted by Berkshire's dollar holdings. The stocks in the portfolio were refreshed quarterly on the basis of the latest 13F, and the portfolio was rebalanced monthly to keep constant weights. The mimicking portfolio of Berkshire's private holdings was constructed with the procedure described in Appendix B. The systematic Buffett-style portfolios were constructed from a regression of monthly excess returns. The explanatory variables are the monthly returns of the standard size, value, and momentum factors (Fama and French 1993; Asness 1994; Carhart 1997; Jegadeesh and Titman 1993), the betting-against-beta factor (Frazzini and Pedersen 2014), and the quality-minus-junk factor (Asness, Frazzini, and Pedersen 2013). Returns, volatilities, and Sharpe ratios are annualized. "Idiosyncratic volatility" is the volatility of the residual of a regression of monthly excess returns on market excess returns.

Table 3. Buffett's Cost of Leverage: The Case of His Insurance Float

			Spread over Benchmark Rates				
	Fraction of Years with Negative Cost	Average Cost of Funds (truncated)[a]	T-Bill	Fed Funds Rate	One-Month LIBOR	Six-Month LIBOR	10-Year Bond
1967–1970	0.75	0.29	–5.20	–6.03			–5.93
1971–1975	0.60	4.45	–1.18	–2.38			–2.51
1976–1980	1.00	0.00	–7.52	–8.61			–8.88
1981–1985	0.20	10.95	1.10	–0.26			–1.28
1986–1990	0.00	3.07	–3.56	–4.60	–4.79	–4.90	–5.30
1991–1995	0.60	2.21	–2.00	–2.24	–2.46	–2.71	–4.64
1996–2000	0.60	2.36	–2.70	–3.10	–3.33	–3.48	–3.56
2001–2005	0.60	1.29	–0.82	–0.96	–1.05	–1.19	–3.11
2006–2010	1.00	–4.73	–6.94	–7.18	–7.43	–7.73	–8.59
2011–2015	1.00	–2.37	–2.42	–2.48	–2.57	–2.86	–4.68
2016–2017	0.50	0.23	–0.39	–0.46	–0.57	–1.03	–1.85
Full sample	0.63	1.72	–2.97	–3.61	–3.42	–3.64	–4.71

[a]In years when cost of funds is reported as "less than zero" and no numerical value is available, we set the cost of funds to zero.
Notes: The data were hand-collected from Warren Buffett's comments in Berkshire Hathaway's annual reports. Rates are annualized, in percentage points.

Based on the balance sheet data, Berkshire also appears to finance part of its capital expenditures with tax deductions for accelerated depreciation of property, plant, and equipment as provided for under US Internal Revenue Service rules. For example, Berkshire reported $28 billion of such deferred tax liabilities in 2011 (p. 49 of the 2011 Annual Report). Accelerating depreciation is similar to an interest-free loan in the sense that (1) Berkshire enjoys a tax saving earlier than it otherwise would and (2) the dollar amount of the tax when it is paid in the future is the same as the earlier savings (i.e., the tax liability does not accrue interest or compound). Of course, Berkshire does pay taxes, which we discuss in a later section.

Berkshire Hathaway's remaining liabilities include accounts payable and derivative contract liabilities. Indeed, Berkshire has sold a number of derivative contracts, including writing index options on several major equity indexes—notably, put options and credit default obligations. For example, Berkshire stated in the 2011 Annual Report (p. 45),

> We received the premiums on these contracts in full at the contract inception dates. . . . With limited exceptions, our equity index put option and credit default contracts contain no collateral posting requirements with respect to changes in either the fair value or intrinsic value of the contracts and/or a downgrade of Berkshire's credit ratings.

Hence, Berkshire's sale of derivatives may serve both as a source of financing and as a source of revenue because such derivatives tend to be expensive (Frazzini and Pedersen 2012). Frazzini and Pedersen showed that investors that are either unable or unwilling to use leverage will pay a premium for instruments that embed the leverage, such as option contracts and levered exchange-traded funds. Buffett can profit by supplying this embedded leverage because he has unique access to stable and cheap financing.

Decomposing Buffett: Public Stocks vs. Private Companies

Berkshire Hathaway's stock return can be decomposed into the performance of the publicly traded companies that it owns, the performance of the privately held companies that it owns, and the leverage it uses. The performance of the publicly traded companies is a measure of Warren Buffett's stock-selection ability, whereas the performance of the privately held companies may additionally capture his success as a manager.

To evaluate Buffett's pure stock-selection ability, we used Berkshire Hathaway's 13F filings to collect the portfolio of publicly held companies and constructed a monthly time series of the market value of all Berkshire's public stocks, Public_t^{MV}, and the monthly return on this mimicking portfolio, r_{t+1}^{Public}. Specifically, at the end of each calendar quarter (under the assumption that the firm did not change holdings between reports), we collected Berkshire's common stock holdings from its 13F filing and computed portfolio monthly returns, weighted by Berkshire's dollar holdings. The stocks in the portfolio were refreshed quarterly on the basis of the latest 13F, and the portfolio was rebalanced monthly to keep the weights constant.

We could not directly observe the value and performance of Buffett's private companies, but based on what we do know, we could back them out. First, we could infer the market value of private holdings, Private_t^{MV}, as the residual because we could observe the

value of the total assets, the value of the publicly traded stocks, and the cash (see Buffett's balance sheet in Exhibit 1):

$$\text{Private}_t^{MV} = \text{TA}_t^{MV} - \text{Public}_t^{MV} - \text{Cash}_t^{MV}.$$

We then computed the return of these private holdings, $r_{t+1}^{Private}$, in a way that is immune to changes in the public stock portfolio and to splits/issuances by using split-adjusted returns as follows:

$$r_{t+1}^{Private} = \frac{\Delta\text{Private}_{t+1}^{MV}}{\text{Private}_t^{MV}}$$
$$= \frac{r_{t+1}^{f}\text{Liabilities}_t^{MV} + r_{t+1}^{Equity}\text{Equity}_t^{MV} - r_{t+1}^{Public}\text{Public}_t^{MV} - r_{t+1}^{f}\text{Cash}_t^{MV}}{\text{Private}_t^{MV}},$$

where r_{t+1}^{f} is the risk-free T-bill return, r_{t+1}^{Equity} is the return on Berkshire's stock, and the market value of liabilities is estimated as $\text{Liabilities}_t^{MV} = \text{TA}_t^{MV} - \text{Equity}_t^{MV}$.

Note that our estimate of the value of Berkshire Hathaway's private companies includes the value that the market attaches to Buffett himself (because it is based on the overall value of Berkshire Hathaway). To the extent that Berkshire's stock price is subject to randomness or mispricing (e.g., because of the Buffett-specific element), the estimated value and return of the private companies may be noisy.

Given our estimates for Buffett's public and private returns as well as his leverage, we could decompose Berkshire's performance; see Appendix B for a rigorous derivation. Berkshire's excess return can be decomposed into a weighted average of the return of the public stocks and the return of the private companies, leveraged up by L:

$$r_{t+1}^{Equity} - r_{t+1}^{f} =$$
$$\left[w_t\left(r_{t+1}^{Private} - r_{t+1}^{f}\right) + \left(1 - w_t\right)\left(r_{t+1}^{Public} - r_{t+1}^{f}\right)\right]L_t.$$

Berkshire's relative weight, w_t, on the private holdings is naturally given by

$$w_t = \frac{\text{Private}_t^{MV}}{\text{Private}_t^{MV} + \text{Public}_t^{MV}}.$$

Empirically, we found that Berkshire owned 65% private companies, on average, from 1980 to 2017, the remaining 35% being invested in public stocks. Berkshire's reliance on private companies has been increasing steadily over time—from less than 20% in the early 1980s to more than 78% in 2017.

Table 2 shows the performance of Buffett's public and private positions. Both have performed well. Buffett's public and private portfolios have exceeded the overall stock market in terms of average excess return, risk, and Sharpe ratio. The public stocks have a higher Sharpe ratio than the private stocks, suggesting that Buffett's skill comes mostly from his ability to pick stocks, not necessarily his value added as a manager (but keep in mind that our imputed returns may be subject to noise).

Berkshire Hathaway's overall stock return is far above the returns of both the private and public portfolios. The reason is that Berkshire is not simply a weighted average of the public and private components. It is also leveraged, which magnifies returns. Furthermore, Berkshire's Sharpe ratio is higher than those of the public and private parts, which reflects the benefits of diversification (and possibly benefits from time-varying leverage and time-varying public/private weights).

Buffett's Alpha and Investment Style: What Types of Stock?

We have noted that Warren Buffett's returns can be attributed to his stock selection and his ability to apply leverage, but how does he select his companies? To address this question, we considered Buffett's factor exposures:

$$r_t - r_t^f = \alpha + \beta_1 \text{MKT}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{UMD}_t + \beta_5 \text{BAB}_t + \beta_6 \text{QMJ}_t + \varepsilon_t,$$

where MKT is the excess return of the overall equity market, SMB is the size factor (small minus big), HML is the value versus growth factor (high book to market minus low book to market), UMD is the momentum factor (up minus down), BAB is betting against beta, and QMJ is quality minus junk.

As shown in Table 4, we ran this regression for the excess return, $r_t - r_t^f$, of Berkshire Hathaway stock, the portfolio of publicly held stocks inferred from the 13F filings, and the portfolio of private companies computed as described previously. For each of these returns, we first ran a regression on the market return (MKT). Berkshire has a beta of less than 1 and a significant alpha. We next controlled for the standard factors that capture the effects of size and value (Fama and French 1993) and momentum (Asness 1994; Carhart 1997; Jegadeesh and Titman 1993). The size factor, SMB, is a strategy of going long small-capitalization stocks and short large-cap stocks. Hence, a positive loading on SMB reflects a tendency to buy small-cap stocks, so Berkshire's negative loading reflects a tendency to buy large-cap stocks. The value factor, HML, is a strategy of buying stocks of high book value to market value while shorting stocks of low book value to market value. Berkshire's positive loading thus reflects a tendency of buying stocks that are cheap—in the sense of having high book value relative to their market value. The last of the four "standard" factors is the momentum factor, UMD, which corresponds to buying stocks that have performed well relative to peers over the past year (winners) while shorting the stocks that are relative underperformers (losers). Berkshire's insignificant loading on UMD means that Buffett is not chasing trends in his stock selection.

Collectively, these four standard factors do not explain much of the alpha shown in Table 4. Because Buffett's alpha cannot be explained by standard factors studied by academics, his success has to date been considered a sign of his unique skill or simply a mystery.

Our innovation for this study was to also control for the factors betting against beta, BAB, described in Frazzini and Pedersen (2014) and quality, QMJ, of Asness, Frazzini, and Pedersen (2013). A loading on the BAB factor reflects a tendency to buy safe

Table 4. What Kinds of Companies Does Berkshire Hathaway Own?

	Berkshire Stock, 10/1976–3/2017				13F Portfolio, 4/1980–3/2017				Private Holdings, 4/1980–3/2017			
Alpha	**13.4%**	**11.0%**	**8.5%**	5.4%	**5.8%**	**4.5%**	3.0%	0.3%	**7.0%**	4.9%	3.9%	3.5%
	(4.01)	(3.30)	(2.55)	(1.55)	(3.09)	(2.46)	(1.62)	(0.16)	(1.98)	(1.40)	(1.10)	(0.91)
MKT	**0.69**	**0.83**	**0.83**	**0.95**	**0.77**	**0.85**	**0.86**	**0.95**	**0.30**	**0.39**	**0.40**	**0.42**
	(11.00)	(12.74)	(12.99)	(12.77)	(22.06)	(23.81)	(24.36)	(23.52)	(4.46)	(5.63)	(5.72)	(5.03)
SMB		**–0.29**	**–0.30**	–0.13		**–0.19**	**–0.19**	–0.05		**–0.26**	**–0.25**	–0.23
		–(3.11)	–(3.19)	–(1.17)		–(3.73)	–(3.79)	–(0.95)		–(2.65)	–(2.56)	–(1.95)
HML		**0.47**	**0.31**	**0.40**		**0.28**	**0.19**	**0.25**		**0.28**	0.21	0.22
		(4.68)	(2.82)	(3.55)		(5.20)	(3.25)	(4.32)		(2.63)	(1.80)	(1.85)
UMD		0.06	–0.02	–0.05		–0.01	–0.06	**–0.09**		0.08	0.04	0.04
		(1.00)	–(0.25)	–(0.80)		–(0.36)	–(1.66)	–(2.58)		(1.24)	(0.62)	(0.51)
BAB			**0.33**	**0.27**			**0.19**	**0.15**			0.15	0.14
			(3.79)	(3.04)			(4.08)	(3.18)			(1.61)	(1.53)
QMJ				**0.47**				**0.37**				0.07
				(3.06)				(4.55)				(0.43)
$\overline{R}^2$	0.20	0.25	0.27	0.29	0.52	0.58	0.59	0.61	0.05	0.08	0.08	0.08
Obs.	486	486	486	486	444	444	444	444	399	399	399	399

(t-statistics in parentheses)

Notes: This table shows calendar-time portfolio returns. Alphas are annualized. Boldface indicates statistical significance at the 5% level. See also the notes to Table 2.

(i.e., low-beta) stocks while shying away from risky (i.e., high-beta) stocks. Similarly, a loading on the QMJ factor reflects a tendency to buy high-quality companies—that is, companies that are profitable, growing, and safe and have high payout.[8]

Table 4 reveals that Berkshire Hathaway loads significantly on the BAB and QMJ factors, indicating that Buffett likes to buy safe, high-quality stocks. Controlling for these factors drives the alpha of Berkshire's public stock portfolio down to a statistically insignificant annualized 0.3%. That is, these factors almost completely explain the performance of Buffett's public portfolio. Hence, a significant part of the secret behind Buffett's success is the strategy of buying safe, high-quality, value stocks. These factors also explain a large part of Berkshire's overall stock return and of the private part in that their alphas become statistically insignificant when BAB and QJM are controlled for. The point estimate of Berkshire's alpha, however, drops only by about half.

Although Warren Buffett is known as the ultimate value investor, we find that his focus on safe, high-quality stocks may be at least as important to his performance. Our statistical finding is consistent with Buffett's own words from the Berkshire Hathaway 2008 Annual Report: "Whether we're talking about socks or stocks, I like buying quality merchandise when it is marked down."

We emphasize again that being able to explain Buffett's returns by using factors from academic papers written decades after Buffett put the strategies into practice does not make Buffett's success any less impressive. It is interesting, however, to discover the importance of leveraging low-beta, high-quality stocks for the person known as the "ultimate value investor."

A Systematic Buffett Strategy

Given that we can attribute Warren Buffett's performance to leverage and his focus on safe, high-quality, value stocks, we naturally wanted to consider how well one could do by implementing these investment themes in a systematic way. Buffett is known as an active stock picker, but we tried to go back to Buffett's roots and, in the spirit of Graham and Dodd (1934), focused on systematically implemented screens.

We considered systematic Buffett-style portfolios that tracked Buffett's market exposure and active stock-selection themes. First, we captured Buffett's market exposure, $\beta^{Buffett}$, as the slope of a univariate regression of Berkshire Hathaway's excess returns on the market portfolio. Second, we captured Buffett's stock-selection tilts by running a regression of Berkshire's monthly beta-adjusted returns on the factors that help explain its performance, as described in the previous section:

$$\begin{aligned} & r_t - r_t^f - \beta^{Buffett}\mathrm{MKT}_t = \\ & \alpha + m\mathrm{MKT}_t + s\mathrm{SMB}_t + h\mathrm{HML}_t \\ & + u\mathrm{UMD}_t + b\mathrm{BAB}_t + q\mathrm{QMJ}_t + \varepsilon_t. \end{aligned}$$

The regression coefficients are equal to those in the fifth column of Table 4, with the exception that the market loading is reduced by an amount equal to $\beta^{Buffett}$. The right-hand side excluding the alpha and the error term captures Buffett's *active* stock-selection tilts:

$$r_t^A = m\mathrm{MKT}_t + s\mathrm{SMB}_t + h\mathrm{HML}_t + u\mathrm{UMD}_t + b\mathrm{BAB}_t + q\mathrm{QMJ}_t.$$

We rescaled this active return series to match Berkshire's idiosyncratic volatility, σ_I, to simulate the use of leverage and to counter any attenuation bias:

$$r_t^{Active} = r_t^A \frac{\sigma_I}{\sigma_{r_t^A}}.$$

Finally, we added back Berkshire's market exposure and the risk-free return, r_t^f, to construct our systematic Buffett-style portfolio:

$$r_t^{Buffett\ style} = r_t^f + \beta^{Buffett}\mathrm{MKT}_t + r_t^{Active}.$$

This systematic Buffett-style strategy is a diversified portfolio that matches Berkshire's beta, idiosyncratic volatility, total volatility, and relative active loadings.

We similarly constructed a Buffett-style portfolio based on the loadings and volatility of Berkshire Hathaway's public and private equity holdings. Table 2 reports the performance of our systematic Buffett-style portfolios, and Figure 3 shows the cumulative return of Berkshire, Buffett's public stocks, and our systematic Buffett-style strategies. Finally, Table 5 reports correlations, alphas, and loadings for our systematic Buffett-style portfolios and their actual Buffett counterparts.

The performance of the systematic Buffett-style portfolios are comparable to Buffett's actual return. Because the simulated Buffett-style portfolios do not account for transaction costs and other costs and benefit from hindsight, their apparent outperformance should be discounted. The main insight here is the high covariation between Buffett's actual performance and the performance of a diversified Buffett-style strategy.

The Buffett-style portfolio matched the public stock portfolio especially closely, perhaps because this public portfolio was observed directly and its returns were calculated from public stock returns in a method that used the same methodology as our systematic portfolios. Berkshire's overall stock price, however, may have idiosyncratic price variation (e.g., because of the value of Buffett himself) that cannot be replicated by using other stocks. This idiosyncratic Berkshire variation is even more severe for the private part, which may also suffer from measurement issues.

The comparison of Berkshire Hathaway's public stock portfolio and the corresponding Buffett-style portfolio is also the cleaner test of Buffett's stock selection because *both* are simulated returns without any transaction costs or taxes. Indeed, the correlation between our systematic portfolio and Berkshire's public stock portfolio (shown in Table 5) is 73%, meaning that our systematic portfolio explains 53% of the variance of the public stock portfolio. The correlations between the systematic portfolio and Berkshire's stock price and between the systematic portfolio and Buffett's private investments are lower (48% and 26%, respectively) but still large in magnitude. Table 5 also shows that our systematic portfolios have significant alphas with respect to their corresponding Buffett counterparts, whereas none of the Buffett portfolios have statistically significant alphas with respect to their systematic counterparts. This result may have arisen because our systematic portfolios

Figure 3. Performance of the Equity Market, Berkshire Hathaway, and a Systematic Buffett-Style Portfolio. Panel A shows the cumulative return of Berkshire Hathaway's portfolio of publicly traded stocks (as reported in its 13F filings), a corresponding systematic Buffett-mimicking portfolio, and the CRSP value-weighted market return (leveraged to the same volatility as Berkshire's public stocks). Panel B shows the cumulative return of Berkshire Hathaway, a corresponding systematic Buffett-mimicking portfolio, and the CRSP value-weighted market return (leveraged to the same volatility as Berkshire). The systematic Buffett-style strategy was constructed from a regression of monthly excess returns of Berkshire Hathaway stock and the portfolio of publicly held stocks inferred from the 13F filings (columns 5 and 9, respectively, in Table 4). The explanatory variables are the monthly returns of the six factors. The systematic Buffett-style portfolio excess return is the sum of the explanatory variables multiplied by the respective regression coefficients, rescaled to match the volatility of Berkshire's return.

Table 5. Buffett's Returns vs. a Systematic Buffett Strategy

	Berkshire Regressed on Systematic Portfolio			Systematic Portfolio Regressed on Berkshire		
	Berkshire Hathaway	**Public US Stocks (from 13F filings)**	**Private Holdings**	**Berkshire Hathaway**	**Public US Stocks (from 13F filings)**	**Private Holdings**
Sample	1976–2017	1980–2017	1984–2017	1976–2017	1980–2017	1984–2017
Alpha (annualized)	5.4%	0.3%	3.5%	**18.7%**	**10.2%**	**14.2%**
	(1.64)	(0.17)	(0.98)	(5.85)	(5.99)	(4.10)
Loading	**0.34**	**0.43**	**0.23**	**0.34**	**0.43**	**0.23**
	(8.05)	(10.01)	(4.60)	(8.05)	(10.01)	(4.60)
MKT	**0.45**	**0.44**	**0.23**	**0.45**	**0.44**	**0.23**
	(6.87)	(9.59)	(3.45)	(6.87)	(9.59)	(3.45)
Correlation	0.48	0.73	0.26	0.48	0.73	0.26
$\bar{R}^2$	0.29	0.61	0.09	0.29	0.61	0.09

(t-statistics in parentheses)

Notes: This table shows calendar-time portfolio returns. Alpha is the intercept in a regression of monthly excess return. Alphas are annualized. Boldface indicates statistical significance at the 5% level. See also the notes to Table 2.

have similar factor tilts as Buffett's but hold a much larger number of securities, thus benefiting from diversification.

The Berkshire Hathaway stock return does reflect incurred transaction costs and possibly additional taxes, so Buffett's performance is all the more impressive. Given Berkshire's modest turnover initially, transaction costs were probably small in the early days. As Berkshire grew, so did transaction costs, which could account for some of Berkshire's diminishing returns over time. Furthermore, Berkshire initially focused on small companies, which is reflected in a positive SMB loading in the first half of the time period (not shown) but may have been increasingly forced to focus on large-cap stocks. Indeed, Table 4 shows that Berkshire has a negative loading on the SMB factor. Hence, Berkshire's diminishing returns could also be related to capacity constraints.

Assessing the impact of taxes on Berkshire Hathaway's performance is complicated. For Berkshire's private holdings, the joint ownership in a multinational company is associated with tax advantages. For the public stocks, Berkshire could face double corporate taxes—that is, the need to pay tax both indirectly in the portfolio companies' earnings and in Berkshire as it receives dividends or realizes capital gains. However, Berkshire can deduct 70%–80% of the dividends received, defer capital gains taxes by holding on to the positions so that gains remain unrealized,[9] and minimize taxes by allocating earnings abroad as a multinational.[10] Hence, assessing whether Berkshire is at a tax disadvantage overall is difficult.

In addition to the systematic long–short portfolios, we computed a long-only, unleveraged, systematic Buffett-style strategy. At the end of each calendar month, we sorted securities on the basis of the portfolio weights corresponding to our active tilts, r_t^{Active}, and constructed an equal-weighted portfolio that held the top 50 stocks with the highest portfolio weights. Table 2 shows that these simple Buffett-style portfolios also performed well, albeit not as well as when we allowed short selling.

As a final robustness check, we considered Buffett-style portfolios that did not rely on in-sample regression coefficients. Specifically, we created an implementable Buffett-style strategy by using only information up to month t to construct portfolio weights for the next month, $t + 1$. These portfolios performed similarly to our full-sample Buffett-style portfolios and had similar alphas, as described in Appendix C.

In summary, if one had applied leverage to a portfolio of safe, high-quality, value stocks consistently over this time period, one would have achieved a remarkable return, as Buffett did. Of course, he started doing it more than half a century before we wrote this paper!

Conclusion and Practical Implications of the Oracle's Alpha

We have showed just how spectacular the performance of Berkshire Hathaway has been when compared with that of other stocks or mutual funds. Indeed, for the sample we studied, we found that Berkshire Hathaway had the highest Sharpe ratio among all and a higher Sharpe ratio than all US mutual funds that have been around for more than 40 years. We found that the Sharpe ratio of Berkshire Hathaway was 0.79 over the period 1976–2017. Although this Sharpe ratio is nearly double that of the overall stock market, it is lower than many investors imagine. Adjusting for the market exposure, Buffett's information ratio is lower, 0.64. The Sharpe ratio reflects high average returns but also significant risk and periods of losses and significant drawdowns.

If his Sharpe ratio is very good but not super-human, then how did Buffett become among the richest in the world? The answer is that he stuck to a good strategy—buying cheap, safe, quality stocks—for a long time period, surviving rough periods where others might have been forced into a fire sale or a career shift, and he boosted his returns by using leverage. We estimated that Buffett applies a leverage of about 1.7 to 1, boosting both his risk and excess return in that proportion. Thus, his many accomplishments include having the conviction, wherewithal, and skill to operate with leverage and significant risk over a number of decades.

We identified several general features of Buffett's chosen portfolio: He buys stocks that are safe (with low beta and low volatility), cheap (i.e., value stocks with low price-to-book ratios), and of high quality (profitable, stable, and growing stocks with high payout ratios). Interestingly, stocks with these characteristics tend to perform well in general, so these characteristics help explain Buffett's investment.

We created a portfolio that tracked Buffett's market exposure and active stock-selection themes and was leveraged to the same active risk as Berkshire Hathaway. We found that this systematic Buffett-style portfolio performed comparably to Berkshire. Buffett's genius thus appears to be at least partly in recognizing early on that these investment themes work, applying leverage without ever having a fire sale, and sticking to his principles. Perhaps this is what he means by his comment in the Berkshire Hathaway 1994 Annual Report: "Ben Graham taught me 45 years ago that in investing it is not necessary to do extraordinary things to get extraordinary results."

Finally, we considered whether Buffett's skill is the result of his ability to buy the right stocks or his ability as a CEO. We decomposed Berkshire Hathaway's returns into two parts—investments in publicly traded stocks and the private companies run within Berkshire. We found that both public and private companies contribute to Buffett's performance but the portfolio of public stocks performs the best, suggesting that Buffett's skill is mostly in stock selection.

We then asked why he relies heavily on private companies, including the insurance and reinsurance businesses. One reason might be taxes, and another might be that this structure provides a steady source of financing that allows him to leverage his stock-selection ability. Indeed, we found that 36% of Buffett's liabilities consist of insurance float (i.e., insurance premiums paid up front) with an average cost below the T-bill rate.

In summary, we found that Buffett has developed a unique access to leverage; that he has invested in safe, high-quality, cheap stocks; and that these key characteristics can largely explain his impressive performance.

Our results have the following three notable practical implications.

First, we shed new light on the efficiency of capital markets by studying in a novel way the famous coin-flipping debate at the 1984 Columbia conference between Michael Jensen representing the efficient market economists and Warren Buffett representing Graham-and-Doddsville. The 2013 and 2017 Nobel prizes reignited this debate; as a prototypical example, see the *Forbes* article "What Is Market Efficiency?" (Heakal 2013): "In the real world of investments, however, there are obvious arguments against the [efficient market hypothesis]. There are investors who have beaten the market—Warren Buffett."

The efficient market counterargument is that Buffett was simply lucky. Our findings suggest that Buffett's success is neither luck nor magic but is a reward for a successful implementation of value and quality exposures that have historically produced high returns. Second, we illustrated how Buffett's record can be viewed as an expression of the practical

implementability of academic factor returns after transaction costs and financing costs. We simulated how investors can try to take advantage of similar investment principles. Buffett's success shows that the high returns of these academic factors are not simply "paper" returns; these returns can be realized in the real world after transaction costs and funding costs, at least by Warren Buffett. Furthermore, Buffett's exposure to the BAB factor and his unique access to leverage are consistent with the idea that the BAB factor represents reward to the use of leverage.

Third, our results illustrate what investment success looks like in the real world. Although optimistic asset managers often claim to be able to achieve Sharpe ratios above 1 or 2 and many chief investment officers seek similarly high performance numbers, our results suggest that long-term investors might do well to set a realistic performance goal and brace themselves for the tough periods that even Buffett has experienced. Indeed, because Buffett became one of the richest people in the world with a Sharpe ratio of 0.79, most investors should seek to actually deliver a Sharpe ratio somewhere between this number and the market's Sharpe ratio, which was around 0.5 during this sample period, rather than making suboptimal investments in a futile attempt to consistently reach a much higher number.

We thank Cliff Asness, Aaron Brown, John Howard, Ronen Israel, Sarah Jiang, and Scott Richardson for helpful comments and discussions, as well as seminar participants at the Kellogg School of Management, CFA Society Denmark, Vienna University of Economics and Business, Goethe University Frankfurt, and AQR Capital Management. We are grateful to Nigel Dally for providing us with historical 10-K filings.

Disclosure: the authors are principals at AQR Capital Management, a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described in this article. The views expressed here are those of the authors and not necessarily those of AQR.

Endnotes

[1]Based on the original insights of Black (1972) and Black, Jensen, and Scholes (1972), Frazzini and Pedersen (2014) showed that leverage and margin requirements change equilibrium risk premiums. They demonstrated that investors without binding leverage constraints can profit from betting against beta—that is, buying low-risk assets and shorting risky assets. Frazzini and Pedersen (2012) extended this finding to derivatives with embedded leverage, and Asness, Frazzini, and Pedersen (2012) added the risk–return relationship across asset classes. Asness, Frazzini, and Pedersen (2013) considered fundamental measures of risk and other accounting-based measures of "quality"—that is, characteristics that increase a company's value.

[2]Graham and Dodd's *Security Analysis* (1934) is credited with laying the foundation for investing based on value and quality, and Benjamin Graham and David Dodd were Buffett's professors at Columbia Business School.

[3]Value stocks, on average, outperform growth stocks, as documented by Stattman (1980), Rosenberg, Reid, and Lanstein (1985), and Fama and French (1993), and high-quality stocks outperform junk stocks, on average, as documented by Asness, Frazzini, and Pedersen (2013) and references therein.

[4]The French website is http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[5]This result can be seen directly by optimizing numerically among these combinations or via the theoretical result of Treynor and Black (1973) that the highest Sharpe ratio arises as the square root of the sum of the squared market Sharpe ratio and the squared information ratio. See also the generalized result of Clarke, De Silva, and Thorley (2016).

[6]See the Berkshire Hathaway news release of 22 May 2002 at www.berkshirehathaway.com/news/may2202.html.

[7]For example, the 2017 Annual Report provides the float number on p. 7 and states on pp. 6 and 7, "Before I discuss our 2017 insurance results, let me remind you of how and why we entered the field. We began by purchasing National Indemnity and a smaller sister company for $8.6 million in early 1967 . . . , insurance business that usually delivered an underwriting profit. Even more important, the insurance operation carried with it $19.4 million of 'float'— money that belonged to others but was held by our two insurers. Ever since, float has been of great importance to Berkshire. . . . Premiums are generally paid to the company upfront whereas losses occur over the life of the policy. . . . As a result of our emphasizing that sort of business, Berkshire's growth in float has been extraordinary. We are now the country's second largest p/c [property and casualty insurance] company measured by premium volume and its leader, by far, in float."

[8]See Asness, Frazzini, and Pedersen (2013) for details.

[9]For a corporation, capital gains are subject to corporate taxes (with no special provision for long-term capital gains). Capital gains taxes can be deferred from a *cash flow perspective* as long as they are unrealized, but the accrued capital gains tax does lead to an expense from the *perspective of generally accepted accounting principles*. That is, Berkshire Hathaway does not *pay* any taxes for unrealized capital gains, but such unrealized capital gains do lower its reported *earnings* and hence its book value of equity while raising the accounting liability called "principally deferred income taxes."

[10]For instance, Berkshire Hathaway's 2011 Annual Report states, "We have not established deferred income taxes with respect to undistributed earnings of certain foreign subsidiaries. Earnings expected to remain reinvested indefinitely were approximately $6.6 billion as of December 31, 2011. Upon distribution as dividends or otherwise, such amounts would be subject to taxation in the U.S. as well as foreign countries. However, U.S. income tax liabilities would be offset, in whole or in part, by allowable tax credits with respect to income taxes previously paid to foreign jurisdictions. Further, repatriation of all earnings of foreign subsidiaries would be impracticable to the extent that such earnings represent capital needed to support normal business operations in those jurisdictions. As a result, we currently believe that any incremental U.S. income tax liabilities arising from the repatriation of distributable earnings of foreign subsidiaries would not be material."

References

Asness, C.S. 1994. "Variables That Explain Stock Returns." PhD dissertation, University of Chicago.

Asness, C.S., A. Frazzini, and L.H. Pedersen. 2012. "Leverage Aversion and Risk Parity." *Financial Analysts Journal* 68 (1): 47–59. doi:10.2469/faj.v68.n1.1

Asness, C.S., A. Frazzini, and L.H. Pedersen. 2013. "Quality Minus Junk." Forthcoming, *Review of Accounting Studies*.

Black, F. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business* 45 (3): 444–55. doi:10.1086/295472

Black, F., M.C. Jensen, and M. Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*, edited by Michael C. Jensen, 79–121. New York: Praeger.

Buffett, W.E. 1984. "The Superinvestors of Graham-and-Doddsville." *Columbia Business School Magazine* (17 May): 4–15.

Carhart, M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52 (1):57–82. doi:10.1111/j.1540-6261.1997.tb03808.x

Clarke, R., H. De Silva, and S. Thorley. 2016. "Fundamentals of Efficient Factor Investing." *Financial Analysts Journal* 72 (6): 9–26. doi:10.2469/faj.v72.n6.3

Fama, E.F., and K.R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1):3–56. doi:10.1016/0304-405X(93)90023-5

Frazzini, A., and L.H. Pedersen. 2012. "Embedded Leverage." Working paper, AQR Capital Management (November). doi:10.3386/w18558

Frazzini, A., and L.H. Pedersen. 2014. "Betting against Beta." *Journal of Financial Economics* 111 (1): 1–25. doi:10.1016/j.jfineco.2013.10.005

Graham, B., and D. L. Dodd. 1934. *Security Analysis*. New York: McGraw Hill.

Heakal, Reem. 2013. "What Is Market Efficiency?" *Forbes* (1 November): www.forbes.com/sites/investopedia/2013/11/01/what-is-market-efficiency/#5a32dacf1d2d.

Jegadeesh, N., and S. Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1): 65–91. doi:10.1111/j.1540-6261.1993.tb04702.x

Kacperczyk, M., C. Sialm, and L. Zheng. 2008. "Unobserved Actions of Mutual Funds." *Review of Financial Studies* 21 (6):2379–416. doi:10.1093/rfs/hhl041

Lowenstein, R. 2008. *Buffett: The Biography*. London: Duckworth Press.

Martin, G.S., and J. Puthenpurackal. 2008. "Imitation Is the Sincerest Form of Flattery: Warren Buffett and Berkshire Hathaway." Working paper, American University. doi:10.2139/ssrn.806246

Rosenberg, B., K. Reid, and R. Lanstein. 1985. "Persuasive Evidence of Market Inefficiency." *Journal of Portfolio Management* 11 (3):9–16. doi:10.3905/jpm.1985.409007

Stattman, D. 1980. "Book Values and Stock Returns." *Chicago MBA: A Journal of Selected Papers* 5: 25–45.

Treynor, J., and F. Black. 1973. "How to Use Security Analysis to Improve Portfolio Selection." *Journal of Business* 46 (1): 66–88. doi:10.1086/295508

Appendix A: Data Sources and Methodology

Stock Return Data

Stock return and price data are from the CRSP database. Our data include all domestic common stocks (share codes 10 and 11) on the CRSP tape between December 1925 and March 2017. To compute Berkshire Hathaway's stock returns, we value-weighted share classes A and B on the basis of lagged market capitalization (Berkshire Hathaway introduced a share class B in April 1996).

The stock return data for Berkshire Hathaway on the CRSP tape start in 1976. Hence, we have data only for the last 41 years of Warren Buffett's record. He ran various private investment partnerships from 1957 to 1969, started trading Berkshire in 1962, took control of Berkshire in 1965, and started using Berkshire as his main investment vehicle after he closed his partnerships in 1969 (Lowenstein 2008). At the time of this writing, we have been unable to collect data on Berkshire's stock price prior to its introduction on the CRSP tape and Buffett's partnership performance, so our study covers the period 1976 to 2017, which can be viewed as a conservative estimate of Buffett's complete track record and out-of-sample evidence relative to his first almost 20 years of success.

Balance Sheet Data

Our main source of balance sheet data is the Compustat North America database. Because of several errors in the cash item (especially in the quarterly reports in the early part of the sample period), however, we checked and corrected these data with information extracted from the original Form 10-K filings as well as information from Berkshire Hathaway's annual letter to shareholders. Berkshire holds a significant amount of cash on its balance sheet, which we hand-collected from Berkshire's annual reports and Form 10-K filings.

We made the following adjustments: For the end of 1985, the official cash number included a significant amount of cash set aside for the purchases of Capital Cities Communications and the Scott Fetzer Company. Therefore, we used the pro forma consolidated balance sheet presented in note 18 on p. 42 of the 1985 Annual Report. For the end of 1987, we used the restated cash figure mentioned in the 1988 Annual Report, note 1(b), p. 25. For other balance sheet items, we also focused on annual balance sheet data.

13F Holdings Data

We downloaded holdings data for Berkshire Hathaway from the Thomson Reuters Institutional (13F) Holdings Database, which includes holdings of all US entities exercising investment discretion over $100 million or more and filed with the SEC. The data on Berkshire's public stock holdings run from 1980 to 2017.

Mutual Fund Data

We collected mutual fund returns from the CRSP Mutual Fund Database. The data run from 1976 to 2017. We focused our analysis on open-end actively managed domestic equity mutual funds. Our sample selection procedure followed that of Kacperczyk, Sialm, and Zheng (2008); see their appendix for details about the screens that were used and summary statistics of the data.

Appendix B: Decomposing Berkshire Hathaway's Returns

To decompose Berkshire Hathaway's returns into its public equity part, private equity part, and leverage, we first defined the private equity return as

$$r_{t+1}^{Private} = \frac{r_{t+1}^{f}\text{Liabilities}_t^{MV} + r_{t+1}^{Equity}\text{Equity}_t^{MV} - r_{t+1}^{Public}\text{Public}_t^{MV} - r_{t+1}^{f}\text{Cash}_t^{MV}}{\text{Private}_t^{MV}}.$$

Rearranging this expression so that the overall Berkshire return is on the left side yields

$$
\begin{aligned}
r_{t+1}^{Equity} &= r_{t+1}^{Private}\frac{\text{Private}_t^{MV}}{\text{Equity}_t^{MV}} + r_{t+1}^{Public}\frac{\text{Public}_t^{MV}}{\text{Equity}_t^{MV}} - r_{t+1}^{f}\frac{\text{Liabilities}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}} \\
&= \left(r_{t+1}^{Private}\frac{\text{Public}_t^{MV}}{\text{Private}_t^{MV} + \text{Public}_t^{MV}} + r_{t+1}^{Public}\frac{\text{Public}_t^{MV}}{\text{Public}_t^{MV} + \text{Public}_t^{MV}} \right) * \\
&\quad L_t - r_{t+1}^{f}\frac{\text{Liabilities}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}},
\end{aligned}
$$

where we used

$$
\begin{aligned}
L_t &= \frac{\text{TA}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}} \\
&= \frac{\text{Private}_t^{MV} + \text{Public}_t^{MV}}{\text{Equity}_t^{MV}}.
\end{aligned}
$$

The excess return of Berkshire could now be written in terms of the weight of the private holdings, $w_t = \text{Private}_t^{MV} \Big/ \left(\text{Private}_t^{MV} + \text{Public}_t^{MV}\right)$, as follows:

$$
\begin{aligned}
r_{t+1}^{Equity} - r_{t+1}^{f} &= \left[w_t r_{t+1}^{Private} + \left(1 - w_t\right) r_{t+1}^{Public} \right] L_t - r_{t+1}^{f}\left(\frac{\text{Liabilities}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}} + 1 \right) \\
&= \left[w_t \left(r_{t+1}^{Private} - r_{t+1}^{f} \right) + \left(1 - w_t\right)\left(r_{t+1}^{Public} - r_{t+1}^{f} \right) \right] L_t - r_{t+1}^{f}\left(\frac{\text{Liabilities}_t^{MV} - \text{Cash}_t^{MV}}{\text{Equity}_t^{MV}} + 1 - L_t \right) \\
&= \left[w_t \left(r_{t+1}^{Private} - r_{t+1}^{f} \right) + \left(1 - w_t\right)\left(r_{t+1}^{Public} - r_{t+1}^{f} \right) \right] L_t.
\end{aligned}
$$

The Berkshire equity excess return, therefore, depends on the excess returns of private and public holdings, their relative importance, and the degree of leverage.

Note that the 13F holdings data and mimicking portfolio returns, r_{t+1}^{Public}, start in 1980. Our way of estimating returns from private holdings, however, produced very noisy estimates for the first three years of the sample. Also, there were several outliers in the imputed $r_{t+1}^{Private}$ in the first years of the sample, with several returns exceeding +100% monthly. Therefore, we focused most of the analysis on $r_{t+1}^{Private}$ in the period 1984–2017, for which our method produced less noisy estimates.

Appendix C: Implementable Systematic Buffett Strategies

We constructed systematic Buffett-style portfolios that tracked Warren Buffett's market exposure and active stock-selection themes. We did this step as in Table 2, except here, the analysis is implementable in real time (i.e., out of sample).

At the end of each calendar month t, using data up to month t, we first captured Buffett's market exposure, $\beta^{Buffett}$, as the slope of a univariate regression of Berkshire Hathaway's

excess returns on the market portfolio. Second, we captured Buffett's stock-selection tilts by running a regression of his monthly beta-adjusted returns on the factors that help explain his performance:

$$r - r^f - \beta^{Buffett}\text{MKT} = \alpha + m\text{MKT} + s\text{SMB} + h\text{HML} + u\text{UMD} + b\text{BAB} + q\text{QMJ} + \varepsilon.$$

We required at least 60 monthly observations to run the time-series regressions. The regression coefficients have the same interpretation as those in column 3 of Table 4, with the exception that the market loading is reduced by an amount equal to $\beta^{Buffett}$. The right-hand side excluding the alpha and the error term captures Buffett's *active* stock-selection tilts:

$$\tilde{r}^A_{t+1} = m_t\text{MKT}_{t+1} + s_t\text{SMB}_{t+1} + h_t\text{HML}_{t+1} + u_t\text{UMD}_{t+1} + b_t\text{BAB}_{t+1} + q_t\text{QMJ}_{t+1}.$$

We rescaled this active return series to match Berkshire's idiosyncratic volatility to simulate the use of leverage and to counter any attenuation bias:

$$r^{Active}_{t+1} = \tilde{r}^A_{t+1}\frac{\sigma_{t,I}}{\sigma_{t,\tilde{r}^A}},$$

where $\sigma_{t,I}$ is Berkshire's idiosyncratic volatility, estimated from data up to month t.

Finally, we added back Buffett's market exposure and the risk-free return, r^f_t, to construct our systematic Buffett-style portfolio:

$$r^{Buffett\ style}_{t+1} = r^f + \beta^{Buffett}_t\text{MKT}_{t+1} + r^{Active}_{t+1}.$$

Note that in our notation, the subscript t indicates that quantities are known at portfolio formation date t.

Our systematic Buffett-style return, $r^{Buffett\ style}_{t+1}$, is a diversified portfolio that matches Berkshire Hathaway's beta, idiosyncratic volatility, total volatility, and relative active loadings. These portfolios use only information up to month t to construct portfolio weights for the next month, $t + 1$.

We similarly constructed a Buffett-style portfolio from the loadings and volatility of Berkshire's public and private equity holdings. Table C1 shows the results for these Buffett-style portfolios.

In addition to the systematic long–short portfolios, we computed a long-only, unleveraged, systematic Buffett-style strategy. At the end of each calendar month, we sorted securities on the basis of the portfolio weights corresponding to our active tilts, r^{Active}_t, from data up to month t and constructed an equal-weighted portfolio that held the top 50 stocks with the highest portfolio weights, also shown in Table C1.

Table C1. Implementable Buffett-Style Strategies

	Buffett-Style Portfolio			Buffett-Style Portfolio Long Only		
	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings	Berkshire Hathaway	Public US Stocks (from 13F filings)	Private Holdings
Sample	1981–2017	1985–2017	1989–2017	1981–2017	1985–2017	1989–2017
Beta	0.64	0.68	0.29	0.83	0.83	0.93
Average excess return	39.9%	19.3%	20.8%	9.6%	9.3%	8.6%
Total volatility	29.6%	18.4%	27.5%	13.4%	13.5%	14.8%
Idiosyncratic volatility	28.0%	15.2%	27.2%	4.4%	4.5%	5.7%
Sharpe ratio	1.35	1.05	0.75	0.72	0.69	0.58
Information ratio	1.24	0.90	0.68	0.65	0.55	0.24
Leverage	6.93	3.84	6.03	1.00	1.00	1.00
Subperiod excess returns						
1976–1980						
1981–1985	87.1%	41.2%		17.8%	22.8%	
1986–1990	27.9	10.2	41.3%	9.9	11.4	3.3%
1991–1995	63.9	32.5	54.5	14.5	12.9	15.5
1996–2000	42.0	22.3	17.6	12.8	18.4	15.6
2001–2005	29.8	17.6	12.5	–1.2	–2.4	1.3
2006–2010	0.2	4.1	–10.3	1.9	0.8	0.7
2011–2015	35.1	22.3	22.4	11.4	11.5	10.8
2016–2017	45.2	34.1	18.0	14.9	14.7	14.0

Note: See the notes to Table 2.

This page intentionally left blank

PART II

Asset Allocation: Diversify, Diversify, Diversify

Why *Not* 100% Equities
The 5 Percent Solution

This page intentionally left blank

Why *Not* 100% Equities

A diversified portfolio provides more expected return per unit of risk

Clifford S. Asness

Thaler and Williamson (T&W) ask a provocative question in "College and University Endowment Funds: Why Not 100% Equities?" [1994]. Although they talk about endowments, their question is really more general. T&W present strong evidence documenting the historical superiority of investing in 100% equities compared to a more common investment policy of 60% equities and 40% bonds (60/40).

Documenting the historical superiority of 100% equities versus 60/40 does not completely answer T&W's original question. A recommendation that endowments invest in 100% equities instead of a 60/40 blend actually mixes two distinctly different recommendations: 1) a recommendation that endowments should take more risk than 60/40, and 2) a recommendation about how to implement this riskier strategy (i.e., 100% equities).

Whether a long-term investor should take more risk is a fascinating and sometimes contentious subject that I do not address. The question of whether one should invest 100% in equities is the subject of my article. Recommending 100% equities ignores the benefits of diversification, and even a long-term investor who agrees with recommendation (1), and therefore wishes to take more risk, should generally not own 100% equities.

I argue that an investor willing to bear the risk of 100% equities can do even better with a diversified portfolio. A diversified portfolio historically delivers more return, while not increasing risk (measuring risk along several different dimensions). This is shown most clearly when an investor is willing to lever, although even without leverage, 100% equities would rarely be optimal.

Furthermore, regardless of which portfolio is ultimately chosen, this article argues that choosing how much risk to bear, and constructing a set of portfolios with the most

Reprinted from *The Journal of Portfolio Management*, Winter 1996, Vol. 22 (2), with permission.

Exhibit 1 Stocks Versus Bonds Versus 60/40

expected return for a given amount of risk, are separate tasks. Choosing a portfolio of 100% equities based on their historical realized return misses this separation.

The Case for Equities

Exhibit 1 follows T&W and displays the value of a dollar under three different investment strategies as it grows from January 1, 1926, to December 31, 1993. The top strategy is to invest only in the S&P 500. The bottom strategy is to invest only in a portfolio of long-term corporate bonds. The middle strategy is to invest in a portfolio consisting of 60% the S&P 500 and 40% long-term corporate bonds (rebalanced monthly to 60/40).[1]

Evidence like that presented in Exhibit 1 is commonly used to argue for the superiority of equities, and T&W's empirical evidence is very much in this spirit. Over this period (1926–1993), a dollar invested in equities grew to $800, a dollar invested in the 60/40 portfolio grew to $330, and a dollar invested in corporate bonds grew to a paltry $40. Why would a rational investor ever invest in bonds?

Other arguments for 100% equities involve looking at the probability of outperformance over a given interval of time, and as the length of the interval goes to infinity. Over only a few ten-year intervals do equities underperform 60/40, and over even fewer twenty-year intervals do they underperform. Mathematically, if equity returns are drawn from a sample of past equity returns, then as the investment horizon grows infinite, equities' probability of outperforming goes to one. Again, the superiority of equities seems assured.

Modern Finance

Perhaps the most important lesson of modern finance is that under certain assumptions, the choices of 1) which risky assets to hold, and 2) how much risk to bear are independent choices. Under some simple assumptions, an investor chooses a portfolio of risky assets (in this case stocks and bonds) to maximize the portfolio's Sharpe ratio.[2] This ratio is the portfolio's expected return less the risk-free rate, divided by the portfolio's standard deviation.

Given this maximal Sharpe ratio portfolio P, the investor then chooses the proper mixture of P and risk-less cash. This mix will vary from investor to investor because of differing tolerances for risk, but the relative weights among risky assets will stay constant. Feasible portfolios that maximize expected return for a given amount of risk are said to be efficient.

Exhibit 2 presents some summary statistics for portfolios consisting of stocks and bonds from 1926–1993. The compound returns in Exhibit 2 restate the results of Exhibit 1 — equities dramatically outperform bonds, and also clearly outperform the 60/40 portfolio. This is not an apples-to-apples comparison, however.

Stocks are represented by the S&P 500. Bonds are represented by the Ibbotson total return series for long-term corporates. The 60/40 portfolio is a combination of 60% the S&P 500 and 40% long-term corporates, rebalanced back to 60/40 every month. The compound return is the annualized total return from each strategy, assuming monthly returns are reinvested in the strategy. The standard deviation is the annualized standard deviation of monthly returns over the 1926–1993 period.

Stocks are represented by the S&P 500. Bonds are represented by the Ibbotson total return series for long-term corporates. The 60/40 portfolio is a combination of 60% the S&P 500 and 40% long-term corporates, rebalanced back to 60/40 every month. The levered 60/40 portfolio invests 155% each month in the 60/40 portfolio, and –55% each month in the one-month T-bill.

Stocks have been more volatile than bonds over this period, and more volatile than the 60/40 portfolio.

Constructing a new portfolio makes this comparison more fair. Imagine an investor has already determined that 1) the 60/40 portfolio is the optimal portfolio of risky assets, and 2) the desired amount of risk is the same as a 100% stock portfolio (where risk means standard deviation). For a $1 investment, a new portfolio can be constructed by purchasing 20.0/12.9 = 1.55 dollars of the 60/40 portfolio, and financing this with 55 cents of borrowing.[3]

Exhibit 3 repeats Exhibit 2, including this portfolio called "levered 60/40." Exhibit 4 repeats Exhibit 1 including the levered 60/40.

Exhibit 2 Portfolio Statistics 1926–1993 (%)

Portfolio	Compound Return	Standard Deviation
100% Stocks	10.3	20.0
100% Bonds	5.6	6.8
60% Stocks, 40% Bonds	8.9	12.9

Exhibit 3 Effect of Leverage (%)

Portfolio	Compound Return	Standard Deviation
100% Stocks	10.3	20.0
100% Bonds	5.6	6.8
60% Stocks, 40% Bonds	8.9	12.9
Levered 60/40	11.1	20.0

Exhibit 4 Stocks Versus Bonds Versus 60/40 Versus Levered 60/40

An equivalent-risk portfolio of stocks and bonds outperforms a 100% stock portfolio over the 1926–1993 period. While a 100% stock portfolio grows to $800, the levered 60/40 portfolio grows to $1,291. Even though a 100% bond portfolio grows to only $40, using bonds in conjunction with stocks and leveraging leads to an investment that grows to $1,291, while the investor who owns 100% stocks must bear the same risk and receive only $800. So much for bonds being inappropriate for long-term investors.[4]

Are the Risks Really Comparable?

Perhaps standard deviation is not the proper measure of risk for a long-term investor. We can compare 100% stocks to levered 60/40 using some other possible risk measures.

Worst Cases

Leverage is often thought of as imprudent or "inviting disaster," so in evaluating a levered strategy, looking at worst cases might be particularly appropriate.

Exhibit 5 presents the worst month and worst year for 100% stocks and for the levered 60/40 portfolio over both the entire period (1926–1993) and the post-war period (1946–1993). For each comparison, except the worst post-war year, the levered 60/40 portfolio actually outperforms the 100% stock portfolio. In the worst post-war year, the levered

Exhibit 5 Worst Cases

Period	Portfolio	Worst Month	Worst Year
1926–1993	100% Stocks	9/31 (–29.7%)	7/31–6/32 (–67.6%)
1926–1993	Levered 60/40	9/31 (–27.8%)	7/31–6/32 (–66.6%)
1946–1993	100% Stocks	10/87 (–21.5%)	10/72–9/73 (–38.9%)
1946–1993	Levered 60/40	10/87 (–17.2%)	10/72–9/73 (–44.2%)

Exhibit 6 Percent of Outperformance

Portfolio	% of 10-Year Periods Outperforming Unlevered 60/40	% of 20-Year Periods Outperforming Unlevered 60/40	% of % of 30-Year Periods Outperforming Unlevered 60/40
100% Stocks	76.9	89.8	100.0
Levered 60/40	76.4	88.5	100.0

60/40 portfolio underperforms the 100% stock portfolio by 5.3 percentage points, while in the worst post-war month it outperforms by 4.3 percentage points. It is clear that looking at worst cases is not going to rescue 100% stocks.

Probability of Outperformance

Much of the long-term argument for 100% equities rests on their probability of outperforming over a long period. Exhibit 6 looks at the percent of rolling ten-year, twenty-year, and thirty-year periods that 100% stocks outperforms the original unlevered 60/40, and the percent of times the levered 60/40 portfolio outperforms unlevered 60/40.

The results for levered 60/40 are entirely comparable with the results for 100% stocks. Thus, the levered 60/40 portfolio also achieves the much ballyhooed long-term consistency of 100% equities.

I have already shown that the levered 60/40 portfolio has higher total returns, comparable standard deviation, and comparable "worst cases" to 100% equity. It is comforting to know that one doesn't give up long-term consistency to achieve these virtues.

What If You Can't or Won't Lever?

Modern finance says first construct the optimal portfolio of risky assets, and then choose how much to either lend or borrow of the riskless asset. A levered portfolio of 60% stocks and 40% bonds outperformed a 100% stock portfolio from 1926–1993, but it could be argued that this is unrealistic because many investors either can't or won't lever.

If investors can't risklessly lever, each investor will no longer choose to own the same portfolio of risky stocks. The investment process, however, can still be thought of as having two steps: 1) constructing a set of efficient portfolios (portfolios with the most expected return for a given amount of risk), and 2) choosing which of these efficient portfolios to hold.

Exhibit 7 adds small stocks to the assets we have already analyzed. We see that a 60/40 portfolio of small stocks and bonds outperforms 100% stocks with less risk and no leverage. Bringing small firms into the picture is somewhat arbitrary, and certainly introduces the possibility of data-mining, but the superiority of small firms is not the point of Exhibit 7.

Although the 100% stock portfolio has outperformed the 60/40 portfolio, it is not necessarily the portfolio that does so with the least risk. Even if investors can't lever, they

Exhibit 7 Returns Including Small Stocks

Portfolio	Compound Return	Standard Deviation
100% Stocks	10.3	20.0
100% Bonds	5.6	6.8
100% Small Stocks	12.4	30.7
60% Small Stocks, 40% Bonds	10.7	19.1

should still choose a portfolio with the least risk given an expected return. Perhaps it is small firms, or perhaps some other portfolio is optimal, but the point is the same.

Finally, if investors believe in diversification and wish to take more risk than that of the assumed optimal 60/40 portfolio, but can't or won't lever, they can look to T&W for a solution. T&W describe what to do with an investor who wishes to try to time the market. They point out that these investors should have more than 100% equity exposure when they particularly like the market. T&W do not recommend explicit leverage to achieve more than 100% equity exposure; rather they suggest implicit leverage through stock index options. If investors want to avoid explicit leverage, but still take advantage of the results of this article, they could follow a similar options-based strategy.

What If You Have a Long Time Horizon?

You might want to argue about the use of annualized monthly standard deviation as a measure of risk. There are cases where this point has merit.

If stock and bond returns are independent from period to period, and investors are risk-averse with constant relative risk aversion, then a longer time horizon changes nothing. The relative weight of stocks and bonds would be constant across otherwise similar investors with different horizons. The solution changes only if the volatility of stocks increases more slowly with time than does the volatility of bonds.

There is some evidence that stock returns do in fact exhibit negative long-run autocorrelation, and thus volatility does go up more slowly than if returns were independent (see Fama and French [1988]). Many have challenged the existence of this negative autocorrelation, but if it does exist, it could lead to stocks being a larger part of optimal longer-horizon portfolios (see Samuelson [1992] or Kritzman [1994]).

Siegel [1994] presents an argument along these lines. He shows efficient frontiers obtained from combining a portfolio of stocks with a portfolio of bonds (he assumes that there is no riskless asset). He constructs these portfolios (using data from 1802 to 1992) assuming a variety of different time horizons, and assuming that equity returns are not independent through time.

Siegel shows that as time horizons grow longer, under his assumptions, a risk-averse investor would put a larger and larger fraction of wealth in equities. In fact, for his moderately risk-averse investor, a levered position in stocks is called for if the investor's horizon is thirty years.

Siegel recognizes that bonds and stocks are both part of an efficient portfolio. He simply argues that the frontier of efficient portfolios changes as horizon lengthens. In fact, for any case Siegel examines, a position in bonds exists in almost every portfolio.[5]

Under Siegel's assumptions, for some investors with certain time horizons this position in bonds could be a short position. That's all right — the main goal of this article is not to sell bonds (my tide not withstanding), but rather to defend the separation of portfolio construction (i.e., defining the efficient set) from the choice of how much risk to bear.

One Realization versus Equilibrium

One last interesting thing to consider is the combined effects of 1) observing only one realization of historical returns, and 2) survivorship bias, on the arguments at hand. I have shown that historically a diversified portfolio of stocks and bonds is optimal, but it is still an open question as to why equities have done as well as they have. This is known as the equity premium puzzle (see Mehra and Prescott [1985]).

If the historical results had shown that 100% equities dominates any combination of equities and bonds, including optimally levered combinations, would we all then agree to avoid bonds going forward? Realized returns are always a combination of expected returns and unexpected returns. Furthermore, all we get to observe is one long history of these returns. It is quite possible that equities' observed stellar performance is the result of observing one good realization *and* the presence of long-term market survivorship bias.

For estimating expected future returns, one alternative to history is equilibrium. In equilibrium, expected returns are set so that the market clears. Bonds represent part of the total portfolio of invested wealth, and thus their expected returns should be set so that they are held in the market portfolio. In other words, in equilibrium one holds a diversified portfolio of stocks and bonds (and everything else).[6] Thus, looking at no data, we (ex ante) expect an investor to own bonds.

Say history had been different, and the levered 60/40 portfolio was beaten by 100% equities. A believer in equilibrium might still be unconvinced. The combination of observing only one realization, survivorship bias, and a prior belief in equilibrium raises the hurdle for abandoning bonds. The fact that historically bonds have in fact added value to a diversified portfolio only strengthens the case against 100% equity.

Conclusions

Given the ability to lend or lever, the decision on the optimal mix of risky assets and the decision on the amount of risk to take are separable. Furthermore, this optimal mix of risky assets is the same for each investor.[7] Given no ability to lever, the optimal mix of risky assets can be different for different investors, although construction of the set of efficient portfolios and the choice of which efficient portfolio to hold can still be separated. Under either set of assumptions, it is only under very special circumstances that 100% stocks is the optimal portfolio.

The argument that endowments, or other long-term investors, don't take enough risk is neither damaged nor supported by my analysis. I challenge only the recommended implementation of 100% equities, and only when it is based on realized return superiority and not on an analysis of return versus risk.

Clearly, an investor could believe both 1) in the benefits of diversification, and 2) that long-term investors should take more risk. This investor would then want to act

accordingly: i.e., own a high-expected return, optimally diversified portfolio. The issue then becomes how to capture the benefits of diversification in a practical manner, while retaining high expected returns.

While I address many practical concerns, additional research is called for on implementing this diversified high risk/return portfolio. Issues involving leverage, taxes, and bankruptcy risk are all interesting and deserve further work. Ways to avoid explicit leverage such as buying bonds along with riskier-than-normal equities, and employing option strategies are all promising avenues. Furthermore, the best implementation will in all probability be different for different types of institutional and individual investors.

In an answer to Thaler and Williamson's question, "Why Not 100% Equities?", I argue that the two steps necessary in creating a portfolio must be examined separately. Step one is constructing the feasible set of efficient portfolios (portfolios that are investable and have the most expected return per unit of risk). Step two is choosing which of these portfolios to own (i.e., what risk/return trade-off maximizes utility).

A long-term investment in 60/40 may, or may not, get step two wrong (i.e., not take enough risk). An investment in 100% equities almost certainly gets step one wrong (it is not an efficient portfolio).

Endnotes

The author thanks Fischer Black, Kent Clark, Tom Dunn, Ken French, Brian Hurst, Bob jones, Larry Kohn, Bob Krail, John Liew, Paul Samuelson. Jeremy Siegel, Larry Siegel, Ross Stevens, Steve Strongin, and Richard Thaler.

[1]All data in this paper come from Ibbotson Associates. We follow Thaler and Williamson and focus on the 60/40 portfolio as an alternative to 100% equities. Of course, 60/40 is itself not necessarily optimal.

[2]See Sharpe [1994] for an explanation of Sharpe ratios.

[3]I assume borrowing is done at the one-month T-bill rate. The 55 cents of borrowing is determined in-sample as the amount of leverage that creates a 60/40 portfolio with the same standard deviation as 100% stocks. Alternatively, if the analysis is started in 1931, and always uses the last five years' standard deviation to form the levered 60/40 portfolio (i.e., out-of-sample), the forthcoming inferences do not change.

[4]My example is not countered by more conservative borrowing assumptions. For instance, if I assume borrowing at the T-bill rate + 50 basis points, the levered 60/40 portfolio falls to a 10.81% compound return, still an advantage of 49 basis points over 100% stocks.

This analysis follows Thaler and Williamson by using the Ibbotson long-term corporate bond series. In principle, a market index bond return, not a long-term bond return, would be preferred. It is comforting that my conclusions are considerably strengthened if the long-term corporate series is replaced with an intermediate-term bond series.

[5]One point on the continuum of efficient portfolios would indeed be 100% equities.

[6]In particular, this is implied by the capital asset pricing model. Other equilibrium models do not necessarily imply that investors all hold the same portfolio. In all equilibrium models, however, the market clears and every "net long" asset gets held by someone.

[7]Certain assumptions (multivariate normality, for instance) are necessary to obtain the result that investors choose mean-variance efficient portfolios.

References

Fama, E.F., and K. French. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economics*, 96, No. 2 (1988), pp. 246–273.

Kritzman, Mark. "About Time Diversification." *Financial Analysts Journal*, January-February 1994, pp. 14–18.

Mehra, Rajnish, and Edward C. Prescott. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 15 (1985). pp. 145–161.

Samuelson, Paul A. "At Last, A Rational Case for Long-Horizon Risk Tolerance and for Asset-Allocation Timing?" In Frank Fabozzi and Robert Amott, eds., *Active Asset Allocation*. 1992, Ch. 19.

Sharpe, William F. "The Sharpe Ratio." *Journal of Portfolio Management*, Fall 1994, pp. 49–58.

Siegel, Jeremy J. *Stocks for the Long Run*. Burr Ridge, IL: Irwin Professional Publishing, 1994.

Thaler, Richard H., and J. Peter Williamson. "College and University Endowment Funds: Why Not 100% Equities?" *Journal of Portfolio Management*, Fall 1994, pp. 27–37.

This page intentionally left blank

The 5 Percent Solution

Clifford Asness and Antti Ilmanen

The Much-coveted 5 Percent Real Rate of Return Is Difficult to Achieve, but for Investors Willing to Use Derivatives and Leverage There Is a Potential Way to Do It.

Institutional Investors Are in a Quandary

They commonly target 5 percent real annual returns, or 7 to 8 percent nominal returns. Starting from today's prices for stocks and bonds, the likelihood of actually achieving those returns is low.

When economist John Maynard Keynes was criticized for his shifting policy views, he is believed to have responded: "When the facts change, I change my mind. What do you do?" Unfortunately, institutional investors have been reluctant to publicly accept the new fact of those inconveniently low market yields. Instead of facing the grave reality of past promises being unaffordable in a low-return environment, these investors — or their sponsors — use overly optimistic return expectations as a convenient way to keep kicking the can further into the future.

In recent years some investors have gingerly lowered their long-run return targets, but few institutions outwardly expect less than a 4 percent real return or 6 to 7 percent nominal return on their overall portfolios. Over the past decade and a half, such expectations have generally not been fulfilled, and most investors will likely be disappointed yet again over the coming decade. In fact, those with simple, traditional portfolios like 60–40 U.S. stocks and bonds are even more likely to be disappointed going forward.

Sadly, we cannot change these facts. In the following pages we document the challenge (record-low forward-looking yields), review common institutional responses (highly equity-centric portfolios) and give our recommendations (more effective diversification). Specifically, we propose balanced risk allocations across traditional market premia, truly diversifying return sources from liquid alternative risk premia, and improvements in portfolio construction and risk control. Admittedly, some of these suggestions entail the use of direct leverage, which is an obvious risk, although one that needs to be compared with

Reprinted from *Institutional Investor*, May 2012, with permission.

the near-complete concentration in equity market risk found in most institutional portfolios. We believe that prudent use of leverage is a manageable and rewarding risk to take. We think the ideas presented above can give the investors who adopt them their best chance of getting close to 5 percent long-run real returns.

Current market yields and valuations make it very unlikely that traditional allocations will achieve 5 percent real return in the next five to ten years. These forward-looking measures have been correctly sending the same message since the late 1990s. They are sending an even stronger message today.

As a proxy for expected long-term real returns, we compute the prospective real yield of the traditional 60–40 U.S. stock and bond portfolio. Our estimate of the real equity yield is a simple average of (i) the smoothed earnings yield (the so-called Shiller price-earnings ratio, inverted to become a yield) and (ii) the sum of the current dividend yield and 1.5 percent, an assumed real rate of growth for dividends per share. The real bond yield is the difference between the long-term Treasury bond yield and a measure of long-term expected inflation. The figure below ("Time to Get Real") presents the 60–40 weighted average of the two real yields since 1900.

Model Citizens Alternative Portfolios Share One Common Risk

Of course, we are not the only ones recommending something other than a 60–40 portfolio of stocks and bonds. The low-return environment has changed common institutional investment practices since the 1990s. Some pioneers, like Yale University, adopted the "endowment model" and diversified into various alternative asset classes, combining reliance on the equity premium with faith in illiquidity premia and in hedge fund alpha. This investment model gained popularity in the early 2000s after the equity market bust. By now the majority of U.S. pension funds and other institutions invest some portion of their portfolios in alternatives.

The degree of external management varies widely across institutions and approaches. Most pension funds and endowments have relied on external managers while raising the share of passive mandates over time. However, some major institutions with large internal staffs have increasingly relied on in-house management. For the traditional 60–40-type allocation, Norway's Government Pension Fund Global has led this trend, and in private markets and other illiquid assets, some large Canadian pension plans have been at the forefront.

Based on the above distinctions, the main alternatives to the externally managed 60–40 investment model could be called the Yale, Norway and Canada models. Yet they all share one important commonality with 60–40: Their performance still depends largely on equity market direction, though some of that may be masked by the use of private investments. Thus, although the return experience over the past decade has been mixed, the diversification experience has been disappointing, as all of these portfolios have moved surprisingly in sync.

We firmly believe that a very different approach gives investors who employ it a more reliable way to achieve relatively ambitious return targets, while mitigating large losses that could lead to procyclic capitulation (throwing in the towel when everyone else is throwing in the towel). Our approach emphasizes effective diversification and less concentrated equity market risk. To achieve this goal, some use of direct leverage is needed. In contrast, most common investment models choose concentration and avoid direct leverage, even while embracing embedded leverage: leverage that is built into the securities, like equities, or even into the investment structure, like private equity.

– C.A. and A.I.

Until the 1990s it was relatively easy to achieve 5 percent long-run real returns. The long-run average real yield since 1900 is 5 percent, and realized returns matched the promise of this prospective return as the 60–40 portfolio delivered, on average, close to a 5 percent real annual return. Indeed, this historical experience may have contributed to the 5 percent real return becoming such a widely used target for institutional investors with 60–40-like portfolios. (Skeptics might note that trading costs and fees, not included here, were higher in the past and would have reduced realized returns in past decades even more than now. Moreover, 60–40 only evolved into an institutional standard over time; until the 1960s stocks were considered speculative investments.)

Unfortunately, that favorable environment belongs to the previous century. Since 1998 the ex-ante real yield of 60–40 has been below 3 percent most of the time, making the task of investors that much harder. At first — say, during the period of 1998 to 2000 — most investors took no notice because stock markets boomed and return prospects were wrongly judged on past performance, extrapolating the future long-run equity premium from what it had been in the past, rather than on prospective yields, which at the same time were falling as a result of excessive valuations. In addition, while equities were getting very expensive (low yields), bond yields were relatively high. After the tech boom turned into a bust and bond yields fell, investors began to pay attention to forward-looking returns, but hardly enough.

Currently, the prospective real yield on the 60–40 portfolio is 2.4 percent, its lowest level in 112 years. Roughly speaking, the ex-ante real yield on stocks is 4 percent and bonds is zero percent — both below their long-run average levels, with bonds well below.

There are really only three possibilities going forward. First, that there has been a permanent change in fundamentals, such that real equity earnings growth will be sustainably stronger than in the past, making expected real returns from here much higher than the equity yields imply. In the second scenario these yields will remain near their current levels, and the expected real return of the 60–40 portfolio will be permanently much lower than it was in the past. In the third scenario these yields revert toward their historical averages, delivering higher prospective returns after the reversion, but as this occurs investors will experience a period of even worse realized returns as rising asset yields cause capital losses. Both the second and third scenarios strike us as plausible, but we are skeptics of the first, while recognizing that dreams die hard.

Of course, there will be market rallies, some quite substantial, along any of these paths, but that does not change the fact that these three are the only possible long-term future scenarios.

It may surprise some that the prospective real return on 60–40 is now at a record low. After all, weren't things worse, for instance, near the peak of the tech bubble, when equities were so expensive? Though equities are expensive today, they were far more expensive in early 2000 (in the sense of high prices versus fundamentals, yielding a low expected real return). Yet in early 2000 bonds were quite cheap on a historical basis. Today is relatively unique in that both stocks and bonds are expensive at the same time.

Notice that throughout our history, while sometimes low and sometimes high, the prospective real return on 60–40 has never gotten seriously close to negative. However, there is another asset whose near-term outlook is even worse and in fact negative. The lower thin line in the figure below shows the ex-ante expected real yield on cash (defined as Treasury bills less a short-term measure of inflation). This cash yield is currently near record lows, –2.1 percent, reflecting the central bank's attempts to stimulate the economy by pushing investors into riskier assets or providing cheap financing for direct investments. This effort

Source: AQR Capital Management.

sustains the low prospective return environment for all kinds of assets, but it can also offer direct benefits for those who can borrow (finance their positions) at such low rates. A key distinction, at least for the short term, is that one could borrow at a real rate of –2.1 percent and buy a 60–40 portfolio with a real yield of 2.4 percent, raising the prospective return on this levered portfolio to 4.5 percent. This positive carry on risky assets may balance some of the valuation concerns, for the time being.

Investors have two broad choices in how to respond to the stark news delivered above. They could take a very long-term view and accept that the 5 percent real return target is unlikely to be achieved in the next five to ten years but perhaps is still a reasonable very long-term goal, making plans according to these lower expectations. Or they could take action. This article is about a set of actions, based specifically on our recommendations, of course! In particular, we recommend:

- Harvesting a broad set of return sources, far broader than the typical set that relies heavily on the equity risk premium;
- Implementing a series of portfolio management methods we label "alpha in portfolio construction";
- Putting in place the risk control necessary to see this, or any approach, through the tough times.

We freely admit that all our recommendations fall into a strange category. They are all "alpha" in the sense of a deviation from the market portfolio of wealth that we expect will add to performance, stability or both. But one should not pay alpha prices for the things we propose. Some are strategic allocation recommendations, or bets on a broad set of systematic strategies, while others are techniques for combining these allocations and managing their risks. In addition, as in all forms of alpha, they are zero-sum. We do not claim, nor could anyone claim, to fix the problem of low expected real returns for everyone. But on the other hand, we are not just pointing to the magical, expensive alpha of outperformance and saying, "Solve the problem by adding 2 to 3 percent a year." That is not a recommendation;

it's a hope, and typically a very elusive and expensive one. We believe our recommendations are concrete — clearly not doable for all investors at once but certainly doable for a large subset and at low cost.

We now delve deeper into our first recommendation: identifying diverse return sources and harvesting them cost-effectively. Later sections will discuss the value added from portfolio construction and risk control. It is useful to think of the return sources of any portfolio as a pyramid with three layers, starting from the base, with the highest-capacity and lowest-cost sources, and moving up to the top, with the lowest capacity and highest costs.

- Market risk premia form the base and are the rewards for stable long-only holdings in major asset classes. The equity premium is the best known and most important, while other examples include the term premium (what long-term bonds earn over cash), credit premium and commodity premium. These premia have high capacity and low fees when accessed, say, through index funds.
- The middle layer has attracted increasing attention as investors have learned to appreciate that much of what is marketed as alpha can be better understood as systematic alternative risk premia. Some of the many examples include the value premium, the premium to basic convertible or merger arbitrage, the premium to carry strategies and the extra expected return from accepting illiquidity.
- True alpha is elusive and scarce, and is the top of the pyramid. It has inherently low capacity and is a zero-sum game, so it must be earned at the expense of other investors. If true alpha is found, it should justify higher fees.

The average CIO faces myriad choices when constructing a truly diverse portfolio, but it is really simpler than it's often made out to be. We have recommendations for each of the three layers.

At the base layer of market risk premia, we believe risk-balanced allocations should be favored. Instead of depending mostly on the equity premium, as a 60–40 portfolio does, investors should consider putting together what has come to be called a risk parity portfolio, in which the importance of several market risk premia are balanced. One liquid version of a risk parity portfolio includes one third of the risk budget in global equities for growth, one third in global government bonds for deflation protection and one third in real assets (commodity futures and inflation-linked bonds) for inflation protection. Such a portfolio can offer more robust performance across macroeconomic scenarios than a 60–40 portfolio, which excels in a strong growth/stable inflation scenario and tends to suffer amid weak growth and either high inflation or deflation. More generally, the risk parity portfolio tends to offer a higher risk-adjusted return over the long term.

Importantly, the goal of risk parity portfolios is balanced risk allocation, not balanced dollar allocation. This results in more effective diversification than found in risk-concentrated portfolios (such as the equity-dominated 60–40) and can lead to lower volatility and higher risk-adjusted returns. This volatility reduction may be capitalized — converted into higher returns — by moderate use of leverage.

In addition to very thorough global diversification, many risk parity portfolios are also more effective at managing risk through time, applying the same "risk not dollars" approach by leveraging less when market risk is higher. In the extreme, a risk parity portfolio might employ no leverage in times of exceptional market risk.

Investors can, of course, overlay tactical asset-class views on top of their risk parity portfolio. We do not argue that the parity portfolio is always conditionally optimal, but we believe strongly in risk parity as a better strategic base for such tactical views than a portfolio, such as 60–40, that is quite strongly tilted to benefit only in one type of economic scenario (growth, for example). Long-run success relies more on getting the long-run allocations right than on large timing bets.

We recognize that it may seem contradictory to recommend risk parity when bonds are historically even more expensive than stocks. Again, we stress that we recommend risk parity as a better strategic allocation for the long term. We understand the tactical case for a bond underweight, but any such underweight should start from the superior strategic allocation of risk parity as described above, and we recognize that when starting from superior diversification, market timing is not easy. Moreover, risk parity critics often miss that bonds offer liability and tail hedging services, besides enabling better risk-balancing. Real-world risk parity portfolios are also better diversified and more robust than narrow stock-bond portfolios. Last, risk parity portfolios that target constant volatility over time will cut the sizing of bond positions if bond volatility rises.

In the second layer of the pyramid, alternative risk premia should be preferred to traditional sources of alternative exposure. When most investors think about "alternatives," hedge funds, private equity and various other illiquid investments come first to mind. However, these investments tend to offer more equity market exposure than desired in truly alternative returns. The correlation of both major hedge fund indexes and private equity indexes with global equity markets exceeded 0.8 in the past decade.

Although hedge fund marketing is all about alpha, a drill-down into the industry track record suggests that hedge funds deliver a combination of an embarrassing amount of market risk premia (simply being long stock market risk), alternative risk premia and some alpha — not so much, but by many measures at least positive.

As for private equity, academic research suggests that most funds are even weaker on the key dimensions of performance, risk, liquidity and costs. To clearly outperform public markets, investors need to identify top-quartile managers in advance. Otherwise they are purchasing very expensive, leveraged and illiquid forms of plain market beta.

As opposed to these typical alternatives, we prefer truly diversifying alternative risk premia. Terminology varies, as alternative betas are sometimes called dynamic betas, hedge fund betas, smart betas, or exotic betas. These are long-short strategies that seek to capture the "good" systematic return sources that many hedge funds harvest (as opposed to the equity premium, which is "bad" if it is delivered at 2 percent management and 20 percent performance fees). Certain style strategies — for example, the value tilt, momentum tilt or low-beta tilt — have historically delivered attractive long-run returns in and across virtually all asset classes studied. Other approaches — such as merger arbitrage and convertible arbitrage — are inherently asset-class-specific, but straightforward, diversified versions have delivered strong historical results. Unlike most hedge funds and private equity funds, these dynamic strategies tend to have low correlations with equity market direction.

Seeking alternative risk premia at non-alpha prices is one of the most important ways to help investor performance. These are potentially rewarding and highly diversifying return sources. They are often called alpha, but they really aren't alpha in the sense of unique insight or genius. However, to the extent an investor has little exposure to these return sources, they can be considered alpha in that they are value added to a portfolio and when

implemented correctly are uncorrelated to traditional markets. The key is that they should not be bought for alpha prices.

Alternative risk premia are high-capacity and liquid. They do require some use of what we call the dirty words of finance: leverage, short-selling and derivatives. They do come with their own set of risks. Yet they are time-tested, and we think they should be a big part of long-horizon portfolios trying to achieve 5 percent real returns, as long as they are not accessed through long-biased, high-fee hedge funds or through traditional active management, where they are usually packaged as long-only investments at an implicitly very high fee.

The top layer is alpha. Of course, you should take it if you can find it, but some cynicism is warranted. The overall assumptions about alpha are too heroic for the real world, and, again, it can't save the pie (alpha adds to zero across everyone). Admittedly, our recommendations cannot save the pie either. For everyone who does risk parity, someone else must overweight equities more than the market. For everyone who adds a long-short value strategy, someone must take on more growth risk, and so on. But our recommendations are available in far higher capacity to the investors who would follow them than literal alpha in the conventional sense.

True alpha can still help or even save individual plans, so pursue it based on your own honest assessment of your ability to find it (net of high fees and with open eyes as to whether it really comes from alternative beta premia) and of your possession of the resources required to do so.

One theme we want to emphasize at the risk of forcing double, maybe triple, duty on the word is a very different idea of alpha: that is, alpha in portfolio construction (and risk control to follow). This does not involve the rare skill of traditional alpha at the top of our pyramid but rather represents the skillful combination and management, including cost control, of the various components of the portfolio. Long-run investment success requires identifying attractive return sources, harvesting them cost-effectively, diversifying among them aggressively and overlaying smart risk controls. Many investors focus too much on the first activity at the expense of the others. We believe that portfolio construction, risk management and cost control are the low-hanging fruit of managing a long-term portfolio. It's far easier and more plausible to add impactful value to the whole portfolio net of fees through these concepts than the more typical pursuit of alpha.

Here are some specific suggestions for improving a portfolio (alpha in construction and control):

- The prospectively low-return environment underscores the importance of cost-effectiveness, whatever returns investors are harvesting. When it comes to external management, it is essential to not pay alpha prices when it's not really alpha. Fair fees depend on the return source.
- Allocate by risk, not dollars. To achieve effective diversification, it helps to use meaningful measurement units. Measuring portfolio shares by dollar allocations can be highly misleading. The 60–40 dollar split between stocks and bonds may sound balanced but is actually roughly a 90–10 risk allocation given the greater risk of equities. Risk parity investors have taken this message to heart, but it really applies to every investor. Even if at the end of the day you decide you are comfortable with a 90–10 risk allocation into equities, it is better to invest so with open eyes.

- Ensure that you are building a truly diversified portfolio across investments. Investing in alternative products that are highly market-directional, because many such investments are highly correlated with the equity market, only provides the illusion of diversification.
- Effective diversification almost assuredly requires some amount of leverage, short-selling and derivatives. All should be used in a prudent manner. They involve risks, but we think these risks have proven to be manageable even through some vicious downturns. Return-seeking investors who cannot or will not use these tools are resigned to letting equity market direction drive their portfolio performance.

One of our important messages is that you get to choose your risks — for instance, two paths to high expected returns are concentration in aggressive assets (generally equities) or prudent use of leverage applied to a more diversified portfolio. Both concentration and leverage are risky, and nobody should tell you different. Sadly, you don't get to choose high expected strategic returns without one of these risks. But choosing to concentrate in equities simply because it is the more common choice does not make it any less scary (in the absolute; it may be less scary relative to your peers' making a similar poor choice).

Risk control is ultimately about surviving to the long term and not sabotaging it along the way. Although diversification is central to our investment philosophy, investors may want to supplement it with other means of risk control that help them stick to their game plan when faced with unexpected losses. These means can be quantitative: concentration limits, drawdown control rules, rebalancing to target a stable level of portfolio risk over time, tail hedges (though rarely the explicit purchase of insurance). They can also be psychological: education, preparation and precommitment. What follows is some hard-learned, and occasionally hard-earned, risk control advice for those tough investment environments when spine is truly needed to sustain supposedly long-run allocations.

In the wake of the financial crisis of 2008–09, when many supposedly super-long-term investors acted as liquidity takers, not liquidity providers, we have asked ourselves the obvious question: Why? One simple answer gets to the heart of what it means to be a long-term investor. Long-term investors are often painted as having above-average risk tolerance and a natural edge in being liquidity providers (contrarian buyers of risky assets) in bad times. Both ideas contain a seed of truth, but they may not be right for a single investor.

An investor who chooses to accept above-average risk will sustain above-average losses in bad times and will not be well positioned to buy more when bargains appear. A contrarian investor has to begin a plan in good times and retain some dry powder to have any chance of acting as a contrarian liquidity provider in bad times. Choose which one you want to be — or, at least, what combination you want to be. But don't assume you can be both to the full extent of your endurance without realizing that they will tax you at precisely the same moment. As simple and obvious as it sounds, we think many very long-term institutions "double-counted" their resolve and thought they could be higher long-term risk-takers and liquidity providers in a crisis, and both to the limit of their endurance. But both draw on the same "budget for pain," and this must be recognized up front.

At worst, the result of institutions trying to be both is procyclic capitulation — losing faith and selling risky holdings near the market bottom. We can never truly banish this, as we do not know ex ante the possible depths of a crisis. But by not double-counting

Source: AQR Capital Management.

how long-term you are, essentially thinking you can have a very aggressive strategic asset allocation and be the one to provide liquidity near the nadir, we think you greatly increase the chance it does not happen to you!

Our next concrete recommendation, after not double-counting what it means to be "long term," is to consider your own self-imposed drawdown control methodology on the overall portfolio. Investors generally face three choices in tough times: Never reduce risk because of losses; reduce risk to control losses subjectively, using judgment on the fly; or reduce risk because of losses in a systematic fashion (what we call having a drawdown control methodology) and add risk back using a similarly systematic methodology.

Frankly, we started out our careers with a lot of sympathy for the first approach, particularly for strategies in which value investing is a big part, as value positions often get more attractive after suffering. Although we are still theoretical fans of that method, we have seen too many instances where a resolve never to cut (or even to add to) risk becomes a mad dash to control losses on the fly at the worst possible time. It does seem clear that no matter what their stated plans to hold the line, all investors have some breaking point, and we believe our collective biases can lead this subjective break to happen at the worst times. This is all magnified when leverage or a high-volatility portfolio is chosen.

Let's delve further into the second approach. We are all subject to many of the same investment biases we describe throughout this article, biases that will feel most acute during the toughest times. Although this subjective method may seem somewhat useful for reducing risk when in pain, we have found it to be completely useless or even a detriment when adding back risk after the pain has started to abate. Choosing when to add back risk is at least as hard to manage well as cutting risk in the first place.

All considered, in most practical instances — again, particularly in the presence of leverage or an aggressive portfolio posture — we prefer the third approach, a systematic drawdown control methodology of early intervention and modest risk cuts, in a hope never to have to cut draconically near a bottom, to either plan one or plan two. Note again that this is a form of zero-sum "alpha," as the world as a whole cannot cut risk.

Last, this might be a stretch to call "risk control," but while we're talking about throwing in the towel at the worst time, an additional high-level recommendation would be to alter or at least soften the focus on three-to five-year evaluation periods for managers and styles. These evaluation periods are death to returns, and nobody ever notices.

Financial market data abounds showing short-run (within a year) momentum patterns and multiyear reversal (value) patterns. Yet investors often make asset-class allocation decisions and manager fire-hire decisions using a three-to five-year evaluation period. In short, they act like momentum investors at reversal (value) frequencies. Return-chasing — allocating toward winners or away from losers — at multiyear horizons and procyclic capitulation after disappointing performance are among the most common ways investors can sabotage their own long-run performance.

The figure "Riding the Wave" gives an admittedly self-serving example of the simulated three-year returns of a simple quant long-short strategy since the mid-1950s. Despite a very appealing long-run performance — the dotted line about 5 percent in excess of cash — three-year returns exhibit waxing and waning fortunes. (Note that recent results are well within the historical range.) Few investors have the resolve to stay with a strategy through its waning periods.

Of course, we do not counsel ignoring investment performance forever, nor do we counsel never switching asset classes or managers. All we counsel is that if three- to five-year returns are major criteria, and they often are, they are frequently being used in a statistically backward manner that should be acknowledged and perhaps changed.

Before concluding, let's ask the obvious question: If we're right, why don't more people listen to us?

First, we dismiss the highly unlikely answers that we may be wrong, unconvincing or simply not well liked.

Consider why more people don't leverage a diversified portfolio rather than focus on equities. In theory, it might be optimal for investors to let equities be 90 percent of portfolio risk. If the equity premium offered a uniquely high Sharpe ratio (reward for risk) among return sources, proportionate to its dominant risk, it would be reasonable to depend so heavily on one return source. In practice, however, other asset classes and many long-short investment strategies have historically offered comparable or higher Sharpe ratios. Also, looking ahead, prospective real equity returns are below the long-run average, and it is hard to make a case for a uniquely high reward for this single risk. If investors let equity market direction dominate their portfolio performance, they are doomed to follow the roller-coaster ride of market gyrations with little recourse and a smaller reward than usual.

Yet most investors still choose concentrated equity market risk, despite the better rewards of leveraged diversified portfolios. Investors dislike leverage for both bad reasons (the appearance of speculation) and good ones (any levered portfolio is vulnerable to the danger of having to deleverage at fire-sale prices). This risk is real but can be managed (*Institutional Investor*, May 2010). Concentration risk cannot be managed in any analogous way and, unlike leverage risk, brings a lower risk-adjusted return, not a higher one.

Another main reason for equity domination is familiarity — "everyone does it" — and the resulting lack of peer risk (recall Keynes's safety in failing conventionally). Underlying this familiarity are some fair reasons, such as a strong theoretical basis, as well as extensive empirical evidence over 100 years in numerous countries. These reasons together enable

investors to sustain their long equity bias through several years, or even a decade or more, of disappointing performance. Any other return source may lead to more time inconsistency (a nicer way to say throwing in the towel at the worst time).

So, while still dismissing the unlikely answers initially mentioned above, we think the difficulty in implementing these suggestions comes down to lack of familiarity, the agency problem of failing conventionally being the superior way to underperform and some real, but in our view overdone, concerns about leverage. The bottom line is that without these barriers we think our recommendations, even given their large capacity, would, like all forms of alpha, be largely arbitraged away, so these barriers may be things to lament, but for those with more freedom to act than the average investor, they are also reason to celebrate.

In conclusion, traditional, simple asset-class allocations — say, 60–40 stocks and bonds — are likely not going to make 5 percent real returns from here given that forward-looking real returns are at half this level. The standard universe of "alternative asset classes" is not likely to fill the gap, as it tends to repeat the problem of concentration in equity risk, just at a higher fee. Nevertheless, we believe that some investors can still achieve the stated 5 percent goal, or at least far closer to that, if they embrace a modest amount of innovation, as we detail in this article, and thoroughly prepare themselves to see it through. No single idea will do the trick, but investors should consider a well-balanced combination of market premia and alternative risk premia, pursue alpha through portfolio construction and employ thoughtful risk controls. Each of these can help investors get closer to the 5 percent real return target. Together they may even make the target realistic for some, and the diversity of ideas should give investors a less rocky ride.

This page intentionally left blank

PART III

To Time or Not to Time

Bubble Logic
Fight the Fed Model
Style Timing: Value versus Growth

This page intentionally left blank

Bubble Logic

Or, How to Learn to Stop Worrying and Love the Bull

Clifford Asness[1,2]

A bull market, and the incentives of those who make their living from bull markets, can create its own form of logic. This article explores some of the stories that encourage the purchase or retention of stocks or mutual funds and the logic behind these stories. Some of these stories are honest attempts to explain new phenomenon, and may or may not prove true going forward. Some seem to be unintended falsehoods that come from an incomplete or lazy application of economic reasoning. Finally, some seem less well intended. The stories, and the logical analyses behind them, generally originate with Wall Street (both sell side and buy side), sometimes riding the coattails of academia, and are often readily absorbed by investors engaged in wishful thinking. Such wishful thinking has led to a stock market, and the growth/tech sector of the market in particular, that is priced so expensively that even very long-term investors will probably end up disappointed.

I. Introduction

First, full disclosure. I am a principal of a boutique investment manager with a value orientation. Thus, for about two years now with only brief interruption, we have had our assets handed to us. I believe that before a person rants and raves they should fully disclose that they may be typing with a jaundiced keyboard. Part of this article's thesis is that self-serving incentives often color what passes for independent analysis and research of financial markets, and since I am not immune to this bias, it would be particularly hypocritical of me not to declare my stance at the outset.

That said, my goal is to tick through many of the bromides about investing that currently are conventional wisdom on Wall Street and Main Street. Some of these popular wisdoms are explanations for new phenomenon that may or may not be true but must

Copyright © 2018 AQR Capital Management, LLC. This paper initially appeared on SSRN, August 2000.

be critically examined. While some are correct, but misinterpreted, some are just dead wrong. While some are just benign silliness, many can prove harmful. Like in most businesses, these bromides exist largely to sell a product, in this case to sell equities. Of course, wishful thinking and the human desire for a free lunch makes the consumer/investor very susceptible to this sales pitch. A basic theme is that although Wall Street research is made to look like independent science, and the financial media is made to look like neutral journalism, they are biased towards keeping you buying or holding common stocks. There is nothing wrong with that. That is their business. However, the investor needs to keep his eyes open. Furthermore, it is distinctly possible, and in my opinion likely, that the acceptance of these "wisdoms" has led to a stock market, and the growth/tech sector in particular, that is so expensively priced as to probably disappoint even very long-term investors.[3]

This article is meant to stimulate thought and debate, and should be taken that way. It certainly contains facts, but it also contains a healthy dose of my opinions (occasionally intentionally provocative), and I try my best to distinguish between the two. Some of the areas I address are very topical, while some apply to any era. I certainly do not claim to have all the answers. It is a lot easier to point out the fallacies in others' arguments than to figure out the answers. Still, when fallacies rule the land, somebody has to point at the naked emperor.

The article is divided into six parts. Part I is this introduction. Part II examines the properties of a long-term investment in equities, and the implication of today's high prices for this investment going forward. It is the most mathematical/technical part of the article, but is essential for examining today's stock market. Part III is lighter in tone and rigor, examining some creative ways sometimes used (and abused) to defend today's high stock prices. Part IV briefly examines growth vs. value investing. Part V encompasses several miscellaneous topics that might not fit anywhere perfectly, but clearly fit the theme of bubble logic (the impact of new trading technologies, the effect of stock splits, the selectiveness of mutual fund advertising, and finally, the notion that it is "different this time"). Finally part VI concludes.

Each main part contains sub-sections that begin with a paraphrased quotation summing up the popular wisdom I am about to critically examine. Admittedly, this article is a mix of some hardcore valuation analysis, and a lighter anecdotal survey of the market. While perhaps an awkward mix at times, I argue that both are necessary to understand today's stock market. In general, the most theoretical issues are relegated to the extensive footnotes, which the interested reader can peruse to their taste. Finally, if any readers are more annoyed than amused by some of the sarcastic humor herein (especially in the second half of this article after the heavy lifting topics are done) then I do apologize. Have pity on a partially gored bear.

II. Long-term Equity Investing

"Equities Always Win over the Long-term"

One of the most prominent ideas behind the bull market is that over any long-term horizon (let us identify 20 years as the long-term horizon) an investor cannot lose in the stock market. This belief is clearly held by many investors today. There are many sources for this belief, but perhaps the most recently influential is Jeremy Siegel's article, *Stocks for the Long Run*,

documenting that over long periods the equity market's premium over inflation and over alternative assets like nominally risk-free cash, has been very large and consistent. Siegel's data is rock solid and correct.[4] However, let us look at a graph of some alternative data. This one comes from Robert Shiller (author of the article *Irrational Exuberance*) and shows the price-to-earnings ratio (P/E) of the S&P 500 for over 100 years. Note, the E in this case is the 10-year average of trailing S&P 500 real earnings, not last year's trailing earnings.[5]

First, it is exceedingly important to note that there is nothing magic about the long-term that makes equities pay off so consistently, it is just math.[6] As with any volatile asset the longer the observation period, the more noise tends to cancel out and the more accurately we observe the true average return (the long-run expected or required return of equities). The long-run average equity market return vs. inflation (the real equity return), and the average equity return vs. cash or bonds (the risk premium), should be positive since investors should require a positive expected return for investing in risky or riskier assets. Thus, the twin observations, (1) that average equity returns are positive, and (2) that we become more certain of realizing a positive return over long periods, are not exciting findings, they are what we expect to find. The exciting finding has been exactly how positive these returns have been and, to a lesser extent, that deviations around the average over long periods have been somewhat smaller than we would expect. Siegel finds that the real return on U.S. equities has been about 7% over the long-term, and this is higher than most theories say it should be.[7] The gross compound return of the S&P 500 has beaten inflation in 100% of the 20-year periods from 1926 to 2000 (measured at overlapping monthly intervals), and has beaten U.S. T-bills in 96% of these same periods. Clearly, equities have been strong and consistent long-term performers.

Over any period, including the long-term, the return on an equity investment will be a function of the price you pay for it (let us deal with price in terms of P/E), the price at which you ultimately sell it, the dividends it throws off in between, and the earnings growth over the period.[8] The data underlying the "long-term argument for equities" rests on a period when P/Es (defined as above) averaged about 15, peaked at around 30 (not counting the latest bull market run), and examined on this same scale are now residing in the low- to mid-40s. Other versions of P/E, and in fact just about any other credible valuation measure, tell a similar story.[9] Equities historically always returned a reasonable amount (or even a superb amount) over the long-term. One reason is that they were almost always priced

reasonably. To assume that this long-term consistency will now exist independent of pricing is simply to believe in voodoo. It will happen because it has always happened is not a strong argument. Equities have gone up more than inflation (i.e., a positive real return) in all 20-year periods, and beaten cash (i.e., a positive risk premium) in almost all of these periods, not because of magic, but largely because throughout the period we study they were generally priced reasonably, or even cheaply, vs. their earnings and dividends prospects. That is not necessarily the case anymore.

In an interesting recent paper (working paper, June 2000), Professors Fama and French take on the problem of estimating the market's risk premium (which they define as the expected return of the broad stock market over commercial paper). Among their other results, they find that because of high stock prices today, the expected risk premium of stocks over high quality commercial paper is now approximately ½ –1 ½% (the range accounting for reasonable degrees of optimism or pessimism about earnings and dividend growth going forward). This compares to an average historical risk premium of about 6% from 1872 to 1999.

I am addressing here the simple question of whether equities will always win (beat inflation and/or short-term cash) over the long-term.[10] Fama and French's estimated equity premium of ½ –1½% is an expected value. Real life always varies from expectations, and Fama and French's estimate can be thought of as the center of our future expectations, or put differently, our best guess based on current prices and growth estimates of how the future will turn out (i.e., there is a 50% chance things work out better, and a 50% chance things work out worse).[11] Put simply, the very low current expectation for equities compared to history means there is very little "cushion." If things work out slightly worse than expected, equity returns can now easily be less than commercial paper over the long-term and can even be less than inflation (i.e., negative real returns). When equities were priced for much higher expected returns they had a very large cushion against negative returns. This cushion is largely gone now.

Now, to be balanced, there might be reasons to justify today's very high prices. Long-term (not highly transitory) large increases in real earnings growth, perhaps driven by productivity growth, and the ability to ultimately turn this earnings growth into free cash flow to investors, could justify higher prices (I examine this more later). Some also argue that today's low inflation environment justifies a higher P/E.[12] However, when equity prices (P/Es) are 3 their historical average and far above their prior historical maximum, the burden of proof in the debate is on those who claim they are still a low risk (or no risk) long-term investment. It is not enough to say "they always have been and thus will always be." None of what I say is a proof that equities will do poorly over the next 20 years. All I argue against here is the idea that the stock market *must* do well, and that the price you pay does not matter to the probability of this occurrence.[13] If equities are priced to offer considerably lower expected returns compared to history, then they are far more susceptible to negative shocks that can leave even their 20-year returns lower than short-term cash or even inflation. Returns over shorter time horizons, e.g. 10 years, that might be relevant to many or most investors, are even more threatened. Again, the cushion is substantially reduced (as is the reward even in the expected case).

Finally, let us step back for a moment. Does it fit intuition that as the world has bought into the "stocks cannot lose over the long-term" argument, investors have become price insensitive (why care about price if you cannot lose?), and thus bid up prices to the point where equities suddenly can lose over any term? It sure fits my intuition.[14] Countless times

we see researchers find patterns in the stock market that have existed for a long time and then continued to work for a short time after discovery. But then they go away as victims of their own success. That is, too much money chases the effect. The current price of the stock market seems a prime candidate to be just such a case.

"But, My Estimates of Expected Stock Returns Are Based on Solid Long-term Data"

Fama and French (and many others) find that expected stock returns going forward are lower than stock returns have been in the past. However, when making asset allocation decisions, many investors still estimate expected future returns by using a long-term average of historical returns (the most common period employed is 1926 to the present). On first examination this seems eminently reasonable. However, this suffers from an important paradox. The bull run of the last 5 years (taking us from a P/E around 20 to over 40), and even more so the last 20 years, non-trivially raises the long-term average realized return of stocks, and hence the estimate some use going forward of the expected return. However, this more than doubling in P/Es in 5 years almost assuredly reduces the expected return of stocks going forward. That is, just as users of this method are estimating a higher expected return for stocks, it is in fact lower, and lower for the precise reason they are raising their estimates![15]

Going to extremes can make this issue even clearer. Imagine all stock prices went up 100x tomorrow with no change in fundamentals.[16] Hopefully, we would all agree that paying about a 3000–4000 P/E for the S&P 500 right now would make a very poor long-term investment (wow, a P/E that even the authors of *Dow 36,000* would not love). However, the historical average return on stocks would skyrocket once this 10,000% return was added. Obviously, in this case it would make little sense to use the historical average to forecast the long-term future. The historic average would be incredibly high, and the future would look incredibly poor. While far less extreme, the real-life situation today is analogous. Finally, this issue is very clear when applied to bonds. Few would think a bond's expected return has gone up if its yield falls, yet when its yield falls its price and thus historical average return rises. Stocks are really no different, but all the dust kicked up in equity analysis can obscure this highly analogous relation.

"It Is Silly to Compare Today's P/Es to Those before the Great Depression"

Some question the very act of examining a figure like the earlier long-term P/E chart. They argue that comparing today's P/E to historical averages is misleading, or more prosaically put "driving through the rearview mirror," as times are different now. Today's S&P 500 P/E (using the Shiller data) is about 44, while the historical average P/E from 1872–1999 is about 15. Critics say that it is naïve to assume that we will return to about a 15, as life is better now. As one example, the July 18th, 2000 *Wall Street Journal* had an interesting and thoughtful article on this topic. Quoting Jeremy Siegel from that article,

"When we look back over the past century and say the average price-earnings ratio was 14 we're talking about a period that includes the Great Depression, two world wars, and the double digit inflation of the 1970s. Saying we'll go back to a 14 P/E means saying we have learned nothing about how to better manage the economy." Well, there certainly may be some truth to this observation, but let us dig a little deeper. These are the average P/Es over different periods:

Average P/Es Over Different Periods

Period	Average 10-Year P/E	Average 1-Year P/E
1891-Present	15.5	14.5
1946-Present	16.6	14.9
1970-Present	16.6	15.4
1980-Present	18.5	17.0
June 2000	**43.9**	**32.2**

For balance, I show both the Shiller 10-year P/Es (that correct for significant biases in the 1-year trailing P/E), and the more traditional 1-year trailing P/E (which is also calculated from the Shiller data). Comparing recent P/Es to the longest-term average P/E is startling (43.9 vs. 15.5, and 32.2 vs. 14.5). However, this is not an artifact of including very old data. Looking at the 1946-present averages (no world wars, no Great Depression) the comparison is hardly any different. Even looking 1980–present does not change the story much (and the current euphoric period is getting more and more weight in this shorter-term average). While memory of these cataclysmic events might have depressed P/Es even after they ended, it is certainly not the case that one needs world wars and a Great Depression to get reasonable stock valuations. Rather, even compared to other modern times, it is today's high P/Es that are the exception not the rule.

Interestingly, in this same article, Jeremy Siegel was quoted as saying he thought reasonable P/Es for the S&P 500 might be in the range of 20–25.[17] This can be (and actually was) interpreted as bullish because 20–25 is not 15, and thus Siegel is agreeing (correctly) that it is naïve to say that we must return to the historical average. However, falling from today's 32.2 to 25.0 is a –22% return (I assume Siegel's using 1-year P/Es), and falling from 32.2 to 20.0 is a –38% return. I will examine this more later, but a fall of this magnitude for the broad market, and in all likelihood a much harsher fall in the growth/tech sector, in most circles would be labeled bearish. While I think Professor Siegel's observation is accurate (that there is no real reason we must return to historical average P/Es) and certainly interesting, it is somewhat amazing that this can be interpreted as the bullish case!

Essentially, it is true that nobody should point at the historical average and say we must get back there. That is as naïve as observing that equities have historically done well and assuming they must do well going forward. However, to summarize, two points are important. One, you do not need to include the Great Depression or a world war to get low average P/Es. Today's high P/Es are about 2x to 3x (depending on which P/Es you use) the average P/Es measured only over the modern era (1946–present). Second, as Siegel's quote supports, today's P/Es are still dangerously above reasonable levels (not just vs. history, but from actually examining the math behind expected returns – more on how to do this later).

Finally, I would like to discuss the relevance of the Great Depression. Revisiting Siegel's quote, "Saying we'll go back to a 14 P/E means saying we have learned nothing about how to better manage the economy." Again, perhaps a 14 P/E is an extreme prediction, but it is also precarious to rest too hard on how much we have learned about the economy. Students of the Great Depression uncover tremendous parallels to today, including a belief then that we had learned a lot about managing the economy (I know this talk of ominous parallels makes me sound a bit like Hal Lindsey, but bear with me). The quotes below are all (except for the last one) from a *New York Times* article by Floyd Norris called "Looking Back at the Crash of 1929." The first quote is Norris describing a Times editorial from October 1929 that blasts speculators, but then assures us that the Fed will protect us from the consequences of our own folly:

> ... it may be useful to recall an editorial published in *The New York Times* in the midst of the 1929 crash, on Oct. 26. It heaped scorn on those who had participated in the "orgy of speculation" that had sent prices so high amid talk of a new era and permanently high stock prices. "We shall hear considerably less in the future of those newly invented conceptions of finance which revised the principles of political economy with a view solely to fitting the stock market's vagaries."
>
> But after blasting the speculators, *The Times* took a much more sanguine view of the economy's future. The Federal Reserve had "insured the soundness of the business situation when the speculative markets went on the rocks."

Sounds a lot like the current view that we have a "Greenspan Put" where the Fed will save us from a crash so we can safely trade/invest like a crash cannot happen. On the impact of technology (in the 1929 case it was radio):

> Then, as now, there was talk that an exciting new technology had rendered the old economic laws irrelevant. Then, as now, stock connected to that technology zoomed skyward, but even companies that had nothing to do with the technology saw their stock prices benefit.

Norris's list of parallels continues. Like today, pre-1929 dissenting voices were laughed at, the country was obsessed with stocks, and "dumb money" crushed "smart money":

> By 1929, such cautionary voices had been discredited, and the stock market had become a force unto itself, propelled by dreams – and the reality – of quick wealth. "Playing the stock market has become a major American pastime," reported *The Times* in a magazine article published on March 24, 1929. The article noted that the number of brokerage accounts had doubled in the past two years, and added, "It is quite true that the people who know the least about the stock market have made the most money out of it in the last few months. Fools who rushed in where wise men feared to tread ran up high gains."

Then as now, Wall Street came to the defense of stock prices. The following discusses bankers' reaction to a severe "dip" in early 1929, and the subsequent recovery from this dip that signaled to many that all was well again with the bull market (the parallels to the spring/summer of 2000 NASDAQ recovery make me start looking over my shoulder for four horsemen).

> "Responsible bankers agree," The Times quoted an unnamed broker as saying that day, after the recovery began, "that stocks should now be supported, having reached a level that makes them attractive." The responsible banker in question, it turned out, was Charles Mitchell, the president of National City Bank, a predecessor of today's

> Citibank. He defied the Fed, and lent out all the money the speculators wanted. Soon prices were back on their upward course. By the August peak, the Dow was 35 percent above the low reached during the March sell-off.

Responsible bankers agreed, and choosy moms chose Jif, but we all know what happened next (a tremendous crash and bear market, the Great Depression, and a full 20 years of effectively zero real return on stocks).

Finally, I must add one other quote that makes clear the parallel between today and 1929 both in abandoning traditional valuation methodology, and in assuming the Fed will bail us out of any crisis with easy money:

> Once stock prices reach the point at which it is hard to value them by any logical methodology, stocks will be bought as they were in the late 1920s – not for investment, but to be unloaded at a still higher price. The ensuing break could be disastrous because panic psychology cannot be summarily altered or reversed by easy-money policies.

Note the author's cynicism regarding whether a central bank can actually save us from an overvalued and declining market. The quote is from 1959, by Alan Greenspan.[18]

While the parallels are interesting, there are of course some things that are very different. However, this is not necessarily good news. If somebody asked me the riddle, "what's the single biggest difference between June 2000 and September 1929," I might be compelled to answer "the price." Measured using the Shiller 10-year scale today's P/E is about 43.9 vs. 32.5 in September of 1929, and measured using 1-year P/Es today's value is 32.2 vs. 20.4 in September of 1929.[19] Furthermore, at the end of September 1929, CPI inflation stood near zero, and 10-year bond yields hovered around 4%, so these crutches sometimes used to defend today's high stock prices were even lower back then. Finally, recent real earnings growth (circa early 2000) has been strong. 1-year compound growth has been about 12% (vs. a 3.5% historical average), 5-year annualized compound growth has been about 4.0% vs. a 1.9% historical average, and 10-year annualized compound growth has been 4.5% vs. a 1.5% historical average (the historical average compound growth falls with time-horizon due to the effects of volatility on compound growth). In other words, recent real earnings growth has been strong. However, in September of 1929, the relevant numbers were 18.3% for 1-year growth, 10.2% for 5-year growth, and 5.4% for 10-year growth, all better or much stronger than today. Apparently, then as now, investors were looking at recent growth, and pricing stocks as if this growth would go on forever. Only then, the growth was even stronger, and the price was not as high.

On many qualitative issues (rampant speculation, total faith in the Fed, extreme belief in new technology, Wall Street trying to jawbone a stock recovery, etc.) today seems very much like 1929. On other fronts the comparison is more favorable for today (we probably live in a safer world, security regulation protects us better from outright fraud, etc.). Finally, on the pretty important issue of price, today seems significantly worse than 1929. Now, comparisons to other times are dangerous as things can and do change. I do not want to fall into the "representitiveness bias" documented by students of behavioral science where one over-relies on perceived similarities. To this end, most of the rest of this work focuses on forward looking estimates of stock market returns, not simple historical comparisons. However, while slavish devotion to history makes little sense, and similarities can be overstated, it is perhaps at least as dangerous to completely ignore the lessons of the past.

"Stock Prices Are High as Investors Are Willing to Accept a Lower Risk-Premium Today"

Many reasonable analysts looking at today's equity prices reach the almost unavoidable conclusion (as did Fama and French in the work cited earlier) that the equity market's current prospective expected return is quite low compared to history. However, rather than forecast a severe drop in stock prices (or a long period of price stagnation) that would restore the expected market return to more normal levels, some argue that the lower expected returns are in fact here to stay. The idea is that investors now recognize that stocks are less risky vs. other assets than previously believed, and thus should offer a lower return premium going forward, and thus have higher prices today (remember, a higher price today leads directly to lower expected returns going forward). Actually, this explanation is theoretically reasonable. It might be true, and could explain the very high valuation levels we see today. However, there are at least two giant holes in this argument.

Hole #1: Investors show no signs of accepting lower stock returns going forward. Investor surveys all point to their expected future returns being higher, not lower, than historical experience. It is hard to reconcile equity investors being happy with bond-like expected returns, with the countless ads for on-line brokers implicitly promising you a private island if you will only trade with them. When compounding at ½ –1 ½% over commercial paper it takes a long time to pay off the mortgages on those islands. In addition, the existence of inflation-protected government bonds offering about a 4% guaranteed real return, high quality municipal bonds whose tax equivalent nominal yields currently approach 10%, and very high yields on equity-like low grade bonds makes it even more unlikely that investors would now be consciously willing to accept very low prospective returns on stocks. Apparently, it is not general risk-aversion that is low, but only risk-aversion when it comes to buying equities. Quoting Jeremy Siegel (*Journal of Portfolio Management*, Fall 1999), "This divergence between increased historical returns and lower future returns could set the stage for some significant investor disappointment, as survey evidence suggests that many investors expect future returns to be higher, not lower, than in the past."

Hole #2: As mentioned earlier, a large part of the long-term consistency of stock returns comes from the fact that they have historically had a high average return. If that average (or expected) return goes way down then stocks are going to have some long periods of negative performance as their cushion is gone (see the discussion of Fama and French's findings earlier). Then, if equities can lose, they become risky again, and the circular argument that they should be priced super expensively because they have no long-term risk disappears completely.[20]

It is rational to observe that the historical equity risk premium in the U.S. has been high, and maybe conclude it has been too high.[21] And maybe, just maybe, part of the gigantic bull market we have seen is a permanent rational lowering of this risk premium. However, it strains credulity to explain the majority of this bull market as coming from equity investors now being perfectly happy to make nothing or just a bit over bonds going forward.

"Earnings Growth = Stock Return"

This section examines the long-term returns of companies that are growing fast, and expected to continue this pace for quite a while. I provide examples that demonstrate that the long-term expected return from investing in these companies is not even close to equal to

their expected growth rate. This counters what seems to be a widely held (though thankfully not universal) view that if a company's earnings grow at 30% per year, by investing in it, you will make about 30% a year. More generally, there appears to be a strong belief held by many that to make money investing, one must invest "where the growth is." This is not true. While our general prosperity is certainly linked to the overall economies' ability to grow, this does not mean that investing in specific fast growth companies, or indices of these companies, is automatically a good idea. In fact, it should be immediately clear that any prediction of return that ignores the price you pay has to be wrong. Furthermore, if the confused belief that earnings growth = stock return is responsible for all or some of investors' current exuberance over stocks, then this misconception may be responsible for the low level of expected stock returns going forward.

Let us use an example of a specific fast growing stock. I will pick on Cisco, probably the poster child for high tech blue chips, and try to determine what an investor today in Cisco can realistically (or even optimistically) expect as a long-term return.[22] As of June 2000, Cisco is trading at about a 140 P/E vs. 1-year trailing earnings. Wall Street analysts are currently forecasting (using the median IBES estimate) Cisco to grow earnings-per-share (EPS) at 30% per year for the next five years. The following graph is the actual annualized 3-year compound EPS growth Cisco has achieved over the last ten years or so (a 3-year period is used to smooth short-term fluctuations).[23]

Cisco is obviously a phenomenal company considering the spectacular growth rates it has generated for so long, and people who have faith in Cisco the company (not Cisco the price) are perhaps correct. However, there is a clear trend down in their EPS growth, and all economic intuition says this should occur. As a company gets bigger, and competitors gear up (both happening here in spades) it is natural for any company to slow its growth. Now, Wall Street analysts are not known for their restraint, and considering the trend above, their forecast of 30% growth going forward for another five years may be optimistic. But, for now, let us assume it is a good forecast[24]. Of course, that assumption does not mean you will earn a 30% return on your investment in Cisco. To estimate long-term expected return

more assumptions are needed. I will assume that Cisco continues to outgrow the market for another 5 years after its first 5 years of 30% EPS growth, as its growth linearly declines to normal market growth rates thereafter. I assume normal growth is 6% nominal EPS growth or about 3% real growth assuming inflation stays at 3%. In other words, growth in years 6–10 is linearly declining from 30% to normal growth of 6%, and it is steady at 6% from year's 11 onward. To be clear, under these assumptions nominal growth each year is as follows (scenario (1) in the following table):

	Cisco's Growth over Different Years (%)																		
Scenario	Year 1–5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Year 20	IRR	20-Year Annualized Compound Growth (%)	Multiple of GDP in 20 Years
(1)	30	26	22	18	14	10	6	6	6	6	6	6	6	6	6	6	7.5	14.5	4.6
(2)	54	46	38	30	22	14	6	6	6	6	6	6	6	6	6	6	10.0	22.4	17.3
(3)	30	29	27	26	24	23	21	20	18	17	15	14	12	11	9	7	9.0	20.8	13.3

Combined, these assumptions mean Cisco has above normal earnings growth for the next ten years. Over the full next 20 years, these assumptions lead to compound annual growth of 14.5% for Cisco, and real total EPS growth that is roughly 5x that of real GDP (assuming real GDP grows at a compound rate of 3% per year). Even for a great company, it is optimistic to assume such powerful long-term growth given the sea of competition. I also assume that when Cisco's growth slows they start paying dividends (or buying back shares) in line with the historic behavior of firms with comparable growth rates (and eventually settling to a 50% payout ratio, which is about the average for the S&P 500 from 1950–1999).[25] With these assumptions, and using today's price, I discount back the cash flow to investors and find a current internal rate of return (IRR) on an investment in Cisco of 7.5%.[26] If all my assumptions pan out then this is the return a long-term investor should expect on their Cisco investment buying in at today's price.

Of course, while still assuming my optimistic earnings growth assumptions come true, the next question is whether a 7.5% long-term return is enough. The answer seems to be a strong no. First, for perspective, note that this IRR is not that far above today's cash and government bond yields, and probably a similar expected real return (with a lot more risk) than inflation-protected government bonds. Now, most investors would probably expect and demand a higher return on Cisco than on the broad stock market as Cisco is more volatile. Say investors demand a 3% risk premium for the market over cash (itself far lower than history, but a bit higher than the Fama-French current estimate) and 1/3 above that for Cisco. So investors should require about 4% above cash to own Cisco, or a long-term return of about 10% today. If the above assumptions are right, Cisco's IRR has to go up 2.5%. Unfortunately, if the price moved down to get us there today, Cisco would have to immediately fall about 72%.

Now, getting more aggressively optimistic on my earnings growth assumptions will make the situation better, but the assumptions have to be truly heroic to get to a 10% IRR at today's price. For instance, in scenario (2) I assume Cisco grows EPS at 54% for the next five years and then slows to normal market growth linearly by year eleven. Now, the IRR does get to 10%. However, instead of growing 5x faster than real GDP for 20 years, Cisco's real EPS now grows about 17x more than real GDP grows for 20 years.[27] Of course, besides

leading to 20-year growth of mythological proportions, 54% per year is also well above their growth of the last 3 years, is way above the usually optimistic Wall Street estimates, and must be sustained for five full years going forward starting from today's huge base. Instead of assuming larger then 30% EPS growth over the next five years, I can get more optimistic by assuming that Cisco's above normal growth lasts longer. For instance, in scenario (3), instead of linearly declining to normal market EPS growth by year 11, instead I assume this decline occurs more slowly lasting until year 20. In this case, Cisco's IRR goes to 9.0% (still below the required amount), and I am now effectively assuming that Cisco grows real earnings for 20 years at a compound rate 13x real GDP growth.[28]

In fact, my entire analysis may be *way* too kind to Cisco. Ignoring for a moment that 10% is probably an unrealistically high estimate, what if you actually told their investors that they should expect a 10% long-term return? If they believed you, I think many (perhaps most) would bolt for the door as they expect, require, and in fact demand, the 30–100% annual returns they have been receiving. It is a real paradox that many Cisco investors would laugh at you if you told them they were only going to make 10% per year going forward, yet you need exceptionally optimistic assumptions just to get to a 10% long-term expected return. Eventually, something has to give, as "long and strong" gives way to "long and wrong" ("long and strong" is one of the deeper pieces of analysis you will often find on Internet chat rooms devoted to growth stocks).

Bottom line, while rational people can disagree, I think the case against Cisco as a long-term investment is reasonably strong given today's prices. However, it is not nearly as strong as the case against the entire NASDAQ 100. The entire NASDAQ 100 looks very much like a slightly less extreme version of Cisco, and while this analysis can certainly be wrong for one company (though unlikely, Cisco could surprise us with sustained 54% growth or even more), it gets much less plausible to assume this type of long-term growth for an entire index of 100 large companies. Can it happen? Of course, anything can. Perhaps the CEOs of these 100 companies are all children of Lake Woebegone? But, we must ask whether it is rational for an entire market to be priced with this as the base case.

Let us contrast the analysis of Cisco, with analysis of an "old economy" stock like the Ford Motor Company (Ford is far from the only example, and like Cisco, is only meant as an example). Ford is assumed by Wall Street analysts to have 5-year earnings growth going forward of only 8.2%. Pretty anemic huh. However, they are also selling for a P/E vs. trailing earnings of about 8 (i.e., you pay 8x last year's earnings for Ford vs. 140x last year's earnings for Cisco), and they have a current dividend yield of 4.3% (vs. a zero yield on Cisco). I make the same assumptions for Ford as for Cisco (i.e., they match Wall Street's growth expectations for 5 years, and then slow from 8.2% to 6% growth over years 6–10, ultimately then growing at 6% in perpetuity along with the overall economy). This is scenario (1) in the following table:

		Ford's Growth over Different Years (%)																	
Scenario	**Year 1–5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**Year 20**	**IRR**	**20-Year Annualized Compound Growth (%)**	**Multiple of GDP in 20 Years**
(1)	8	8	7	7	7	6	6	6	6	6	6	6	6	6	6	6	12.7	6.8	1.1
(2)	6	6	5	5	4	4	3	3	3	3	3	3	3	3	3	3	10.5	4.1	0.7

Instead of the 7.4% IRR I estimated for Cisco, Ford's IRR is 12.7% in the base case scenario. Remember above when I tried to find how optimistic I had to get about Cisco's earnings to get to a 10% IRR, well for Ford the question is how pessimistic I need to be to get to 10%. In scenario (2) I assume Wall Street is wrong, and instead of 8.2% Ford only grows earnings at a nominal 6% for the next 5 years, linearly slows down from 6% to 3% during years 6 to 10, and then grows at a nominal 3% from year 11 onward. These assumptions mean Ford is matching real GDP growth for 5 years, slowly declining for years 6–10 to zero real growth (or 3% nominal growth), and then staying at zero real growth *forever* in a presumably growing economy. Well, at these assumptions, far more pessimistic than Wall Street's assumptions, Ford's IRR is 10.5%. In other words, to get Cisco (or similarly the entire NASDAQ 100) up to an IRR of 10% I have to be far more *optimistic* than Wall Street, and far more optimistic than seems economically reasonable. To get Ford *down* to near an IRR of 10%, I have to be far more *pessimistic* than Wall Street. Note, I am not favoring old economy Ford over new economy Cisco based only on price while ignoring growth. I give Cisco tremendous credit for amazing growth going forward, and penalized Ford harshly for sluggish growth going forward. By doing so, I acknowledged that in a real sense Cisco is the "future" and Ford is the "past." This analysis favors Ford (by an obscene margin) over Cisco not because of a myopic focus on price, but because of simple recognition that price matters, and the right amount to pay for growth is not unbounded. In other words, at these prices, Cisco might be the "future" when it comes to earnings growth, but in all likelihood, Ford is the "future" if one cares about long-term stock returns. Again, I only use Ford and Cisco as examples of a more general market phenomenon. Any one company or even industry segment can possibly grow more (or less for Ford) than necessary to justify today's prices, but it is much more difficult for broad indices to achieve this.

It is easy to imagine a relatively new investor getting caught up in the exuberance of buying these stocks thinking that buying companies with great current earnings growth automatically means making a lot of money. Frankly, to various degrees, investors are often explicitly or implicitly told this by Wall Street and the financial media. Criticizing this new investor would probably be too harsh. However, one can certainly criticize the legion of "licensed" strategists/analysts out there who simply ignore the math. I do not think any of the bullish strategists and analysts would remain publicly bullish on the NASDAQ 100 while simultaneously predicting substantially below 10% long-term returns. On the other hand, I do not know many who would be willing to publicly predict that the earnings of Cisco, or far more unlikely the entire NASDAQ 100, will grow at 17x real GDP over the next 20 years, or that real GDP will grow at enormous rates for an extended period of time.[29] However, with only some small wiggle room on assumptions, that is the mathematical choice. Yet, Cisco is on almost every "must own" recommendation list I see. Go figure.

Let me remind the reader again that all of the above analysis of Cisco assumed (at least) that Wall Street's exceptionally optimistic earnings growth forecasts for the next five years do come to pass (and I added to the optimism by assuming the tremendous growth only slows gradually after year five). Even with these assumptions I still derive unacceptably low long-term expected returns. In other words, high earnings growth does not necessarily equal high long-term investment returns, and depending on the price paid, can certainly accompany low or negative returns. Let us now be more sober for a moment. What if things are not as optimistic as Wall Street forecasts? What if over the next 20 years Cisco, or the entire NASDAQ 100, posts growth rates below my very optimistic assumptions. In fact, history says this is the very likely outcome. Historically, studies have shown that the earnings of both fast and slow

growing companies on average "regress to the mean" quicker than what is priced into the market, and as I mentioned earlier, Wall Street's growth forecasts tend to be optimistic, and have little historical power for forecasting horizons past one or two years. Quoting Barton Biggs of Morgan Stanley in a recent missive (June 2000),

> The big-cap, sacred-cow tech stocks in the U.S. and Europe have been nicked but not ravaged, and no one wants to take the risk of being out of them. The multiples of EBITDA, earnings and sales are so elevated on these marvelous companies that they have to be discounting compound earnings growth at 20–30% a year for at least the next five years. This would be a feat never before accomplished by companies of this size. In fact, Bernstein's studies show that, based on the history of the last 40 years, there is only one chance in about seven that a "recognized" high-growth tech stock can sustain that exalted status for five years, and only one chance in 14 for 10. "Dwell on the past and you will lose an eye. Ignore the past and you will lose both of them."

Though I am highly sympathetic to this more realistic outlook, I will not replace my optimistic assumptions with more realistic assumptions, as the IRRs I find with the optimistic assumptions are depressing enough. Suffice it to say that if growth is less than my very optimistic forecasts, given today's prices for these companies, it will be very very bad. If the world turns out wonderfully and my optimistic forecasts are attained, it is merely very bad. Finally, if one is a super optimist, perhaps it is only bad.

I do not want to shout fire in a crowded NASDAQ market, but please make sure your smoke detector is working (and check where someone might be blowing this smoke).

"But I Am a Really Long-term Investor"

It is instructive to examine my analysis[30] of Cisco further. I actually projected Cisco's growth, and valued that growth over an infinite horizon. Under my assumptions, by buying at today's price, you get an unacceptably low return over this period (I think we can all agree infinity is long-term). Often it is said that companies will "grow into their valuations" over the long-term. This is misguided. There is no concept of growing into it. Alternatively, an investor might say, "sure I see your valuation argument, but I'm young and have a long time horizon, so naturally I want growth stocks." Well, unless they are Ponce De Leon, their time horizon is less than infinity. If the price is way too high relative to the long-term prospects (even if those prospects are great) then a long-term horizon does not save you. In a very real sense, at today's prices, the world has it backward thinking that stocks are only safe if held for the long-term. If my analysis is correct, then a short-term investor may still do well (or poorly) as nobody knows what will happen in the short-term. However, if my analysis is correct, a long-term investor is in big trouble with a high probability, because in the long-term, irrational valuation loses.

"The Long-term Will Be O.K. as We Have Entered an Era of Sustained Spectacular Earnings Growth"

Unless long-term, non-transitory earnings growth is much stronger than historical experience, investors are currently faced with a difficult choice. Either they must believe that going forward the expected return on the stock market is far lower than history because

market participants are generally content with this low return, or they must believe that we are going to have a significant drop in price (perhaps a quick crash, perhaps a protracted bear market) that will return prices to where expected returns are again attractive. Rather than face this uncomfortable choice there is another option, a loophole if you will. They can believe that long-term real earnings growth will be truly spectacular going forward, and thus stocks have an acceptable long-term expected return even at today's prices. In this section I examine how realistic this hope is for the S&P 500, for which we have 125+ years of data, and what it means for long-term returns. Although lack of data does not permit a similar study the situation appears even grimmer for the growth/tech sector.

The trailing 1-year P/E of the S&P 500 is now approximately 30, and the IBES median forecast of the next five years nominal earnings growth for the S&P 500 is about 17% per annum (capitalization weighting the individual median forecasts[31]). Let us assume this growth occurs. Next, as in my earlier analysis of Cisco, let us assume that over the next 5 years (i.e., years 6–10) earnings growth linearly slows to ultimately reach an assumed long-term rate of 6% per year in year 11 and beyond (or about 3% real at today's approximately 3% inflation rate). Scenario (1) in the following table sums up these growth assumptions:

	S&P 500's Growth over Different Years (%)																	20-Year Annualized Compound Growth (%)	Multiple of GDP in 20 Years
Scenario	Year 1–5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Year 20	IRR		
(1)	17	15	13	12	10	8	6	6	6	6	6	6	6	6	6	6	9.1	10.0	2.1
(2)	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	6.3	4.5	0.7
(3)	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7.7	6.0	1.0

These optimistic assumptions[32] imply about a 10% nominal or 6.8% real compound per annum growth in S&P 500 earnings over the next 20 years. In turn, if this occurs, my model would estimate an IRR for the S&P 500 of just about 9%, or if inflation stays around 3%, a long-term real return of just about 6%. While a bit less than long-term historical experience, and perhaps less than investors today really expect and demand, this would still be

a healthy long-term return. The next task is to see how reasonable this earnings forecast seems versus history. First, let us look at the history of real earnings growth for the S&P 500 using data from 1871–2000. The above figure plots the rolling prior 20-year compound per annum real earnings growth of the S&P 500 (the thick gray horizontal line represents 6.8% 20-year real earnings growth)[33].

The average, maximum, and minimum, compound per annum growth rates for real S&P 500 earnings over the 1891–2000 and 1946–2000 periods are as follows:

Statistics for 20-year Growth of Real S&P 500 Earnings (%)	
1891–2000	
Compound Real Earnings Growth	1.4[34]
Maximum Real 20-year Earnings Growth	7.5
Minimum Real 20-year Earnings Growth	–6.6
1946–2000	
Compound Real Earnings Growth	1.5
Maximum Real 20-year Earnings Growth	6.8
Minimum Real 20-year Earnings Growth	–1.3

Going forward, the 20-year real compound growth rate implied by the prior assumptions was 6.8% per annum. This means that if these assumptions pan out, the next 20 years (2001–2020) will about match the very best 20-year growth rates ever achieved. Now, what are the chances this actually occurs? Well, let us look at some more data. The following figure shows each year's real earnings divided by the average real earnings over the last 20 years:

This figure can be interpreted as a measure of whether very recent earnings are strong or weak vs. the last 20 years (i.e., are we near a local high, low, or neither?). The average for this figure is 116% from 1891–2000, the maximum is 192% and the minimum is 24%

(from 1946 on these figures are 123%, 184%, and 72% respectively). While not at a maximum, we can see that today's figure of 149% is impressive by historical standards. In other words, current earnings are well above their 20-year average as we have been experiencing relatively good times.

While strong current earnings growth certainly has been a good thing, it might be the case that extremely strong growth over the next 20 years is more difficult when starting from a high base, and far easier starting from depressed times. I test this hypothesis. The next table repeats the earlier table, but also includes these same statistics looking only at 20-year periods that began with earnings above the trailing 20-year average by at least 116%, and then by at least 149%. In other words, columns 3 and 4 examine only 20-year periods that started with current earnings above 20-year trend by at least an average amount (116%) and a large amount equal to today's value (149%):

Statistics for 20-year Real S&P 500 Earnings Growth (%)

	All Periods	Starting at >116%	Starting at >149%
1891–2000			
Compound Real Earnings Growth	1.4	0.5	–0.6
Maximum Real 20-year Earnings Growth	7.5	4.0	2.7
Minimum Real 20-year Earnings Growth	–6.6	–6.6	–6.0
1946–2000			
Compound Real Earnings Growth	1.5	1.3	0.9
Maximum Real 20-year Earnings Growth	6.8	4.0	2.7
Minimum Real 20-year Earnings Growth	–1.3	–1.3	–1.3

The data shows that when earnings are starting from a very high base, it is much more difficult to achieve exceptional 20-year growth going forward.[35] Over the full period the maximum 20-year growth for real S&P 500 earnings was 7.5%, and 6.8% over the post-war period. Excluding the periods that started at below average earnings vs. 20-year trend (i.e., only including those 20-year periods starting with current earnings above 116% of 20-year average earnings), the average 20-year growth falls slightly, but the maximum achieved falls from 7.5% to 4.0%. This 4.0% figure is far below the 6.8% real growth on the S&P 500 we need going forward. In fact, going further and looking at column 4 on the far right, we see that when starting from as high a base as today (which occurs for about 20% of the 20-year periods historically) average real earnings growth is actually *negative* over the next 20 years (and only averages 0.9% post-war), and has never been greater than a maximum of 2.7%. Summarizing, achieving a 6.8% compound real growth rate going forward would match the best post-war 20-year period ever, and come very close to the best 20-year period in about 125 years. However, when starting from such a high base as today, the best 20-year real S&P 500 compound earnings growth rate for over 125 years has been only 2.7% per annum. All considered, when compared to history, compound real earnings growth over the next 20-years of 6.8% seems very unlikely. And remember, in the historically unlikely event it does occur, we only get to an IRR of just about 9% on the S&P 500.

It is worth dwelling on this a bit. Matching the best growth in history, a feat *never close* to attained when starting from good times, only gets you to mediocre long-term returns, almost assuredly below the inflated expectations of most investors today.

What happens if we bow to the evidence and relax these very optimistic growth assumptions? All of the analysis so far gives tremendous credence to analysts' 5-year forecasts of earnings growth, and then goes on to assume this above normal growth only slowly moves down to normal over years 6–10 (and even my assumption of what is normal is a high estimate vs. history). This is highly tenuous. For instance, Fama and French examine this issue in their work cited earlier and conclude that earnings and dividend growth are best approximated by a random walk, and thus the best guess of future growth in any year is simply long-term average growth. Similarly, Bogle (*Journal of Portfolio Management*, Summer 1995) advocates using the simple average earnings growth rate over the last thirty years to forecast the future. What happens if these authors are right? In other words, what happens if earnings growth going forward is not spectacular? Well, it is not pretty. Scenario (2) assumes a 1.5% percent compound real earnings growth rate (the historical average) for all future years and the IRR on the S&P 500 drops to 6.3%. Note, while it might seem pessimistic, real 20-year growth of 1.5% is above the post-war average growth rate of 0.9%, and way above the full-period average growth rate of –0.6%, when starting from very good times like today. The IRR in this scenario is below commercial paper, and below the real return available on inflation-protected government bonds. Going back to the more optimistic 3% growth rate ad infinitum in scenario (3), the IRR recovers to an anemic 7.7%. In other words, if Fama and French, and Bogle, and my historical analysis of earnings growth, have any validity, reasonably optimistic estimates of the current long-term expected return of the S&P 500 might fall between 6.3% and 7.7%. Compared to inflation-protected government bonds the risk-premium is negative or just a hair above zero. Note, I am still avoiding the more pessimistic, but historically reasonable, case that starting from such a high base as today the next 20 years will see below average earnings growth.

None of this is a proof that tremendous real earnings growth (6.8% per annum or higher) will not occur over the next 20 years. Furthermore, the evidence above has to be considered more anecdotal than statistical as we do not get to observe enough 20-year periods to make solid statistical assertions. One can certainly argue that despite the historical evidence, times are now so different that massive long-term earnings growth is possible. While this cannot be disproved, I would just mention two caveats. One, we, like everyone throughout history, have the hubris to think that the present is radically different than the past. The last 125 years have seen an incredibly impressive array of technological advances, and all that is contained in the data above. Two, it is often difficult to remember how much things can change in 20 years. Twenty years ago was 1980, high inflation and a deep recession made optimism a four letter word. Ten years ago we were just waking up to our false belief that Japan had found the answer and was going to rule the world (and own every Monet in existence). To think that because things feel (and are) very good now, we can forecast extreme unprecedented earnings growth for a full 20 years going forward, strikes me as very dicey. To make it the base case for pricing the entire stock market (and not even an exceptionally attractive base case) strikes me as very scary.

III. Creative Defenses for the Price of Stocks Today

"Fine, I See the Math, But I Am Bullish Anyway Because of ..."

What is the solution for the bullish analyst/strategist? They face a formidable dilemma. I first found that tremendously optimistic forecasts of long-term spectacular earnings growth are necessary to justify current stock prices (either that, or one must believe that investors are now perfectly happy with bond-like returns on their equity investments). I then examined this possibility, and found that this spectacular sustained long-term real growth is highly unlikely going forward. This work was done on the broad market. The growth/tech sector is currently priced to even lower long-term returns despite using today's super optimistic earnings forecasts (remember the Cisco analysis). For the growth/tech sector, Wall Street's gigantic forecasts must actually be substantially exceeded before long-term returns become remotely acceptable. What is a bull supposed to do in the face of such mathematical torture?

Well, ignore the math of course! Just go ahead and set a high "price target" for one of a host of reasons. It is the paradox of this bull market that there is incredible focus on the short-term, all under the banner of long-term investing. Be bullish because the economy is currently great, because the Internet is the future (along with the children), because this quarter's earnings are rising fast, because the Fed has to stop ruining the fun soon, because we are "oversold," because earnings season will soon be upon us (or behind us), because summer's here and the rallying is easy, because prices have recently dipped, because despite the jump in standard CPI the core rate is still benign (switch core and standard CPI as needed), because we are in a Goldilocks economy,[36] because someone somewhere just "reiterated" a strong buy, because all the bad news we ever hear is "company specific," because the earnings slowdown you heard about was only because parts and labor were not available not because sales are suffering,[37] because the market is ready to put this Microsoft break-up stuff behind it, because we have reached a technical bottom, because this is a presidential election year, or just because we are all having a good hair decade. But for God's sake stay away from the math.[38]

While it might seem more complicated than the above reasoning, the math boils down to evaluating stocks as you would the purchase of any private business. What is this business worth in terms of the cash flow it will pay me over a long-horizon going forward? How does the value of this future cash flow compare to the price I must pay today? How does this compare with other alternative assets in terms of both expected return and risk? Note, this approach is not about what I can sell it for tomorrow, or next month, or next year.

Obviously, I am setting up a straw man as not all analyst/strategists brazenly ignore the math. Frankly, if an analyst does the math, but genuinely forecasts absolutely huge earnings growth, like the NASDAQ 100 growing 17x faster than real GDP growth for 20 years (and hence, eventually, hopefully, having dramatic growth in cash flow to shareholders) then I strongly disagree with them, but must respect their integrity and methodology. On the other hand, a bullish analyst who tells a story, but refuses to do the math, well, that is another thing entirely. This type of analyst is just focusing on the short-term (like those stories listed above) in order to avoid the uncomfortable unpopular conclusion that comes from the math. Personally, I wish this type of analyst would just come out and

say, "I think earnings announcements and the Fed holding off are going to take this mania up a notch, so hop on board the momentum train," as it would be more honest. But, of course, they cannot do that (obviously this would pose some problems as they all claim to be long-term focused). Instead, they go on TV or write a report, extol how great things are and what wonderful times we live in, list many of the reasons above, and never, ever, ever, touch the math. Why, because the math says that although things are wonderful, they are not in the ballpark of wonderful enough, and buying into a mania is not a good long-term decision.

I will leave it to the reader to form their own opinion about whether most bullish analysts truly have done the math and are willing to make the long-term forecasts needed to justify today's prices, or are just spinning a short-term profitable yarn. As usual, I think my own opinion is obvious. However, please, when you watch a strategist talk bullish while presenting only short-term stories and vague references to how wonderful things are, and definitely not doing the math, at the very least be extremely cynical.

"We Recommend Stable Tech Stocks with Earnings, Not Speculative Internet Stuff'

Decrying the Internet bubble, while simultaneously advocating a flight to "safe" solid tech stocks with large and growing earnings is currently a very popular way to sound prudent and rational (i.e., avoid those stocks with no earnings) and aggressive (i.e., own lots of tech) at the same time. It is somewhat of the "in" opinion to have among strategists. Protect your reputation by being prudent, but do not prick the bubble your firm so depends upon. Neat trick. No need to give up the dream, just stick with the boys, and the groupthink, that got you there. Tech is the place to be, the driving force of the economy, you have to participate. Again, just do not do the math.

None of the analysis I have conducted in this article, or the stories I have been examining, have focused on Internet stocks. Often the Internet is mistaken for the entire stock market bubble (for those willing to call it a bubble). However, Internet stocks have little to do with my ugly prognostications, or alternatively the extremely unlikely assumptions I find necessary to forecast good times for stocks going forward. Internet stocks provide a convenient scapegoat behind which to hide for those who want to avoid the math but still seem to prudently worry about valuation. In other words, if you want to look attractive, stand next to the truly ghastly. Again, given different and more aggressive assumptions, someone can disagree with my assertions about valuation. A strategist who recommends established big cap tech while eschewing the Internet, because he did the math and truly believes in phenomenal long-term earnings going forward for these companies, must be respected if not agreed with. However, do not let anyone avoid the math with this Internet trick.

Finally, in fairness I must say that a bearish strategist is in a real bind. A few years is not that long for a mania to last, but it can be career limiting or ending for a bearish strategist even if they are ultimately right. Thus, the pressure to appear bullish, by hook or by crook, in a mania must be acute. Steering your clients away from the most speculative, most dangerous stocks (i.e., perhaps the Internet), and maintaining your job, might appear to be the best course among difficult alternatives.

"The Days of Outsized Returns Are over, Going Forward the Stock Market Will Return More Like the 10–12% It Has Throughout History"

Another version of this is simply saying that after years of outsized gains, the market is now "fairly valued," leaving the impression of historically reasonable expected returns going forward but judiciously avoiding an explicit forecast. These comments are close cousins in spirit to decrying the Internet bubble while recommending big cap tech stocks. They sound prudent and sober, but are obviously still very bullish. Well, not to be a broken record, but please do the math! For the S&P 500 I found about a 9% long-term IRR using very very optimistic assumptions. When compared to the last 125 years I find that these assumptions have to be deemed unrealistic. Using more reasonable assumptions I find IRRs in the 6.3% to 7.7% range (effectively a negative or near zero risk-premium). In either case, this is not 10–12%. Looking at Cisco, and also extrapolating to the similarly priced NASDAQ 100, I find IRRs in the mid 7% range, based again on Wall Street's very optimistic forecasts. Any less aggressively optimistic analysis would find expected returns on these growth stocks to be well below bonds and cash.

It sounds prudent to tell people to calm down, and not expect the heady gains of the past few years to continue (and in fact, this comment alone is prudent). Calling something "fairly valued" seems utterly reasonable (obviously those nuts screaming cheap or expensive are not reasonable). However, it is only the appearance of prudence and the absence of mathematics that explicitly or implicitly combines this sentiment with forecasts like a 10–12% return on the stock market going forward. Again, if a strategist wants to make this forecast, then please also include an *incredibly* optimistic forecast of sustained long-term earnings growth and an explanation of why it will occur. It is not unreasonable to be an optimist, even an extreme one. However, it is unreasonable to skip the math in an effort to appear to be a prudent bull.

"We Have Just Lived through a Bear Market"

Wall Street and the financial media would have us believe that having lived through the first half of 2000 we are all scarred grizzled veterans of a bear market. Implicitly, this means that stocks are beaten down, there are bargains everywhere, and the bubble has popped. That is, it is safe to go back in the water.

First, the analysis in this article has been done largely on prices holding near the end of this period. No convenient choosing of say, March 27th, 2000 (the NASDAQ peak), was done here. If the very optimistic assumptions I have made earlier in the article are right, stocks are still priced way too high after the "bear market."

Second, speaking of conveniently choosing March 27th, 2000, it is the Knights of Bull who would have us focus on the drop from the peak. Focusing on peak to trough drops is certainly the most dramatic way to examine the market, and conveys the strongest impression of low prices and hence bargains. However it is not the most useful way to examine the market. Suffice it to say that stocks are approximately unchanged in price for the first six months of 2000, and way up in price over any reasonably longer period. This does not constitute a bear market, by any acceptable definition, no matter what even sillier heights they might have hit in-between.

With the exception of some pure Internet stocks, there has so far been no correction in broad market stock prices, rather there has been volatility, and a pause in the market's advance.[39] The fact that so many are so easily convinced that they have lived through major trauma is perhaps one of the scarier indicators of how spoiled and unrealistic investors have become. But, like any feel good Freudian, Wall Street and the financial media are more than willing to put investors on the couch, tell them it is not their fault, and counsel them on dealing with their "pain." Well, I certainly may be wrong about the valuation of the market today, but I am not wrong that the broad indices are about where they were six months ago[40], and much higher than any time further back.[41] No pain, just a lot of volatility, whining (admittedly, half of it from me), and noise, ultimately, signifying nothing.

"Tech Stocks Are Immune to Higher Interest Rates"

One of the main opponents of this expensive stock market has been rising interest rates. In defending the prices of technology stocks, many argue that interest rates really do not affect these stocks. This argument says that interest rates really only effect firms who finance themselves with bonds, or who's business depends on the price of money (like a bank). Well, this is not true.

There are two main reasons why interest rates can affect stock prices. One reason is because a rise in real rates makes future cash flows worth less. A second reason is because a rise in real rates presumably can slow the economy, and firms' earnings with it. Technology stocks should be immune to neither of these effects. Furthermore, they might in fact be very sensitive to both. One, because much of their growth is in the future, a rise in real discount rates should acutely affect their present value. Two, given how reliant their current pricing is on mammoth future earnings growth, any potential slowing of that growth should impact today's stock prices, probably seriously. With valuations so excessive these stocks do not have much room for even a small slowdown. The term often used, but perhaps too mild, is "priced for perfection." Simply put, tech stocks should be far from immune to the negative effects of higher interest rates.

A more difficult question is whether technology stocks are more or less sensitive to interest rates than so-called "old economy" stocks. I have no definite answers here. Technology stocks are clearly more sensitive to the discount rate effect (longer dated cash flows), but perhaps (as many argue) less prone to an earnings slowdown as demand for their products is so strong. But again, given their valuations, any slowdown no matter how small might severely affect their stock prices. Also, one could certainly see much greater price pressure on tech companies in a slower economy (i.e., customers shopping around more). Frankly, I do not know how these opposing forces shake out.

One point I will address is the often heard idea that old economy stocks are more sensitive to interest rates since they finance themselves with bonds as opposed to the more equity biased technology stocks. The reasoning is that if rates rise, these firms will have to pay more in interest charges, and thus earnings will go down (note, as a side issue, this ignores that the firms can alter their capital structure – as one can presume tech companies have altered theirs towards equity as their cost of equity capital is so low). However, another effect is ignored. If you own a bond, and interest rates rise, you lose money, right? Well, if a firm has financed themselves with bonds, when interest rates rise, they should make money, right? The answer is yes, as they have shorted a bond. Think of a firm as a

valuable set of assets (tangible and intangible). These assets are owned, literally divided up, by the stock and bond owners. Say interest rates rise, but for now, assume the assets maintain their value. Well, if the bonds go down in value, the equity has to go up! Simply put, they are short bonds, a declining asset.

On the other hand, there is a reason that old economy stocks may be more sensitive to interest rate hikes than technology ("new economy") stocks. By issuing more bonds, the equity of old economy stocks becomes a more levered claim on their assets. Thus, if the assets fall the same amount for technology stocks and old economy stocks, while ameliorated by the value of their bonds going down (as discussed above), we expect the value of the old economy stocks to fall more as they are a more levered claim on the falling assets. Of course, we would expect this effect for any change in asset value, and the fact that old economy stocks are generally less volatile than new economy stocks might make us doubt this theory.

Bottom line, I do not know which type of stock is more interest rate sensitive. However, I do believe that all theory, and empirical fact, says that both old and new economy stocks should generally (all-else-equal) fall as real rates rise. Sometimes this happens, but recently we have often seen a "rotation" from old to new economy as rates rise that actually causes the new economy stocks to go up. The story is often that investors are "fleeing to the less interest rate sensitive technology stocks." Amazingly, they rise in price as their fundamental value is almost undoubtedly falling. This is almost certainly the result of a host of forces that favor a "rotation" from one stock into another, rather than the selling of stocks for bonds and cash. For instance, equity mutual funds are loathe to own cash or bonds (ask Jeff Vinik) as they focus on "benchmark return/risk" not actual return/risk, individual investors do not want to give up the stock dream for boring cash or bonds, and Wall Street would much rather have you "rotate" than leave. To summarize, new or old economy stocks should (on average, and barring other news) go down as real interest rates rise. Which type should go down more, I do not know.[42]

"The Market, and Tech in Particular, Will Rally as Soon as the Fed Stops Hiking"

I included this in my earlier list of short-term reasons to be bullish despite the math, but it is really worthy of its own section.

It is somewhat ironic that as I write this (June 2000) the stock market, and the growth/technology sector in particular, is rallying sharply on the prospects of slower economic growth, and thus a greater chance the Fed will stay their hand. Many strategists and forecasters are jumping on the bandwagon, predicting better times ahead as soon as the Fed stops. The stock market rally of 1995–1999 is often pointed to as example, as it commenced once the Fed stopped its rate hikes begun in February of 1994. Of course, this argument follows on the heels of many bulls telling us that tech stocks are not affected by interest rates (discussed above). Apparently, having failed at this immunity argument, we are now to believe that the natural state of the market is continuous massive rally, only briefly interrupted by a nasty man named Alan. I see several problems with assuming that the rally recommences as soon as the Fed stops.

First, the reason to fear Fed rate hikes is largely that they may succeed in slowing the economy and raise discount rates on future cash flows. If the Fed stops raising rates it is

because they believe they have slowed the economy enough. We should not fear Fed hikes as a cartoon monster or superstition, but for the real effects these hikes can cause. Thus, the current cheer resembles the old joke, "the operation was a success, but the patient died." It is only long-term spectacular real earnings growth that can come close to justifying current stock prices, and that is not going to occur without very strong economic growth. Perhaps the hope is that the slowdown is short lived enough to stay the Fed's hand, but not so much to really slow earnings, but then earnings will not pick up enough to stir up the Fed again, but somehow long-term earnings will still be huge enough to justify today's prices. Then the porridge will be just right. Watching the bullish cheer an economic slow down at these price levels does feel like we have entered a different dimension of sight, sound, and mind.

Second, the comparison to 1994 is exceptionally dicey (as of course is any reasoning by singular historical example, including my earlier analogy to 1929). At the start of 1995 the P/E (using the Shiller data which I extended for a few months as his data ends in early 2000) of the S&P 500 was about 20, and now at the end of June of 2000 it is about 44. I do not know what else to say that could make it clearer that the analogy to 1994–95 is faulty beyond this approximately 120% increase in valuation. Again, there is nothing wrong with asserting that the market today is fairly valued or attractive. However, at the risk of boring even myself with repetition, it must be based on the math, long-term forecasts of earnings, dividends, and discount rates. To come to a bullish conclusion an analyst must be wildly optimistic in these forecasts. It seems clear that rising real rates and a slowing economy would reduce not enhance this optimism.

Wall Street, the financial media, etc., have all obscured the issue of whether the market, and growth/tech stocks in particular, are wildly overvalued, with the short-term issue (circa June 2000) of when or how much more the Fed will raise rates, how much the market will rally when they stop, and which stocks it will affect the most. I guess these stories are more fun to discuss than the math, both because they are simpler, and because they can lead to a bullish conclusion. Furthermore, these stories might or might not have short-term validity (the stories might be self-fulfilling or self-defeating in the short-term, who knows?). However, they are close to irrelevant to the long-term investor.

"Surely You'd Admit That All the Mergers & Acquisition Activity Is Bullish?"

First, I admit nothing. However, the logic here is almost compelling. The idea is that merger mania might be a sign that companies find it more attractive to buy than build, meaning they think other companies are undervalued. Perhaps managers, closer to the ground, betting their own companies' money, have a better perspective than any of us, and perhaps this all adds up to a very bullish verdict. As an example, many think the M&A boom in the 1980s was a response to equity values that by the early 1980s had gotten quite attractive vs. alternatives (like building something yourself). They believe the subsequent M&A boom was tied in with the 1980s bull market, and they may very well be right.

However, it seems that most of the mergers we hear about these days are stock for stock swaps. Now, I have not done a statistical study if this is true, and I do not know for certain if today differs markedly from the past. However, anecdotally, when you see a big

merger, particularly in tech/telecom/Internet, it always seems to be one company swapping their stock for another's.

There is a time-honored idea in economics expressed as "Gresham's Law." The short way to describe it is "bad money drives out good." A classic example is a government on a precious metal standard, who then issues some coins with a lower mix of the precious metal, but still the same face value (i.e., a debased currency). While the government naively thinks consumers and companies will use the old and new coins interchangeably, what happens is that the old money (the good stuff) goes into mattresses, and the new money (the bad stuff) is the only thing used in circulation. So, the bad money drives out the good.

Now, imagine a company who knows their shares are tremendously overvalued, but so are their targets. What would an acquirer use for a merger? Well they would not use cash (good money) to buy something overvalued, rather they would use their stock (bad money). Just like consumers owning good money (the company's cash) and debased currency (the company's stock) the companies hoard the good money and spend the bad. Gresham's Law in action. In fact, for some companies, their very high assumed growth is actually intimately related to their assumed ability to grow by merger, buying up super expensive small competitors with their own super expensive stock. Neat idea. However, while I cannot prove it will fail, historically this type of strategy has been a bust since the firm of Ponzi merged with Scheme.

Why then are they doing these mergers if the assets they are buying are not undervalued? Well, many possible reasons. First, perhaps they really believe in the synergies the merger will bring. That is, the two companies are such a good fit that the combination will be worth more than the sum of the parts. Historically, it strikes me as the biggest real synergy in most mergers is laying off people and combining overlapping operations, and that does not seem to be the main driving force today. Rather, the synergies today are supposed to be about things like "networking effects" where hitting critical mass, bringing together diverse media, etc., all add in a non-linear way. Of course, the networking effects today may be real, and perhaps synergies are more important now than ever. On the other hand, historically the search for non-downsizing synergies through mergers has been only a bit more successful than the 1000+ year search for a way to turn lead into gold. However, rather than actually turning lead into gold, historically a far more successful endeavor has been convincing the market that years down the road you will do it. Finally, for the cynics out there (and in here), the funny-money merger boom might be at least partially driven by ego, empire building, the fact that mergers allow all kinds of creative accounting (how many firms now report superb earnings *before* merger costs?), and last but not least, the fat fees Wall Street makes on this activity.

While not directly on topic, I want to mention a few related corporate finance instances of bubble logic. How many firms have we seen planning to do an IPO, only to pull it when the market goes down? These firms never say, "we are waiting to sell until we get massively overvalued again," but rather the answer is always that they are pulling the IPO because of "too much volatility," as if they would have pulled it if the volatility were upward. Also, how many spin-offs do we see of high tech areas (e.g., wireless, Internet, media, etc.). The reasons given range from making it easier for Wall Street to value the company's separate parts compared to the difficulty of valuing the whole (while they are at it, perhaps they could use smaller words and bigger type in their annual reports), and being able to compensate the people in these divisions better. My favorite explanation was an executive

(unnamed) who said in a statement that his company is exploring a tracking stock or spin-off of portions of its phone operations to "allow more efficient management and focus on business customers." Why on earth does this not apply to their divisions that are not so richly priced? I guess for these divisions inefficient management and a lack of focus on customers is fine. Of course, if the areas being spun-off are undervalued vs. the enormous growth opportunities (as the companies would undoubtedly tell you), the firms could just retain the divisions and implicitly reap the benefits of this undervaluation, as opposed to giving it away to investors. Again, a cynic might say that these firms know the stuff they are spinning off is massively overvalued, and are just getting out of Dodge (at least for the portion they sell).

Obviously, I have a fairly cynical view, but rather than bullish, much of the recent corporate finance activity strikes me as yet another indication we are in a mania.

"Valuation Schmaluation, If You Had Worried about Microsoft's Valuation 15 Years Ago, You Would Have Missed Making a Fortune"

The idea here is simple. If you care at all about valuation (I am not talking full fledged distressed value investing, but just caring about what you are paying for a stock) then you would have missed investing in Microsoft (and some others). Thus, it is silly to worry too much about valuation! Now, this logic is precisely the same as pointing to the winner of a lottery and declaring that lotteries are a good investment. Everyone remembers the lottery player who won, and the super expensive growth stock that was really worth the price (or much more). On the other hand, a host of losers (lotteries and expensive stocks) fade from memory. A large litany of academic work shows that systematic high P/E investing is not a great long-term idea. Common sense knows that lotteries are usually not positive expected return investments. You can invest in them either for entertainment, or desperation, but be prepared to lose. Arguing by singular example (or a small handful of them) based on what has happened ex post, is very dangerous. The success of Microsoft no more proves that valuation does not matter, than it proves dropping out of Harvard is always a great idea.[43]

"Do Not Try to Time the Market"

Well, one way to get people to ignore the current high price of the market is to explicitly tell them it is a sin to actively change their exposure to stocks. In fact, this is probably the most commonly heard piece of conventional wisdom on Wall Street. Avoid market timing! First, let me say that 99/100 times this is actually a very good piece of advice, and Wall Street probably does the investing community a service by popularizing it. The transactions costs, tax effects, and general unpredictability of the market, all make timing a dicey proposition for the individual and the professional. All else equal if I had to chose between giving a friend the above advice, or the opposite advice of "actively and often try to time the market" it is a no brainer, keep your hands off the portfolio.

However, let us put our cynical hats on for a moment. Wall Street (buy and sell side) is in the business of selling you stocks, and they do not want you leaving the market. Let us

rephrase the advice "Do Not Try to Time the Market" another way. How about, "Ignore the Price of What I Am Selling You and Buy No Matter What." If you think about it, it is the same advice. If your salesman told you to ignore the price of any other purchase than common stocks because "it will all work out over the very long-run," you would run clutching your wallet. While perhaps usually good advice, "do not try to time the market" cannot mean ignore price entirely, as in the extreme this is obviously silly. However, making great long-term returns without any imposition of effort or vigilance (i.e., having to watch prices for opportunities or bubbles) is obviously a seductive siren's call. If being price sensitive means timing the market, and timing the market is a cardinal sin,[44] then prices have no anchor to reality. If one is looking for possible causes of a financial bubble, then the "ban" on market timing must be a prime candidate.

In fact, the most common reasons Wall Street gives us for avoiding market timing are quite silly (the good reasons are listed above). Even if they lead to good outcomes, silly reasons should not be tolerated (we should not have to fool ourselves into doing what is right). Let us talk about two of the more common anti-market timing rationales.

Reason #1: If you timed the market and managed to only miss the few best days for the market, you give back all the positive returns of stocks while retaining most of the risk. This argument is found everywhere (mutual fund ads, stories in the media, advice from financial planners, etc.). It seems every firm has their own version of this parable.[45] The numbers are supposed to shock you, and on the face of it they do. You do not have to miss many of the best days to lose a lot of the return from being invested in stocks. However, this is really a very silly argument. First, it postulates a very wacky, extreme market timing strategy. Even those people who do try and time the market probably do not do it by going to all cash from 100% equities for just a few select single days. After postulating this ridiculous strategy, those who advance this line of reasoning against market timing then make the minor assumption that one then gets this timing pathologically wrong by choosing amongst the thousands of possible days, the literally worst possible ones to be out of the market. The analyst doing this exercise is then shocked, shocked to discover that pursuing this crazy extreme market timing strategy, and getting it astoundingly wrong, appears to seriously hurt returns (and is also shocked, shocked to discover gambling is going on in the market). Interestingly if one carries out these calculations assuming one only avoids the worst days for the market (the opposite of the normal calculation), long-term returns are increased by a similarly dramatic amount. This is not shocking, and fair. If one pursues an extreme strategy the consequences are high but relatively symmetric. Essentially, this argument against timing constitutes some very silly and intentionally selective mathematics, paradoxically supporting some good advice.

Reason #2: Even if you have perfectly terrible market timing, that is, investing your savings for the year on the worst possible day each year (i.e., the high day for the market that year), if you keep investing over a long horizon you still do much better than someone who might have much better market timing (i.e., investing on the low day for the year) but was in the market for a shorter time than you. Thus, the refrain is, "it is not the market timing, but time in the market that counts." This argument is more mathematical trickery. Where reason #1 used a market timing strategy that is way too extreme, reason #2 uses one that is way too tame. Assuming terrible market timing (investing on the worst day of the year) sounds pretty bad, but the only market timing going on here is on the new investment. The main portfolio (i.e., the compounded value of all old investments) is still invested in

the market for the long-term, and very quickly the returns on this main portfolio come to dominate the timing done on the relatively small annual investments. I do not think that stocks must win over any long-term, but they certainly have done great over the period these tests are run over. Thus even with terrible market timing (on the relatively tiny additional investment each year) the person in the market longer generally won. Essentially, we are comparing someone in the market for a longer period but doing a tiny amount of terrible market timing, to someone in the market shorter doing a tiny amount of great market timing. Because the market itself was so strong over this period, and because the amount of true market timing was so tiny, the effect of being in the market for longer dominates. This is not really a test of market timing at all, but a restatement of how wonderful it has been to be invested in this bull market. Earlier, I talked about the long-run argument for stocks, and this is a subject for legitimate debate. However, adding literally a smidgen of market timing to this long-run argument, and discovering it does not matter much, is not exciting news, and is again misleading mathematics.

What is really amusing is when you realize that reason #1 and reason #2 are the exact opposite in spirit and are often mentioned together or at least by the same firm! #1 says do not time the market because the consequences to a misstep are so severe, and #2 says do not bother because the consequences are so miniscule.

What do I think about market timing? Well first, it is generally a bad idea because it is very hard to forecast short-term market movements, and transactions costs and taxes (for taxable investors) will kill you. Second, I think it is very important to distinguish the short-term from the long-term. Perhaps some have effective systems for short-term timing, perhaps not. I am reasonably cynical about the prospects, and thus, without substantial evidence to the contrary, would generally avoid short-term timing. However, over longer-term horizons, I do think making conscious portfolio shifts based on the relative attractiveness of different asset classes can make sense, and especially so when extremes are reached. Note, Wall Street, home of "Do Not Try to Time the Market," implicitly agrees as their strategists are all running around with their changing recommendations for how much stocks, bonds, and cash to own. Looking ahead now, I have no idea what will happen over the short-term. However, over the longer term, it seems pretty clear that either (a) the risk-premium on stocks has permanently come way down, or (b) people are in for a very rude awakening when they realize they do not like holding stocks with very low expected returns, and prices then will sharply fall. Furthermore, I think this analysis applies particularly harshly to the growth/tech sector of the market. I do not know when it will happen, and nobody should ever be certain they are right over any horizon, but either way it looks like a pretty good bet to lighten up on equities now (this does not mean sell them all or go short). That is long-term market timing, and I think done occasionally in moderation it can make sense.[46]

"Dips Are Not to Be Feared, But Are a Buying Opportunity"

Ah, to buy on the dips, one of the most hallowed activities of the last few years. First, I had to put this one after "Do Not Try to Time the Market" as often the same firm, and sometimes the same person at that firm, will give you both pieces of advice. That seems pretty contradictory to me, as buying on the dips is pure short-term market timing at its finest.

Some will argue that dip buying is not short-term market timing as they are simply looking when to enter a long-term investment. It is a common refrain to hear even bulls say they are rooting for the market to dip, "as they have cash to put to work." However, one must ask them what they are doing with the cash until they get their dip to buy? Clearly they believe that the money should be in stocks long-term, and they always have the option of investing now. If they think it is a good idea to hold some money back waiting for a dip, then they are forecasting the stock market's expected risk-adjusted return between now and the end of the dip to be below their alternative (presumably cash or bonds). It is 100% certain that there will be a dip eventually. But, in a long-term rising market (as every bull believes in) the dip might very well occur after the investor has suffered a large opportunity cost from sitting in cash waiting for it. In other words, in a long-term rising market, today's prices might never be seen again, even after some future dip. Thus, someone holding cash waiting for a dip is forecasting that this will not happen to them, and that some time in the near future prices will be lower than they are today. In other words, they are forecasting the *short-term* attractiveness of the market. I could not think of a better definition of market timing.

Now, if you just look at the success of buying dips in isolation, you will find it works as buying in general has worked for a long while. However, to evaluate if there is any content to the idea of buying dips, as opposed to just buying, you must show that after dips it is a better time than normal to buy (i.e., that it is an effective type of short-term market timing). You must show that what you gain or lose from being on the sidelines waiting for a dip is more than made up for by your extra returns after buying a dip. Otherwise, you are just arguing again to invest more in stocks (an interesting but separate argument). Although buying the dips is a simple short-term market timing strategy, I have not seen it analyzed as such (i.e., fairly judging whether can you do better buying the dips than an equivalently risky strategy) and I would be interested in seeing such a study (it might even work).

In my opinion, too many dip buyers think they are in fact value players, only their valuation model is not based on P/Es or IRRs. Their valuation model consists of "if a stock or market has dipped below its all-time high, it is cheap!!" This "buy the dip" mentality may indeed have contributed to the stability of equities we have seen for a few years as any dip is quickly erased. If permanent, it could mean equities really are less risky going forward (and also that their expected returns really are permanently lower). On the other hand, if only temporary, much like an infection not quite killed by an antibiotic, every dip that comes roaring back might make things far worse in the end.

Finally, if I hear one more person refer to buying a 200 P/E stock, that is up 200% in the last year, but is down 5% from its closing high two days ago, as "bargain hunting," I might have to start doing some hunting of my own.

"We Have Heard the Bears Wrongly Scream 'Over-valued' for So Long Now"

Finally, failing all else, we can forget arguing the merits and just laugh at the bearish for being wrong for so long. Well, as you might guess, there are several logical flaws with this plan (but, of course, it can still be pursued for pure entertainment purposes).

It has really not been that long. It was only in December of 1997 that P/Es (from my first figure) crossed their former high of September of 1929, and in fact only in January of 1995 when they crossed and remained above 20. While 3–5 years may seem like a very long time (it certainly does to me), it is not. In terms of markets, 3–5 years is a blip of time. One can certainly argue with the thesis that equities are overvalued, but a legitimate argument is not "you have been saying this for three years now." If one believes that we are in a bubble, then that argument simply uses the mania to justify itself. Argue the merits of the case going forward, but not the recent returns, as over short periods, returns are basically random. Furthermore, once you admit the possibility that the stock market can become overvalued, it becomes very difficult to discuss what limits there are on this overvaluation (again, short-term market timing is difficult). Neither the bulls nor the bears, if things should turn their way for only a short period, should point to recent returns as an indication that they are right regarding valuation.

"OK, So If You're So Smart, Why Doesn't This Bubble Pop?"

This is a darn good question. The prior section made it clear that it is incorrect to justify current stock prices simply by deriding the bears as being wrong so far. However, this should not let a bear off the hook either. Turning it around, while the bulls should do the math to justify their beliefs, a bear who believes the market is priced irrationally high should have a theory about why such irrationality can persist for so long. Barring such a theory, it is probably prudent to assume the market knows more than you do, and that stock prices today are rational. Of course, not surprisingly, I do have such a theory.

Actually, I borrowed the theory. An article called "The Limits of Arbitrage" by Shleifer and Vishny (*Journal of Financial Economics*, 1995) sheds some light on why irrationality can last. The article is complex, but essentially the authors postulate a true arbitrage situation, one in which you are guaranteed to make a risk-free return. However, the authors also postulate some real world complications, namely mark-to-market and bankruptcy risk. Meaning, because of interim fluctuations, you will not necessarily be around to see your arbitrage through to a successful conclusion. The authors go on to make the point that in the real world, even a true arbitrage is not necessarily instantly eliminated because of these risks. Now, to apply this to our question, imagine for a moment that somehow you privately knew for certain that the real return of the stock market would be negative over the next 20 years. What would you do now?

Well, if you run a mutual fund you might very well stay fully invested. Imagine you sell stocks and raise a significant amount of cash and the bubble expands still further (obviously a very real possibility over short-horizons, even if you know the next 20 years will be poor for stocks). When this happens, the marketplace has been quick to punish the under-invested mutual fund manager. However, imagine you do not raise cash and the market declines sharply. Well, you are in the pack, and while your industry might suffer, you will not necessarily suffer relatively (and we all seem to care at least as much about relative as absolute suffering).

If you run an endowment or pension fund you also might not raise that much cash. Much like a mutual fund's shareholders, an endowment's board does not have a 20-year time horizon. You will be evaluated at a far shorter frequency. If you are right, and a crash

ensues, clearly you will be rewarded. But, if the mania rolls on, in a year or two you will find yourself out of a job, and branded a "maverick" (there is an actual term called maverick-risk). If you are right in the short-term, you get a pat on the back and maybe a little something extra in your envelope, but if you are wrong, back to the career drawing board. It is easy to see how even if you possessed a 20-year crystal ball, it might be difficult to act upon this knowledge (of course, you could always hope the crash does not happen for 19 years and 364 days, and then act swiftly and confidently with 24 hours to go).

If you are a sell-side strategist, or worse, a broker, then the incentives to stay bullish are probably more acute. Is there anyone who truly thinks that even the certain knowledge that stocks are priced to offer poor 20-year returns would turn the brokerage crowd bearish? If so, you have more faith than I. This is not necessarily a knock on their morals, but a knock on the incentives and asymmetric loss functions (a geek term for punishing them more for being bearish and wrong, than for being bullish and wrong) we as a group provide them.

Even if you are an individual investor, amazingly, you might still do little with your funds. Twenty years is a long time. The mania can certainly continue in the short- to medium-term, and perhaps you think you can time when to get out? Worse, if it does go on who wants to admit to their friends at cocktail parties that they are missing the party. While this sounds extreme, and probably is, by no means is it certain that individuals, even given this distant foreknowledge, would immediately shed their equities.

Now, let us make the example more realistic and acknowledge that nobody knows for certain what will happen over 20 years, just that the odds are now far worse for equities than normal. Also, add that there is a large cadre of the marketplace seduced by the many fallacies we describe who seem to continue buying no matter what happens. It is quite easy to see how the prospect of poor long-term equity returns could have little immediate effect. Given the strong belief that equities will underperform inflation for the long-term, the rational thing, and the only act of a prudent fiduciary, would be to at least lighten your equity exposure. However, if it were realistic to think that "career risk," "maverick risk," and "asymmetric loss functions" might stay your hand even in a certain world, they certainly can stop you cold in an uncertain one. Frankly, all considering, it is quite easy to see irrationality persisting (and on some days hard to see it ending). Unfortunately, it seems very possible that we are all just doing a "danse macabre," gentlemen and fiduciaries all, waiting for the disaster we know is coming to strike, so we can all go down together.

Of course, the only bright spot is that the inability of many to act on long-term knowledge, even if relatively certain, only increases the long-term benefit to those who can act.

IV. Growth vs. Value Investing

"You Will Be OK, Just Stick to Buying Great Companies"

This quote is related to the earnings growth = stock return fallacy I examined earlier, but because of its simplicity, is perhaps more widespread. A great company is clearly worth more than a crappy one (pardon the technical jargon). Recognizing a great company before the market does is clearly a way to get rich. However, buying a great company after the market knows about it is at best a wash if the market prices things rationally, and could be a negative if, as history (and academic study) seem to show, investors systematically overpay

for perceived greatness. Looking at the long-term data (the longest data I have seen is for 1927–1999), and even including the last few years, value investing, or buying out of favor companies selling for cheap prices, actually outperforms buying expensive companies perceived to be "great." For instance, looking at data from 1929 to 1997, Davis, Fama, and French (*Journal of Finance*, 2000) found that firms selling for higher prices (they defined higher prices in terms of the price-to-article ratio) on average underperformed those firms selling for lower prices. Presumably the higher priced firms were perceived to be "greater" than the cheap firms at the time. Similarly, Lakonishok, Shleifer, and Vishny (*Journal of Finance*, December 1994) found that firms priced expensively, and firms whose sales have been growing relatively quickly, on average make poorer investments than those firms not priced for greatness, or growing as fast. These examples are a drop in the bucket as there is a great body of literature on this topic, most of which finds that investing in cheap, slow growing "not great" firms has generally beaten investing in expensive fast growing "great" firms over the long haul.[47]

As always, any systematic strategy can have poor performance for periods of time (the last two years for instance has seen a great victory for investing in expensive firms). However, at the very least, the historical evidence is that you should not be able to beat the market simply by buying companies generally thought to be great. The simple lesson is that price matters! Please watch carefully as many investors and analysts love to discuss how "great" a company is, but again, do not want to do the math. This does not mean you should not pay up for growth or greatness, but it does mean you better be wary of the price you are paying and know that historically investors have probably paid up too much.

Finally, I have actually heard it said, by individuals and on occasion by recognized experts, that there are certain great companies you "have to own at any price." Well, hopefully only once in this article will I dip from sarcasm to outright rudeness, but that is a dangerously foolish statement. If you hear someone say it, run screaming. Better yet, sell them something.

"Value Stocks Are Dinosaurs, I Would Steer Clear of Them"

Alright, you got me. This quote is really the same as the prior one, just coming at it from the other side. However, it does give me a sneaky chance to examine this issue in more depth.

For several years now value stocks (cheap stocks perceived to have problems, poor growth opportunities, or exposure to some painful risk factor) have generally underperformed growth stocks (fast growing expensive stocks perceived to be "great" companies). Despite the long-term evidence that the opposite is true, many make the simple forecast that this trend will continue. Here I examine the medium-term prospects for value vs. growth using fundamentals, not simple extrapolation of what has been happening lately.[48]

In a recent study, Asness, Friedman, Krail, and Liew[49] (*Journal of Portfolio Management*, Spring 2000) try to build a model to forecast the medium-term returns of value vs. growth stocks. I will follow a similar methodology. First, I form stocks into five groups based on how expensive they are, taking care to make sure each industry is relatively evenly represented in each fifth (so I am not just examining tech vs. everything else, but cheap vs. expensive stocks within each industry). Examining the top one–fifth (the most expensive growth stocks) vs. the bottom one–fifth (the cheapest value stocks) I produce the following figure (ending 6/30/2000)[50]:

The dark line in the figure is the P/E of the median growth stock divided by the P/E of the median value stock. It is measured using the scale on the left, and obviously is always greater than 1.0 (i.e., by construction growth stocks are always priced more expensively than value stocks). However, there is great variance in this number, and in particular note that recently it is at historical highs. In other words, growth is currently priced more expensively vs. value (by a large margin) than at any time over the last 20 years.

Now, the light line in the figure is the difference in expected (IBES median) 5-year earnings growth of the most expensive stocks (growth stocks) vs. the cheapest stocks (value stocks). Using the scale on the right we see that this differential is always positive, ranging from about 6% to as much as 14%. In others words, growth stocks are always expected to outgrow value stocks. Like the ratio of the P/Es, this number also varies through time. Sometimes growth stocks are expected to outgrow value by a large margin, and sometimes the expected margin is smaller. Now, if the reader will only cover up the last two years (oh, if only I could really do that) they will see that while not perfect, these two measures (the ratio of P/Es, and the differential in expected growth) do seem to move together. This is entirely rational. When the market is expecting growth stocks to outgrow value by more than usual, they have historically been priced more expensively than usual. What about the last two years? Well, lately the relationship seems to have broken down. Growth stocks are far more expensive than ever (for the 20 or so years I examine), but are not expected by Wall Street to outgrow value by more than usual.

What is the bottom line? Asness et al. first find, like other researchers, that on average value (by their definition) has defeated growth over these 20 or so years. However, this victory was by no means uniform or consistent. They found the best time for value vs. growth was when growth was the most expensive (i.e., the dark line in the above figure was high) but not giving up too much expected earnings growth (i.e., the light line in the above figure was not high). Furthermore, they found that this model was surprisingly powerful for forecasting value vs. growth at horizons of one year. Of the two effects, the relative valuation levels (dark line) matters more, but both matter. As of now we have a situation where value is priced incredibly cheaply vs. growth by historical standards. However, the valuation differential is not being ameliorated by higher than normal growth expectations.[51]

Thus, the Asness et al. model forecasts that the expected return for value stocks is currently far greater than the expected return for growth stocks (and greater than any time over the last 20 years). While recent value performance has been the worst of times, this model is now forecasting the best of times for the relative returns of value stocks going forward. Like most reasonable models, this forecast is made with far more confidence for medium-term time horizons (say 1–3 years) than for the short-term (say less than 1 year).

There are caveats of course. It is possible that the Asness et al. model is missing something that would justify the current relative price of growth vs. value stocks. Perhaps Wall Street analysts' 5-year forecasts, for some reason, do not express the analysts' true optimism regarding the future earnings of growth vs. value companies. Or, perhaps the analysts do not get it, and the market is correctly pricing far larger and longer excess earnings for growth vs. value compared to what the analysts think will occur. Perhaps neither value investors, nor Wall Street analysts understand today's economy, but the little guy does.

Well, here I will apply a version of Occam's Razor, the idea that the simplest explanation is often correct. Value historically beats growth on average (for at least the 75 years or so we can examine it). Currently, value (as Asness et al. define it) is cheaper than it has ever been, but is giving up far less expected earnings growth than you would normally expect to see at these price differentials. While we can, and should, explore all kinds of stories that might justify this pricing, perhaps the simplest explanation is that history, and the above model, is correct. Perhaps investing in value stocks right now just looks a heck of lot better than investing in growth stocks. Recent popular discourse has made many of the investing fallacies I study more common than ever (e.g., earnings growth = stock return, great companies necessarily make great investments, price does not matter if you are long-term, etc.). All of these fallacies favor the broad market, and growth stocks in particular. While by no means a certainty, the prevalence of these fallacies makes it easier to believe that growth stocks are currently extremely overpriced vs. value stocks. In summary, which of the following is a simpler explanation? First, that value, which wins on average, looks much cheaper than ever today, and using Wall Street's own earnings forecasts is not worse vs. growth than normal, is being rationally shunned by investors who have their own, deeper and better forecasts of long-term earnings. Or second, that investors are caught up in a momentum driven mania, or afraid of being trampled by it, making value currently a historic opportunity for someone with even a medium-term outlook. As usual, our opinion (mine and Occam's) is obvious.

V. Miscellaneous Examples of Bubble Logic

"Technological Advances Make the Market Safer Today"

I want to switch gears now and talk a bit about the mechanical workings of the market itself. It is a common refrain that the individual in the market is made safer today because of the speed at which he/she can gather information and trade. I think this is mostly a myth. I think the myth comes from the fact that many investors believe that in the event of a crisis they can get out. I think many, implicitly or explicitly, know that they own some very overvalued securities, but they are comfortable because the immediacy of information gathering, and the ability to trade near instantaneously, makes many feel they can nimbly avoid the bad times. In other words, there is an illusion of control.

However, what is missed is that near everyone possesses this same technological "advantage," and that getting out of the market before a crisis is a zero sum game. For anyone who gets out, someone has to get in. Put more graphically, in the aggregate, when the $#@! hits the fan, nobody gets out alive. Hearing about bad news quickly on a financial news cable station, or being able to trade immediately in your Ameritrade account, might make you a winner if you are the first, but there is a loser on the other side. If the market is way overvalued in the aggregate the ability to get instantaneous information, and the ability to trade on it quickly, is irrelevant for everything but the distribution of who gets killed.

To be balanced, I have to say that one hypothesis I must give credence to is that technology has led to a more informed investing populace, and a more informed investing populace is a more educated investing populace is a more stable and less prone to panic investing populace. Now, I give credence to this because it sounds plausible and I cannot refute it. It is possible that people are just more patient investors now, partially because of technology, and that the market is safer for this reason (i.e., from a change in psychology and education levels brought about by technology, not the real safety effects of the technology itself). However, to play devil's advocate, there really is no evidence for this, and what anecdotal evidence there is probably comes more from investors embracing the long-run argument for equities and a dip buying philosophy, than any technological edge. To be extreme, I defy anyone to spend any time on an Internet bulletin board related to stocks and then believe these people are making the market more rational (as a bear, the nicest thing I have been called on these boards is moron).

Finally, let us talk specifically about the on-line trading of one's own account. I do not know if many of you readers have played video poker in Las Vegas (or anywhere). I have, and it is addicting. It is addicting despite the fact that you lose over any reasonable length period (i.e., sit more than an hour or two and 9/10 times you are walking away poorer). Now, imagine video poker where the odds were in your favor. That is, all the little bells and buttons and buzzers were still there providing the instant feedback and fun, but instead of losing you got richer. If Vegas was like this, you would have to pry people out of their seats with the jaws of life. People would bring bedpans so they did not have to give up their seats. This form of video poker would laugh at crack cocaine as the ultimate addiction. In my view, this is precisely what on-line trading has become over the last several years (with perhaps some lessons taught only very recently, and not necessarily learned). This is just my opinion, but I think it is very plausible that these "crackhead" traders might be an important part of a multi-year bidding frenzy taking stock prices well past the rational (and I will not even dwell on the paradoxical fact that this bull market, carried on the back of the long-term argument for equities, has spawned a subculture of high turnover day traders). In sum, it is highly arguable whether technological advances have made the market safer, and it may well be that the opposite has occurred.[52]

"Stock Splits Are a Great Buying Opportunity"

First, a joke. I am stealing this from somewhere (I do not remember where), but what is the best "pricing model" for an Internet stock? Well, at $50 it is cheap, at $100 it is fair value, but at $200 it is cheap again as it is about to split 4:1.

I am not going to insult my readers, and myself, by explaining for too long that a stock split is a paper transaction that means next to nothing. I am not going to dwell on the fact that when I overhear conversations about stocks, often times what is generating excitement is that a split might occur, or has just occurred.[53] I will mention that the companies splitting their shares know that for some reason people care a lot about this event, and the cynical among us might notice stock splits of late corresponding to times the company is just aching for some good news (and for some reason manipulating reported earnings is more difficult than usual).

Finally, let us not dwell on the fact that some investors have beepers to alert them when a stock is splitting, as such knowledge can only make us into misanthropes.

"You Can Believe Our Advertising"

Another interesting characteristic of the bubble is how many mutual fund ads we have to read lately advertising triple digit returns. It is interesting to examine what mutual funds get advertised, and contrast that with the reasonably well-established empirical fact that the average actively managed mutual fund underperforms the market.

I thought I might shed a little light on this subject by discussing my "formal" model for how a money manager decides which mutual funds to advertise (obviously, I am speaking in generalities as not all firms do this). Remember, mutual funds almost all have benchmarks (passively managed indices representing their style which the funds are supposed to beat) so companies have three options: (1) do not advertise the fund, (2) advertise the fund without mentioning the benchmark, or (3) advertise the fund and show performance vs. the benchmark. My model for how this decision is made is the following (with the time horizon being approximately the last year):

Step one is obvious, if the fund has done poorly do not advertise it. Step two, if the fund beat the benchmark and has done well in absolute terms, then definitely advertise it and show it vs. the benchmark. Step three, if the fund has failed vs. the benchmark, but is doing

very well in absolute terms, advertise it and leave out the benchmark! This happens all the time, and it is a good bet when you see a fund advertised with a good track record, but no benchmark present, the fund underperformed their benchmark. That is especially prevalent now with so many growth funds posting excellent results, but many not beating growth and NASDAQ/technology benchmarks. Not too many ads for value funds these days are there (even if they beat the benchmark)?

Why should you care about this selective advertising? Well, several reasons. One, there is only a tiny amount of evidence that winning mutual funds continue to win.[54] Thus, the proposition of just investing in the funds that have done well recently is very dicey. Two, by only examining the ads you get a very skewed view of how active management does vs. passive management, and how the broad market has done in general. A pretty big decision for an individual is whether they should go with index funds or actively managed funds (or some combination), and while the weight of the evidence seems to be in favor of index funds, selective advertising can definitely push you the other way. Of course, all of this begs the question of why there is a cottage industry of magazines, web sites, and consultants, devoted to telling you which mutual funds are winning lately? I cannot explain that, but then again, I also cannot explain why the newspaper keeps publishing my horoscope.[55]

Finally, I see nothing really wrong with the fund industry's advertising practices. There is hardly a business in the world that insists on pushing its ugly tough-to-sell products as hard as its attractive ones. Furthermore, if investors insist on shunning anything doing poorly recently, and buying only recent winners, it would be very unfair to blame only the fund companies for the selective advertising practices I discuss. They should not be required to tilt at windmills. My only point, and the implicit point of many of my observations, is that investors should know this practice occurs and use that knowledge when making their decisions.

"It Is Different This Time"

Welcome to the Granddaddy of them all. It is different this time, and the old rules (e.g., valuation) do not apply. Oh yes, the cast of characters is indeed different this time. The Internet, the on-line investor, the 401K, and so on, are all relatively recent developments. Frankly, it is *always* different *all* the time. However, what is the same is far more important than what is different. Earnings and dividends still matter. All else equal, paying more for a stock or stock market must reduce your expected return. The forces of competition still exist making limitless profit growth an unlikely event. In essence, A is A, math still works.

Of all the differences this time, the one I keep thinking about is that on a host of scales, we have never, ever, seen broad market stock prices (or growth vs. value prices) near this high. We are in uncharted territory, and if there is the slightest slip, or even the slightest failure to excel in an unprecedented manner, the long-run will not save us. Yes, come to think of it, perhaps we should be careful what we wish for, as if it really is different this time, it might not be a good thing.

Instead of "It is different this time," I prefer the French *Plus ca change, plus c'est la meme chose*.[56]

VI. Conclusion

Reading this article, one might conclude that I am anti-Wall Street. Nothing could be further from the truth. There is not one government regulation I would offer to fix any of the above. I believe in caveat emptor, and I believe Wall Street is, and should be allowed to be run as a business, selling a product. Furthermore, I strongly believe that we are all far better off with a free and unfettered Wall Street pursuing profits. In fact, it is one of our system's biggest advantages. Even if there are abuses (some of which I detail) I think the alternatives are uniformly worse. I simply believe it is very important for investors to recognize that Wall Street is not an independent source of academic research, rather they are a manufacturer with a huge vested interest in supporting their product. I also think it must be recognized that a host of financial media (e.g., financial T.V. networks, the latest personal finance magazine each week, etc.) are also much better off in an ongoing bull market, and perhaps act with a slant towards perpetuating this state.[57] We all act in our own interest and probably with a bias (intended or not) towards arguments that benefit us.[58] This article is no exception, and thus I do not condemn this activity, I simply point it out, and analyze some of the logic that flows from these observations. I come to the conclusion that these various forms of "bubble logic" have in all likelihood contributed to, or even led to, a situation where stocks are dangerously expensive.

The question of whether we are currently in the grip of a gigantic financial bubble, particularly in the growth/tech sector where I argue many investors have mistaken earnings growth for expected return and great companies for great investments, makes the issues I discuss of no small consequence to our collective prosperity. Put simply, there are really three possibilities for the broad market (with all three being more extreme for the growth/tech sector),

1. Investors understand and are now comfortable with a very low expected return on the stock market going forward.
2. We are in for an exceptionally long period of exceptionally high growth in real earnings that justifies today's market prices.
3. Most investors are not really thinking about either 1) or 2), but are engaged in wishful thinking, believing in hype and slogans, focused on irrelevant short-term stories, or forced to be in stocks by circumstance (e.g., many mutual fund managers), and all this is coming together causing a massive financial bubble. If true, this bubble can only end with a tremendous stock market crash, or a very long period of stagnation.

Earlier in this article I argued strongly that 1) and 2) are highly unlikely. Paraphrasing Sherlock Holmes, "when you eliminate the impossible, the improbable must be true." While "impossible" is far too strong a word, my rejection of 1) and 2) unfortunately leaves 3) as my favored candidate. In fact, while one is never able to prove an assertion about the economy or the stock market as would a logician, when one does the math, the overvaluation of the market, and of the growth/tech sector in particular, is the closest thing to a proof we will probably ever see. Unless we see 20-year growth for the S&P far far in excess of anything ever seen for 125 years starting from similar good times, long-term S&P returns become quite ugly. If we do see such unprecedented growth, the long-term returns become merely acceptable. For growth/tech, if we do see future growth in-line with Wall Street's gigantic and unprecedented expectations, the long-term return to today's buy-and-hold investor is

still exceptionally poor. If we go past perfection, and assume mythological growth for the entire growth/tech market, long-term returns might then achieve mediocrity. Yet, this math is still ignored, and short-term stories, greed, and ignorance still prevail.

I do think more voluntary intellectual honesty on many of the points of this article would benefit Wall Street's customers, and ultimately (perhaps not short-term, but long-term) Wall Street itself. Nevertheless, it is difficult for Wall Street to suddenly sound a siren call of warning when it is not what investors want to hear. It is far easier to focus on a bevy of distracting and ultimately irrelevant short-term phenomena (e.g., the Fed is stopping! earnings this quarter are great!), rather than the math. However, the short-term matters little, and the long-term is too important a thing to be left to bubble logic.

Endnotes

[1]I would like to thank Mark Anson, Rob Arnott, Brad Asness, Jonathan Beinner, Peter Bernstein, William Bernstein, Barton Biggs, Chris Campisano, Mark Carhart, Anne Casscells, Kent Clark, Roger Clarke, Tom Dunn, David Dykstra, Ken French, Deepak Gurnani, Ron Gutfleish, Brian Hurst, Antti Ilmanen, Ronen Israel, David Kabiller, Larry Kohn, Robert Krail, Oktay Kurbanov, Josef Lakonishok, John Liew, Mani Mahjouri, George Main, Todd McElroy, JeffMora, Peter Muller, George de Nemeskeri-Kiss, Tom Philips, Paul Samuelson, Salim Shariff, Robert Shiller, Meir Statman, Ross Stevens, and Steven Thorley for helpful suggestions and comments.

[2]Disclaimer: Please note that the views expressed in this article are purely the opinion of the author. Furthermore, the information set forth herein has been obtained or derived from sources believed by the author to be reliable. However, the author does not make any representation or warranty, express or implied, as to the information's accuracy or completeness, nor does the author recommend that the attached information serve as the basis of any investment decision. This document has been provided to you solely for information purposes and does not constitute an offer or solicitation of an offer, or any advice or recommendation, to purchase any securities or other financial instruments, and may not be construed as such. This document is intended exclusively for the use of the person to whom it has been delivered by the author and it is not to be reproduced or redistributed to any other person. All calculations performed by the author contained herein are subject to error. This document is subject to further review and revision.

[3]A piece by Shawn Tully in the January 24th, 2000 issue of *Fortune* actually covers some similar ground (with a similar viewpoint). As his came first, my work owes a clear debt to Mr. Tully's article. Also, after writing most of this article, a letter to clients by Cambridge Associates entitled "Do the Math" was pointed out to me. This letter parallels many of the arguments of this article (including repeated exhortations to "do the math"), and some of the quotes I use are taken from this piece. I think this piece is excellent and recommend it if you can get your hands on one (and not just because of the similarity in content to my work). However, mine is funnier.

[4]Siegel does not say that stocks will always win over any 20-year period. He only points out how consistently well stocks have done, but gives no silly assurances going forward.

[5]Before discussing the implications of this graph, I feel a need to defend Shiller for a moment. He has been criticized by some for using a 10-year average of real earnings as the E in P/E. The logic of this attack is that since earnings on average go up, a 10-year average of past real earnings will on average understate current earnings, and thus arrive at a larger P/E than just using last year's earnings for E. This is true. But, Shiller makes very little hay about the fact that his version of P/E is currently in the 40s while the trailing P/E over one year (for the S&P 500) is currently in the low 30s. That is not his point. His point is that comparing his series (using 10-year real earnings) to itself over time is an absolutely legitimate exercise. In fact it is a necessary one. Especially

over long periods that include some true earnings depressions, short-term earnings can be very misleading (e.g., in the early 1930s there were some very distressed times when 1-year earnings dropped precipitously thus raising P/E, but in fact these were times better described by a low P/E). By using 10-year real earnings, Shiller simply restates P/E in a more stable meaningful fashion. The fact that P/Es tend to be higher in this method is immaterial. Now, the gap between Shiller's 10-year P/E and the more conventional 1-year P/E is indeed recently larger than normal (though examining the 1-year P/E graph also shows startlingly high recent prices). Again, the 10-year P/E will generally be under the 1-year P/E because earnings grow over time, and thus a 10-year average is usually below current earnings. Because the last ten years' real earnings growth has been somewhat stronger than average, this gap between the 10-year and 1-year P/E is greater than average. This does not mean Shiller's series is misleading. If one believes that this recent strong earnings growth will continue then this is a legitimate reason to argue that today's high P/E might be partly justified, but it does not make Shiller's version of the P/E any less a correct indicator of price (also, note that this view would have been exactly wrong in 1990 following 10 years of below average earnings growth).

[6]Some will say indeed there is magic in long-run equity returns, namely a tendency to "mean revert." The argument states that after good (bad) periods equities tend to offer less (more) attractive expected returns, and this induces a less volatile long-term return to equities than if the stock market truly followed a random walk (with drift). However, mean reversion is not the main explanation for equities consistency over the last 125 or so years. First, there is only modest statistical evidence that such long-term mean reversion exists. Second, over the last 125 years, the high average returns of equities are far more responsible for the stock market's long-term consistency than any contribution from mean reversion. The mean reversion that does exist may be responsible for equities' 20-year variance around their high average real returns being lower than we would otherwise forecast, but this is a second order effect. The high level of the average real returns is first order. For more on this topic, please see a soon to be available paper, "It's the Mean, Not the Reversion" (Asness 2000). Finally, while perhaps an attractive long-term property, anyone long equities right now should pray that there is not a lot of mean reversion in stock returns.

[7]See Brad Cornell's book *The Equity Risk Premium* for an excellent readable review of these theories.

[8]The reinvestment income on the dividends also matters. For an excellent detailed decomposition of what drives equity returns, see Bogle "The 1990s at the Halfway Mark" (*Journal of Portfolio Management*, Summer 1995).

[9]Some examples of other valuation measures are the market's dividend yield, price-to-article ratio, and Tobin's Q. For a nice overview of Tobin's Q in particular, and a bearish view who's vehemence might just exceed my own, see the 2000 book by Andrew Smithers and Stephen Wright, *Valuing Wall Street*. Their article also does an excellent job of pointing out that the stock market (even over the more reasonably priced past) is not immune to periods of negative performance that, for all practical matters, would greatly impact investors with real world investing habits and time horizons (i.e., a 20-year buy-and-hold horizon is probably too long).

[10]This focus on positive returns is somewhat arbitrary. If over 20 years you make just a bit more than zero you are positive, but not economically much better off than making a bit under zero. However, the market has focused great attention on the question/assertion of equities' long-term infallibility as defined in this section.

[11]For technical sticklers, we are ignoring differences between means and medians, and some compounding issues. However, the intuition works fine. For excellent discussion of related mathematical issues see two books by Kritzman, *Puzzles of Finance* (2000) and *The Portable Financial Analyst* (1995).

[12]The argument states that low inflation and interest rates makes the earnings yield (the inverse of P/E) of equities more attractive vis a vis bonds. This is a difficult argument to make as it ignores

the fact that presumably expected nominal growth moves with inflation. In particular, Modigliani and Cohn (*Financial Analysts Journal*, March/April 1979) argue that investors mistakenly make this comparison of equity yields to nominal bond yields. Furthermore, they argue that in the high inflation environment of the late 1970s this mistake (and related issues involving depreciation and liabilities) led investors to systematically undervalue the equity market (they estimated by about 50% at the time). In other words, Modigliani and Cohn rejected the idea that high nominal interest rates meant equity yields must be high (and P/Es low), and correctly forecasted the ensuing bull market. Applying their logic now obviously leads to the possibility of investors overvaluing equities today in our low interest rate, low inflation environment. Finally, Asness (*Financial Analysts Journal*, March/April 2000) looks at the issue empirically and finds that in fact low interest rates do support a higher than normal P/E on stocks, but only for the short-term. For long-term investors a high P/E is still very bad news. In other words Asness finds that in the short-term investors mistakenly act like stock yields should move with nominal interest rates, but in the long-term discover that they should not.

[13]A related abuse of the long-term argument for equities is when people apply it to individual equities. For example, "xyz.com might be massively overvalued but if I hold on long enough I will be fine." Siegel's work does not even begin to apply to these situations. Invest in a massively overvalued single stock and you may get lucky, but the odds are stacked against you, and having a long time horizon will not save you.

[14]For more on this issue, I would refer the interested reader to a very readable article by Jane Bryant Quinn "Wave the Bubble Goodbye." in the April 24, 2000 issue of Newsweek, and to another article by Ms. Quinn entitled "It's Not Dumb to Own Bonds" subtitled "Stocks are risky even if you hold for the long term. Investors also need something safe." in the June 19, 2000 *Newsweek*.

[15]I will not dwell on the technical details here, but one of the main contributions of the Fama and French article cited earlier was developing a methodology to estimate expected stock returns relatively free of this bias.

[16]If it helps to visualize this, just imagine Abby Cohen raises her recommended allocation to stocks by 5%.

[17]Siegel also mentions that perhaps the fact that we are now on the high side of this range (actually well past the high side) can be explained by the relatively good times we are experiencing, perhaps due to technological advances. However, Professor Siegel himself, in his aptly titled piece "The Shrinking Equity Premium" (*Journal of Portfolio Management*, Fall 1999), points out that returns to technological progress historically have gone more to workers in the form of higher real wages than to the value of companies. Quoting from his article, "Optimists frequently cite higher growth of real output and enhanced productivity, enabled by the technological and communications revolution, as the source of this higher growth. Yet the long-run relationship between the growth of real output and *per share* earnings growth is quite weak on both theoretical and empirical grounds."

[18]This quote was taken from the Cambridge Associates piece cited earlier. Their piece also contains many other quotes and parallels for the reader who is interested in more. In particular, my favorite involves the prevalence of books in the mid and late 1920s bearing a positively eerie similarity to Siegel's *Stocks for the Long Run*. One, by Edgar Laurence Smith, entitled *Stocks as Long Term Investments* "proved" that it was close to impossible for the stock market to lose over any 15-year period. Well, I guess every bubble has its Boswell (though in fairness to Siegel he did publish the first draft of his article in 1994, a time of financial distress not a bull market, and he has recently written articles, some of which I quote here, documenting that expected equity market returns are now lower going forward).

[19]Someone really paying attention might note that the gap between the 10-year and the 1-year P/E was even larger in 1929 than today. Meaning, if Professor Shiller did his analysis in 1929, the screams that he was being unfairly bearish by using the 10-year measure would have been even louder than they are today, and of course tragically wrong.

[20]Paul Krugman makes a similar point in a piece called "A Self-Defeating Prophesy."

[21]Even this is debatable. Some authors argue that looking at only the success of the U.S. market over the long-term biases us towards thinking the true equity premium is higher than it really is as ex post the U.S. has been the world's most successful market, and one of the few markets in continuous uninterrupted long-term existence. In fact, that the U.S. would survive at all was by no means a certainty, and probably upwardly biases historically based estimates of the expected return on U.S. equities. The technical term for this is survivorship bias. See Goetzmann and Jorion (*Journal of Finance*, June 1999). It is also another reason why the historical average might be a poor predictor of future stock returns.

[22]While I only pick Cisco as a very recognizable example of a general phenomenon, it should be mentioned that as of June 2000 I am short this company both personally and professionally.

[23]For our purposes, we ignore some of the issues that people valuing a company like Cisco like to fight about (e.g., the dilutive effects of options issued, the impact of the choice of accounting methodology for mergers, etc.) though these issues are probably relevant. The source for the EPS data on Cisco is COMPUSTAT.

[24]A recent study by Chan, Karceski, and Lakonishok (2000) with the working title "The Persistence in Operating Performance Growth" shows that Wall Street analysts have some power to forecast the next two years earnings growth, but even over this short period you still want to discount their optimistic forecasts as they tend to overdo it. Even worse, out further than two years, the analysts have almost no forecasting power. This is a far cry from the optimistic assumption used here that analysts are 100% correct for five full years, and even after five years are still directionally correct.

[25]To estimate each year's dividends I estimate a payout ratio and multiply by that year's earnings. The payout ratio I assume is based on the following function: payout(t) = 10% + 84% * payout(t-1) –38% * earnings growth(t). Payouts above 100% or below 0% are set to that respective boundary. This function is empirically estimated based on the annual payout ratios and growth rates of the S&P 500 through time. It captures that payout ratios are slowly mean reverting, and that payout ratios are lower (often zero) for high growth companies. The function settles to a steady state long-term payout ratio of 50% at 6% nominal growth. The results of this paper are not very sensitive to this specification.

[26]Note, there are complicated issues about uncertainty that we do not address here. I assume that Cisco's earnings evolve steadily and deterministically, and I discount the cash flow to investors from these earnings at the constant IRR that equates the present value of this cash flow to the current price. I also assume that when analysts forecast 30% annual earnings growth for Cisco, they are forecasting that Cisco's earnings in 5 years will be 1.30^5 times their earnings today (it is not fully clear what analysts' are actually forecasting). This is a different, and more aggressive, earnings growth assumption than assuming that the average growth each year is 30% (which would lead to lower future earnings as variance in per year earnings growth lowers total compound growth).

[27]Of course, optimists can also just assume that real GDP grows faster than 3% for the next 20 years. However, even assuming more aggressive growth in real GDP, the required earnings growth for Cisco is still shocking. For instance, if real GDP grows at 5% for the next 20 years, instead of 17x faster, Cisco only has to grow 12x faster than GDP to reach a 10% IRR at today's prices. We will retain the assumption of long-term 3% real GDP growth through this article, though relaxing this assumption has a minimal impact at reasonable levels. At unreasonable levels, long-term phenomenal real GDP growth can perhaps save the day (and perhaps this is what is being assumed by the market).

[28]I experimented with an alternative methodology. At the end of 20 years, assume Cisco is selling for a 15 P/E and calculate the IRR over this period assuming you sell your stake then (i.e., no more infinite horizon). This methodology is much less stable, and more arbitrary, than that employed above. For instance, in our base case scenario (1) (10 year abnormal growth, 30% growth in the

first 5 years) the IRR under this new method was 3.4% (vs. 7.5% in the full analysis), when assuming 54% growth for 5 years in scenario (2) the IRR was 10.5% (vs. 10.0% in the full analysis), and when assuming abnormal growth lasts for 20 years in scenario (3) the IRR was 8.5% (vs. 9.0% in the full analysis).

[29]To put this in perspective, currently, Cisco's trailing earnings are about 0.04% of nominal GDP, Microsoft's are about 0.09%, and GE's are about 0.11%. At 17x real GDP growth, with real GDP growing at 3% per year, Cisco's earnings will be about 0.70% of GDP in 20 years. In other words, in percentage terms, Cisco's earnings will be about 3x the current importance to the economy of Cisco, Microsoft, and GE combined.

[30]When I say "my analysis" I must clarify. I carry out the analysis here, but internal rates of return and discounted cash flow analysis were hardly invented by me. On the other hand, I did invent the Internet.

[31]This methodology is not perfect as capitalization weighting is not perfectly accurate for this task, but it will suffice. In fact, if all earnings were positive, I believe the right way to do this would be to weight the earnings growth rates by dollars of earnings not market capitalization. Because higher P/E firms are probably faster growers, weighting by market capitalization probably overstates the earnings growth of the entire index, and is another potential source of optimism in our approach. Unfortunately, when you hear the expected earnings growth of the market quoted by the financial media and Wall Street, while it is not clear exactly what they are doing, it is highly probably that they are using the biased high forecast. In fact, when they quote the trailing growth of the index, they most probably combine this bias with that described in a later footnote (the substitution bias where they quote the trailing growth of firms in the index today, perhaps added specifically because of recent strong earnings growth, not the firms actually in the index over the period in question). In other words, there is a large chance that the headline growth numbers we hear for the market are perilously close to gibberish (though again, it is hard to know for certain how the quoted numbers are really calculated).

[32]These assumptions are optimistic for many reasons. As mentioned earlier, historically analysts are overly optimistic with their five year growth forecasts. In addition, there is some statistical evidence that 5 year earnings growth is actually negatively autocorrelated, meaning that if earnings grow faster than trend for 5 years, all else equal we would guess slower than trend for the next five years. Thus, clearly my assumptions that the first five years match Wall Street's huge forecasts and that years 6-10 are still above normal are, again, very optimistic. Other optimistic assumptions are explained in the text and other footnotes.

[33]This data series comes from Professor Robert Shiller's website. It is important to note that this data series represents the real earnings of the current S&P 500 firms each year. The 20-year growth of this series is not the 20-year growth of the firms you would have bought at the start of each 20-year period. Since the S&P 500 replaces unsuccessful firms with successful firms, this is likely to bias the compound growth rates we calculated here to be higher than the growth rate of the actual firms in the S&P 500 at the start of each 20-year period. For instance, the July 5th *Wall Street Journal* (page C2) reported that last year's earnings growth for firms currently in the S&P 500 is expected to be 17% this quarter, but only 12% for the firms that comprised the S&P 500 one year ago. I do not know the extent of this bias through time, but it is yet another reason why this analysis, and in fact most public reports of the market's historical earnings growth, probably err on the side of being optimistically bullish. I should note that I am still looking into how S&P reports this data, and there is a small chance they somehow remove this potential bias in the historical data.

[34]The numbers here are smaller than the numbers more commonly used as estimates of average annual real earnings growth for the stock market, since I report the average long-term compound growth not the arithmetic average growth. Variance in annual growth will cause the compound growth to be below the arithmetic average growth. For comparison, arithmetic average annual growth is around 3% over this period.

[35]This result is not driven by large transitory components (real or measurement error) in one year earnings that bias that year up, and thus the next 20-year growth down. First, I lagged the earnings divided by 20-year average calculation by one year (in other words, I used one year old data on earnings divided by 20-year average earnings to decide if the current period was a high or low period) and the results were essentially unchanged. Next, I redefined this variable as the 3-year average of real earnings divided by the 20-year average, and again, results were very similar.

[36]In my version of the story the Goldilocks economy gets eaten by the bears.

[37]While probably better than having no customers, a shortage of parts and labor is a very big deal. Bottlenecks like these are part of the reason why extreme high earnings growth is very difficult to sustain over long periods, and thus these events should not be dismissed as a mere abundance of riches.

[38]Some analysts do claim to have valuation models that justify, say, a 100 P/E for a stock growing at 20% per year in a low inflation environment. The only way I can see this occurring is if you assume the 20% growth goes on for a really really long time (and far longer than I am willing to assume). Frankly, I just do not see how these strategists come up with assumptions that justify >100 P/Es on any stock other than one with a tiny market capitalization going through an exponential growth period or a distressed stock with temporarily near zero earnings. Other ways analysts try to analytically justify prices is through certain heuristic measures; for example, the much discussed PEG ratio (the ratio of P/E to assumed growth – the lower the better). All of these measures are distorted ad hoc attempts to simplify IRR and discounted cash flow analysis. The PEG ratio suffers seriously from ambiguities in how long growth goes on, in whether a ratio is the right functional form to make this comparison, and in our lack of a benchmark for what constitutes a high or low number. As a quick one-stop measure perhaps the PEG ratio has use in relative value, but when it disagrees with thoughtful full analysis, it must be cast aside.

[39]I will focus on this more later, but there has also been little or no correction in stocks' relative pricing. While the Internet has come way down, for the first six months of 2000 the S&P/BARRA Growth index has still beaten the S&P/BARRA Value index.

[40]Specifically, the total return from 12/31/99 to 6/30/00 of the S&P 500 is –0.4% and of the NASDAQ 100 is +1.5% (source Bloomberg).

[41]I do have to admit though, not everything is unchanged from six months ago. The Fed has sharply raised short-term interest rates, inflation seems to be stirring, and the economy seems to be slowing. However, we will see soon that the economy slowing is really good news.

[42]My favorite ridiculous comment of 1999 was when several analysts commented that Internet stocks should go up if rates rise, as they were sitting on a ton of cash whose yields would rise. I will not go through the math, but suffice it to say that whoever said this should be embarrassed. Well, at least these companies do not seem to have the problem of excess cash any more.

[43]Being from the University of Chicago I might just have to rethink this one.

[44]Remember, the first step towards quitting market timing is admitting you have a problem. Seeking help from a higher power also does not hurt (I do not mean Alan Greenspan).

[45]In particular, sometimes months are used instead of days.

[46]This might be a good time to mention that there is nothing necessarily wrong with being short-term. Short-term strategies, and short-term momentum strategies in particular, might have validity (see for instance Jegadeesh and Titman (*Journal of Finance*, March 1993) and Asness (AQR Capital Management working paper, 1999)). However, these strategies are probably not very applicable to the average investor, and tend to go away or become wildly unstable if too many try to follow them. Clearly the majority of us should be focused on the long-term.

[47]While the researchers generally agree that investments in "great" firms have lost out over the long haul to cheaper firms, they do argue over why. Some argue that investors in expensive firms are making a mistake and over-extrapolating recent success too far into the future, and thus pay too much for these firms. Others argue that investors might not actually be overpaying for perceived

greatness; rather, they might be willing to accept a lower expected return on these companies as perhaps these companies are less risky than their "non-great" counterparts. See Fama and French (*Journal of Finance*, June 1992) and Lakonishok, Shleifer, and Vishny (cited above) for two sides of this debate. Finally, some argue that the entire result itself is an accident of the data, and will not necessarily hold up going forward. See Black (*Journal of Portfolio Management*, Fall 1993) for an example of this point of view. In my opinion, this last argument might be difficult to support now that researchers have found the same effect in many countries, and in previously unexamined U.S. data from 1927–1963.

[48]Also, see Laurence Siegel and John Alexander "The Future of Value Investing" (forthcoming in *The Journal of Investing*, 2000) for an excellent review of value's history, and a study of the forward looking prospects.

[49]Hereafter my worthy co-authors will be obnoxiously referred to as "et al."

[50]For presentation purposes, all numbers in this figure are quarterly moving averages.

[51]Chan, Karceski, and Lakonishok (Working paper, March 2000) look at this in an alternative (and interesting) way. While Asness et al. look at IBES forecasts of future earnings growth, Chan et al. look at actual past growth in operating performance. However, while the approach is different, the answer is the same. They find that past growth in operating performance cannot explain the recent performance of growth vs. value stocks. Furthermore, they conclude "...the assumptions about the sustainability of high growth needed to justify large growth stocks' relative valuations are quite bold."

[52]Robert Shiller advances related, and much more fleshed out, arguments in *Irrational Exuberance* (and since he came first, obviously my thoughts were greatly influenced by his work).

[53]There is some academic evidence that a strategy of buying after a split might have validity (it is difficult to disentangle these results from other momentum trading strategies). However, while some self-fulfilling success might occur, the tiny statistical excess returns documented in academic studies are clearly not what all the excitement is about.

[54]See Carhart (*Journal of Finance*, March 1997).

[55]Unfortunately so far, mine keeps saying "short more."

[56]Meaning, "The more things change, the more they remain the same." After holding out this long, the bubble has finally driven me to quote the French.

[57]One can imagine the chief of a financial news network telling a nay-saying photographer not to worry, "you provide the pictures, I'll provide the bull market."

[58]Note, I think this bias occurs naturally and tacitly, I am not a conspiracy theorist. Alan Greenspan, Abby Cohen, and Colonel Sanders are probably not meeting together in Geneva to plot the bubble's continued expansion over a bucket of extra crispy.

This page intentionally left blank

Fight the Fed Model

The relationship between future returns and stock and bond market yields

Clifford Asness

Tune to CNBC or the like for more than about 15 minutes, and you will hear a strategist, portfolio manager, or market pundit of some stripe explaining that the high market multiples of recent times are justified by low interest rates and/or inflation. "Well, Maria, you have to understand—stocks might look expensive, but that is fine because interest rates and inflation are low." Or so the refrain goes. In fact, to many on Wall Street and in the financial media this assertion has been elevated to the status of conventional wisdom.

The most widespread version of this comparison of stocks to bonds is often deemed the *Fed model.* This model, allegedly found in the annals of a Fed report, not named because of any official Fed endorsement, comes in various forms, but generally asserts that the stock market's earnings yield should be compared to current nominal interest rates (the earnings yield, or E/P, is the inverse of the well-known price-to-earnings ratio or P/E).[1]

Letting Y represent the yield on ten-year Treasuries, the model says we should look at E/P versus Y. In its simplest form, it asserts stocks are cheap when E/P exceeds Y, expensive when Y exceeds E/P, and fairly valued when Y and E/P are equal.

Even pundits who are united in their belief in the Fed model do not always agree on what it is telling them. Of course, as recent times make clear, the E in E/P is not a simple observable number. In addition, some adjust the basic comparison of E/P and Y for a growth assumption or a required equity risk premium, or change the functional form of the relationship.[2]

Reprinted from *The Journal of Portfolio Management*, Fall 2003, Vol. 30 (1), with permission.
Recipient of the 2005 Bernstein Fabozzi/Jacobs Levy Best Article Award.

The basic widespread core belief implied by the Fed model, though, is that the stock market's E/P must be compared to Y, and that low interest rates permit a low E/P or, equivalently, a high market P/E (and vice versa). It is this core belief (whether or not it is labeled the Fed model) that I study here.

The evidence strongly suggests that the Fed model is fallacious as a tool for long-term investors. Essentially, the comparison of E/P to Y is erroneous as it compares a real number (P/E) to a nominal one (Y). The important point is that the stock market's P/E does not have to move with inflation since nominal corporate earnings already do so. Empirical evidence supports this theory. Investors forecasting future long-term stock returns would do much better relying on simple P/E, or the like, rather than the Fed model.[3]

While the Fed model fails as a predictive tool for future long-term stock returns, it does work as a descriptive tool for how investors choose to set current stock market P/Es. Even here, however, the simple Fed model needs help. Applying a relationship studied in Bernstein [1997b] and Asness [2000], it is clear that the Fed model relationship must be conditioned on the perceived volatility of stocks and bonds. Without conditioning on perceived volatility, the simple Fed model is a failure over 1926–2001, even to describe how investors set P/Es. Conditioned on perceived volatility, however, the Fed model explains the puzzle of why the relative yield on stocks and bonds has varied so greatly over the last century.

Note that this finding that the Fed model has descriptive power for how investors set P/Es in no way contradicts the finding that the Fed model fails as a predictive tool for stock returns. If investors consistently err and follow a poor model, it is not surprising that this same model fails those investors for making long-term forecasts.

Data and Terminology

The data used in this article include:

- Monthly U.S. CPI inflation (continuously compounded).
- Monthly continuously compounded total real (after inflation) return of the S&P 500 and of the ten-year U.S. Treasury bond from 1871 through 2001. These monthly returns are added together to derive longer-term holding-period total returns.
- The price-to-earnings ratio (P/E) of the S&P 500 based on ten-year trailing earnings. Each month earnings-to-price ratios based on last year's trailing earnings are multiplied by the S&P 500 price index to determine a monthly earnings per share (EPS) estimate for the index. Each EPS estimate is then divided by the level of the CPI, and averaged over the last ten years to determine a ten-year average real EPS figure for the S&P 500. Finally, the current real price index is divided by this average real earnings figure to determine today's P/E ratio. Ten years of earnings are used in an effort to smooth out short-term transient fluctuations (following Shiller [2000]). Unless otherwise indicated, P/E refers to this measure.[4]
- The yield each month on the ten-year U.S. Treasury bond (Y).

All data sources in this article are, unless otherwise mentioned, the same as those used in Arnott and Asness [2003] or Asness [2000].

Arguments for and against the Fed Model

There are a variety of arguments for why P/Es should or should not move with nominal interest rates.

Common Sense Rationale for the Fed Model

At first glance, the Fed model seems to be simple common sense. I will soon disagree with these widely believed arguments, but it's important to give the devil his due (but not to be his advocate).

Argument #1—The Competing Assets Argument. Many reason as follows. E/P, the annualized earnings on stocks divided by the price paid, is the yield you receive on your equity investment.[5] Y is the yield you get on Treasury bonds (ten-year Treasuries for this comparison).

Investors can invest in either stocks or bonds, and thus these are competing assets. Therefore, the comparison of E/P and Y is valid and important. When E/P exceeds Y, stocks are yielding more than bonds and are thus cheap, and when E/P is lower than Y, stocks are expensive. E/P = Y is the implied fair value point.

Argument #2—The PV Argument. There is a slightly more sophisticated (although ultimately similar) version of argument #1. Some correctly point out that the price of a stock today is the discounted present value (PV) of the future cash flows to investors from the company or market in question (the famous dividend discount model or DDM approach). They argue that when interest rates fall, the PV today of future cash flow rises, and P/Es should also rise.

As an example, imagine the yield on the ten-year Treasury bond trading at par value is 10%. Well, viewing the 10% annual yield as income, the P/E on the bond is 1/10% = 10. Now, imagine that the ten-year par bond yield is 4%. Well, now the Treasury's P/E is 1/4% = 25.

Exhibit 1 S&P 500 E/P and Ten-Year Yields

Argument #2 says it would not be surprising to see stocks selling for higher P/Es when interest rates are 4% than when they are 10%, as the P/E on bonds is also higher.

Argument #3—Just Look at the Data. The final argument in favor of the Fed model is empirical. Exhibit 1 shows the stock market's E/P and the yield on the ten-year Treasury over 1965–2001.

Historically E/P and Y have been strongly related (with perhaps a small level shift down in E/P post-1985). The correlation of these two series over this period is an impressive +0.81. It's a rare Wall Street strategist who in the course of justifying the Fed model does not pull out a version of this graph, or an analogous table (showing that stock market P/Es move with either interest rates or inflation). The implicit argument is that high P/Es are fine if interest rates and inflation are low, as this is normal.

Why the Common Sense Is Likely Wrong

It is important to review these pro-Fed model arguments because belief in them is widespread. Yet obviously I have set up this ersatz common sense for a fall.

Let us start with the well-known Gordon model, which expresses the expected nominal return on the stock market as the dividend yield plus the expected growth of dividends:

$$E[R_S] = D/P + G_D \tag{1}$$

where $E[R_S]$ is the expected nominal stock return, D/P is the current dividend yield (current dividends per share divided by current stock price), and G_D is the assumed constant long-term nominal growth rate of dividends. Capital letters will represent nominal (before-inflation) values, while lowercase letters will represent real figures after accounting for inflation (e.g., $g_D = G_D - I$, where I equals inflation and g_D is thus, as a linear approximation, the expected real rate of dividend growth).

Dividend yields can be linked to earnings yields by the payout rate, PAY = D/E, the proportion of earnings paid out as dividends:

$$E[R_S] = PAY \times E/P + G_D \tag{2}$$

Now make some simplifying assumptions. First, use ½ for PAY, which is about its long-term historical average.[6] Furthermore, assume that PAY is constant, so the growth rate of earnings and dividends is the same. Equation (2) can be rewritten as:

$$E[R_S] = ½\ E/P + G_E \tag{3}$$

where G_E is the growth rate of earnings, which equals G_D.

All else equal, if expected nominal stock returns are higher, the higher the earnings yield at purchase (or, equivalently, the lower the P/E) and the higher the expected long-term nominal earnings growth.

Now, expected real stock returns are (approximately) expected nominal returns minus inflation (assumed to be a known constant I):

$$\begin{aligned} E[r_S] &= E[R_S] - I \\ &= \tfrac{1}{2}\, E/P + G_E - I \\ &= \tfrac{1}{2}\, E/P + g_E \end{aligned} \tag{4}$$

This is an important equation. Expected *real* stock returns are a positive function of starting E/P (or a negative function of P/E) and expected *real* long-term earnings growth.

The key issue is what happens when expected long-term inflation falls. Let's make some reasonable educated guesses. First, let's assume that nominal bond yields fall one-to-one with the fall in inflation. Next, as a starting point, let's expect an equal fall in the long-term nominal return on stocks.

In other words, as a starting point it is probably a good guess that the required *real* return on stocks does not go up when long-term inflation goes down. For instance, if expected nominal stock returns were 10% in a 5% expected inflation environment (5% real return), it would not be reasonable to expect 10% in a 2% inflation environment (8% real). Rather, a more reasonable guess is 7% nominal (5% real return).

If inflation falls, but expected *real* stock returns are to stay the same, expected *nominal* stock returns must fall. Equation (3) makes it clear that either E/P must fall (P/E rise), or G_E must fall. Fed model advocates would have you believe that the E/P must change, so when inflation falls, E/P must fall, and P/E rise. Of course, there is another obvious possibility ignored by Fed model proponents. Instead of E/P moving, G_E can move to partly or completely offset changes in inflation.

In fact, simple economic intuition argues that a G_E move is the likely scenario. Imagine a known permanent instantaneous shift in expected inflation.[7] Is it not plausible, at least as a first guess, to forecast that nominal revenue and expense growth move by the same amount (after all, is that not inflation?), and that long-term G_E moves with the change in expected inflation?

For instance, when expected inflation is very low (as in recent times), pricing power is low (for both firms and labor), and profits grow more slowly in nominal terms. To put it differently, isn't it plausible that real earnings growth (g_E) is largely insensitive to the level of constant known inflation, as inflation is a largely monetary (not real) phenomenon?

Empirical tests of the historical relationship of expected long-term inflation and nominal earnings growth are not straightforward. First, there is a dearth of independent long-term periods to observe, and second, inflation expectations are not directly observable over long periods. We can easily observe, however, actual realized inflation and actual nominal earnings growth.

The regression in Equation (5) has on the left-hand side monthly rolling decade-long nominal EPS growth on the S&P 500, and on the right-hand side the corresponding decade-long realized CPI inflation. The regression runs from 1926 through 2001 (t-statistics in parentheses):[8]

$$\begin{aligned} \text{Nominal Earnings Growth} &= 2.2\% + 0.94\ \text{Inflation} \\ &\quad\ (2.13) \quad (3.55) \\ R^2 &= 36.5\% \end{aligned} \tag{5}$$

Over this commonly studied period, realized inflation has been on average almost an exact pass-through to nominal earnings. On average, 94% of decade-long inflation showed up in

nominal earnings growth, explaining 36.5% of the variation. Using only more recent data, this relationship does become weaker, but the strong positive relationship between inflation and nominal earnings growth remains.[9]

If this seems at all counter-intuitive, consider that one of the tried-and-true reasons to own equities is the belief that stocks are a good long-term inflation hedge.[10] This conventional wisdom is equivalent to believing that expected real (not nominal) earnings growth is relatively constant. If stocks are indeed a good inflation hedge, it is precisely because the nominal earnings of companies tend to rise with nominal inflation, making stocks into a *real* asset. A pundit who believes in the Fed model but also believes stocks are a good hedge for long-term inflation is inconsistent.

This point has been made before. Most notably (and over two-score years ago), Modigliani and Cohn [1979] made this point in somewhat the opposite environment from today's. They observed that in the late 1970s investors were using the Fed model (although they did not call it that) and wrongly pricing equities to a very high E/P (low P/E) because interest rates and inflation were high. Using this logic, they effectively predicted the bull market of the 1980s and 1990s.

Also notable, in an excellent survey of many of these issues, Ritter and Warr [2002] conclude that the Fed model makes the error of money illusion or what they call the "capitalization error." Siegel [2002a] also makes many of these same points.

While there is certainly a history of others who have noted that the Fed model is erroneous, its continued popularity indicates this dissenting view is losing in the court of public (and pundit) opinion. Thus, the Fed model must be fought further (even with alliteration if necessary).

Now reconsider the specific common sense arguments #1 through #3 in light of these counter-arguments.

Refuting Argument #1—The Competing Assets Argument. Argument #1 is that stocks and bonds are competing assets, and thus we should compare their yields. Now we see that the yield on the stock market (E/P) is not its expected return. The nominal expected return on stocks should, all else equal, move one-to-one with bond yields (and entail a risk premium that itself can change over time). But this is accomplished by a change in expected nominal earnings growth, not by changes in E/P.

Refuting Argument #2—The PV Argument. Argument #2 is that when inflation or interest rates fall, the present value of future cash flows from equities rises, and so should their price (their P/E). It is absolutely true that, *all else equal,* a falling discount rate raises the current price. All is not equal, though. If when inflation declines, future nominal cash flow from equities also falls, this can offset the effect of lower discount rates. Lower discount rates are applied to lower expected cash flows.

The typical "common sense" behind the Fed model ignores this powerful counter-effect, in effect trying to use lower nominal discount rates, but not acknowledging lower nominal growth. You would be hard pressed to find a clearer example of wanting to both have and eat your cake.

It is indeed possible to think of stocks in bond terms as the Fed model attempts. Instead of regarding stocks as a fixed-rate bond with known nominal coupons, one must think of stocks as a floating-rate bond whose coupons will float with nominal earnings growth. In this analogy, the stock market's P/E is like the price of a floating-rate bond. In most cases,

despite moves in interest rates, the price of a floating-rate bond changes little, and likewise the rational P/E for the stock market moves little.

Refuting Argument #3—Just Look at the Data. Recall Exhibit 1. Historically, when interest rates or inflation are low, the stock market's E/P is also low, and vice versa. This, Fed modelers say, shows that the market does in fact set the equity market's P/E as a function of the bond yield, implying the Fed model is a good tool for making investment choices.

Pundits using this argument assume that because they show that P/Es are usually high (low) when inflation or interest rates are low (high), the Fed model is necessarily a reasonable tool for making investment decisions. This is not the case. If investors mistakenly set the market's P/E as a function of inflation or nominal interest rates, then Exhibit 1 is just documenting this error, not justifying it.

A simple analogy might be helpful. Say you can successfully show that teenagers usually drive recklessly after they have been drinking. This is potentially useful to know. But, it does not mean that when you observe them drinking, you should then blithely recommend reckless driving to them, simply because that is what usually occurs next. Similarly, the fact that investors drunk on low interest rates usually pay a recklessly high P/E for the stock market (the Fed model as descriptive tool) does not make such a purchase a good idea, or imply that pundits should recommend this typical behavior (the Fed model as forecasting/allocation tool).

The pundits often confuse these two very different tasks put to the Fed model. They often demonstrate (each with a particular favored graph or table) that P/Es and interest rates move together contemporaneously. They then jump to the conclusion that they have proven that these measures *should* move together, and investors are thus safe buying stocks at a very high market P/E when nominal interest rates are low.

They are mistaken. The Fed model, in its descriptive form, documents a consistent investor error (or a strange pattern in investors' taste for risk); it does not justify or recommend that error.[11]

Exhibit 2 S&P 500 Decade-Long Real Return Sorted by Interest Rates

To illustrate this point, and to foreshadow the empirical findings on return predictability, Exhibit 2 examines different interest rate environments over 1965–2001. It puts each month over 1965–2001 into one of five buckets based on the end-of-month ten-year Treasury yield. Bucket 1 includes all months when interest rates were in the lowest one-fifth of the entire sample over 1965–2001 while bucket 5 includes all months when interest rates were in the highest one-fifth.

The dark bars represent the average ten-year annualized real return on the S&P 500 for the decades *ending* in the month in question. For example, the first dark bar indicates that the average annual real return on the S&P 500 was an impressive 10.3% for decades ending in any month when interest rates were in the bottom quintile.

Moving to the right, we see a strong relationship as returns drop while interest rates rise, culminating in a paltry 2.0% per year decade-long real return when ending interest rates were highest.

Of course, the dark bars are relatively useless to investors, as they indicate only what *has* happened in decades preceding low and high interest rates. The light bars in Exhibit 2, however, show what happens in the average decade *following* each interest rate environment.[12] Here the story is very different. The best results actually occur in the decades *starting* with high interest rates, and, conversely, buying when rates were lowest actually led on average to *negative* real returns in the next decade.

So, when pundits say it is a good time for long-term investors to buy stocks because interest rates are low, and then show you something like Exhibit 1 to prove their point, please watch the tense of what they say, as what they often really mean is that it *was* a good time to buy stocks ten years ago—as investors are *now* paying a very high P/E for the stock market (perhaps fooled into doing so by low interest rates as I contend)—and the story going forward may be painfully different.

Other Reasons Inflation Might Matter

Now, forgetting these battling "common sense" approaches, there are some other reasons inflation might matter to P/Es. And the potential impact of each of the other reasons is cumulative and possibly offsetting.[13]

Capital gains taxation is not indexed for inflation. Thus, in a high-inflation environment, equities are unfairly burdened with taxation on purely nominal profits, and might be priced to offer higher gross returns (lower P/Es and higher E/Ps) in order to simply maintain the level of net returns after taxes. This would induce a positive correlation between E/P and Y.

Inflation can distort corporate earnings. Depreciation is taken at historical cost, and in inflationary times, cost of replacement is generally higher than recorded depreciation charges, causing the overstatement of reported earnings versus real costs. When earnings are overstated, all else equal, one might expect a lower P/E ratio (higher E/P) on reported earnings. This is, of course, like the capital gains effect above, supportive of the Fed model assertion that E/P and inflation and interest rates are positively linked, although for different reasons from those most Fed model advocates normally cite. In addition, cost of goods sold is also recorded at historical cost, so in this case, when inflation is high, costs are again understated and earnings again overstated.

Interest costs go the other way. When inflation and interest rates are high, accounting methods overstate the cost of any short-term financing; that is, even though this financing may, in real terms, be no more expensive than normal, nominal cost goes up. Similarly, for firms with long-dated nominal liabilities, accounting earnings fail to recognize the gain to shareholders from the reduction in the real value of these liabilities in the face of rising inflation. Thus, earnings are understated along this dimension at these times.

Historically, very high (and also very low or negative) inflation has been associated with uncertainty, perhaps mechanically from the cost of planning in such an environment, but perhaps also from the macroeconomic difficulties and political uncertainty that often accompany inflation extremes. This can cause investors to demand a high risk premium when inflation is high, and thus high inflation is associated with high required real stock market returns (high E/Ps and low P/Es).

For most of my analysis, the assumption is that expected real stock returns move with inflation and nominal interest rates, because investors suffer from the error of money illusion (wrongly comparing a real to a nominal quantity). Of course, the irrational case cannot be distinguished easily from the simple assertion that investors' taste for equity risk changes with inflation, and they demand higher expected returns when inflation is high (set lower P/Es and higher E/Ps). This again would mean the Fed model works for very different reasons from those its supporters generally proffer.

Perhaps most basically, the various contentions constitute an argument, not a proof. Even without the distortions above, there is no QED proof that E/P is a purely real quantity with expected real earnings growth independent of steady-state inflation—merely several arguments and some empirical evidence that make it likely so.[14]

Overall, there are quite a few reasons why inflation might matter to P/Es. Obviously, the net sign and the magnitude of all of these effects are unknown, so testing the Fed model becomes an empirical issue, with the added implication that the answer may be partial.[15]

Forecasting Returns

The central issue is forecasting power.

Regressions

The next logical step is to turn to the data and test whether the Fed model (E/P − Y) or the traditional model (P/E or E/P) has historically been a better tool for investors looking to forecast real stock returns.

Regressions are used to measure forecasting performance. The left-hand side is the real return on the S&P 500 over either a 20-, 10-, or 1-year horizon. The right-hand side is alternatively the E/P of the S&P 500 (the traditional model); the E/P of the S&P 500 minus the ten-year Treasury bond yield (the Fed model); or both the S&P's E/P and the ten-year Treasury bond yield separately in a two-variable regression.

If E/P has univariate forecasting power, it should show up in the single-variable regression, thus supporting the traditional model. If the Fed model has power, this should be seen in the test of E/P − Y. Finally, running the bivariate regression on E/P and Y separately is

useful, as E/P −Y can appear to have statistical power even if only E/P itself has actual efficacy, simply because E/P − Y can be a noisy measure of E/P itself. Also, it is possible that E/P should be compared to Y, but not at the one-to-one ratio of the Fed model.

The regressions are run over different time periods using different forecasting horizons. For forecasting 10-year horizon returns, the regressions are run over 1881–2001, 1926–2001 (the classic Ibbotson period), and 1955–2001 (the modern period when interest rates have been freely floating). For forecasting 20-year returns, the last 1955–2001 period is skipped as it constitutes very few independent periods. For forecasting 1-year returns, the latest 20 years ending in 2001 (the great bull market) are added.

Exhibit 3 provides the results of nine regressions for forecasting ten-year real S&P 500 returns. Each row represents a different regression; t-statistics are in parentheses (adjusted for overlapping observations). A row with values for only E/P or E/P − Y represents a univariate regression, while a row with values for both E/P and Y represents a bivariate regression.

For example, the first row shows that a monthly regression over 1881–2001 of overlapping ten-year S&P 500 real returns on the starting E/P of the S&P 500 reports an intercept of –0.8% (t-statistic of –0.43), a coefficient of 0.95 on E/P (t-statistic of 5.66), and an adjusted R^2 of 30.2%.

Essentially, the message of Exhibit 3 is simple. The traditional model (E/P alone) has strong forecasting power for ten-year real stock market returns, while the Fed model is wholeheartedly rejected. Expected real ten-year returns are higher, the higher the starting E/P (the lower the P/E you buy in at), and this occurs regardless of (and in fact unaffected by) the level of starting interest rates.

The Fed model itself, E/P − Y, seems to have some weak power in the earlier periods, but clearly this is only because E/P is part of E/P − Y. When E/P and Y are tested in bivariate regressions, E/P matters, and the Y part of the Fed model is ignored (with the wrong sign over 1955–2001).

Exhibit 4 shows very similar results, but with even higher R^2s (Arnott and Bernstein [2002] find a similar result for 20-year horizons). In particular, over the 1926–2001 period, the power of simple E/P to forecast 20-year stock returns is truly impressive. Now, at first glance, it again appears there is some supporting evidence for the Fed model; E/P − Y comes in with a 2.30 and 2.78 t-statistic over the two time periods. This again occurs,

Exhibit 3 Forecasting Ten-Year Real S&P 500 Returns

Date	Intercept	E/P	Y	E/P − Y	Adj. R^2
1881–2001	–0.8% (–0.43)	0.95 (5.66)			30.2%
	4.6% (2.39)			0.50 (1.41)	11.9%
	–0.8% (–0.51)	0.95 (4.13)	0.02 (0.06)		30.1%
1926–2001	–2.9% (–0.91)	1.31 (3.85)			34.9%
	5.7% (2.34)			0.47 (1.03)	9.7%
	–2.6% (–0.98)	1.37 (3.07)	–0.13 (–0.36)		35.5%
1955–2001	–2.5% (–0.57)	1.20 (3.08)			29.6%
	6.3% (2.72)			–0.36 (–0.47)	1.4%
	–2.4% (–0.53)	0.85 (0.82)	0.36 (0.44)		31.0%

however, only because E/P − Y is a noisy proxy for E/P. When E/P and Y are tested separately in the bivariate regression, it is quite clear that Y adds very little. Y does have the hypothesized negative sign, but over both time periods its coefficient is roughly one-sixth of that predicted by the Fed model (i.e., the Fed model predicts Y to have an equal but opposite sign to the coefficient on E/P), and is not statistically strong.

Finally, Exhibit 5 presents the shorter-horizon results when the left-hand side of the regression is rolling one-year real returns.

As shown by others, at shorter horizons R^2 values fall dramatically (see Fama and French [1988]). This occurs because the predictable component of stock returns is small but slowly changing, leading to reasonably reliable long-term forecasts, but poor short-term ones. In English, short-term market timing is hard.

Looking at the longest time periods (1881–2001 and 1926–2001), there is a very similar story as for 10-year and 20-year horizon returns. E/P alone has some forecasting ability (as usual, higher E/Ps are better for future returns). E/P − Y (the Fed model) has some power, but again only because it is a poor man's E/P. The period 1955–2001 is the stuff of an efficient market fan's dreams. Basically, nothing has forecasting power for short-horizon returns over this period.

Only by looking at the recent 1982–2001 bull market is there any support for the Fed model. No specification has a very high t-statistic (this is apparently too much to ask of 20-year regressions), but R^2s are high (for one-year forecasts), and in a bivariate regression

Exhibit 4 Forecasting 20-Year Real S&P 500 Returns

Date	Intercept	E/P	Y	E/P − Y	Adj. R^2
1881–2001	1.4% (0.84)	0.63 (2.59)			37.2%
	4.3% (3.85)			0.48 (2.30)	25.5%
	1.6% (0.83)	0.65 (2.52)	–0.09 (–0.24)		37.6%
1926–2001	–2.2% (–1.15)	1.22 (5.69)			65.4%
	4.6% (2.86)			0.64 (2.78)	33.9%
	–1.8% (–0.86)	1.27 (5.92)	–0.22 (–1.38)		68.1%

Exhibit 5 Forecasting One-Year Real S&P 500 Returns

Date	Intercept	E/P	Y	E/P − Y	Adj. R^2
1881–2001	–3.6% (–0.88)	1.38 (2.66)			4.5%
	4.1% (2.34)			0.82 (2.04)	2.4%
	–3.3% (–0.76)	1.40 (2.62)	–0.08 (–0.16)		4.4%
1926–2001	–9.4% (–1.64)	2.35 (3.29)			8.3%
	5.0% (2.37)			1.09 (2.14)	3.7%
	–8.3% (–1.36)	2.42 (3.38)	–0.31 (–0.54)		8.4%
1955–2001	2.0% (0.38)	0.72 (1.01)			1.3%
	6.8% (3.71)			0.57 (0.65)	0.4%
	2.5% (0.49)	0.96 (0.93)	–0.32 (–0.35)		1.3%
1982–2001	1.9% (0.29)	1.65 (1.89)			10.3%
	20.1% (4.57)			4.08 (1.84)	8.8%
	9.2% (1.16)	3.19 (1.66)	–2.05 (–1.02)		11.2%

the coefficient on Y is negative and about two-thirds the size of the positive coefficient on E/P (i.e., two-thirds of the way to the Fed model).

Now, one could simply dismiss this result as a lone and very narrow victory won over a short period for the Fed model (adding Y takes the R^2 only from 10.3% to 11.2%, and E/P −Y still works worse than plain old E/P). This dismissal is probably warranted, although the recent results do give some hope to those using the Fed model for tactical purposes.[16]

The bottom line is that for forecasting long-term stock returns the Fed model is an empirical failure, and the traditional model (regular old P/E) is a success story.

Are We Forecasting Stocks or Stocks versus Bonds?

So far the focus has been on forecasting future real stock returns, and the empirical evidence has strongly favored the traditional model versus the Fed model for this task. Still, this does not address the issue of forecasting relative (stock versus bond) returns.

Simple economic intuition as well as the findings of others (Arnott and Bernstein [2002], for example) indicates that the best and most reasonable forecast of future real bond return is the current real bond yield (Y minus forecasted future inflation or Y − I). Thus, if E/P is a real quantity as argued here, a strong candidate to forecast future stock versus bond returns would be E/P minus the current real bond yield, or E/P − [Y − I].

Furthermore, while E/P − [Y − I] might be a fair comparison, it excludes any risk premium for stocks.[17]

Thus, a very simple formula for *relative* value might look something like something like E/P − [Y − I] − RP (letting RP equal the required risk premium). When that is positive, stocks are probably more attractive than usual versus bonds, although not necessarily attractive on an absolute basis. Of course, this necessitates adding an estimate of expected inflation, and an estimate of the required risk premium, neither an easy measure to observe with certainty. While additional complication is regrettable, such additions are necessary for the equation to make any sense at all.[18]

Essentially, declaring it a relative-value tool does not save the Fed model. Even for this task, the Fed model specification of E/P − Y can be rejected on first principles.

Forgetting the fact that the Fed model is misspecified, even for relative value, an interesting practical question is what Wall Street pundits think they are forecasting with the Fed model. When they say something like "stocks are undervalued according to the Fed model," might they actually sometimes mean "stocks are overvalued, but less so than bonds"? One would hope that in this case they would actually say so, as that would perhaps be useful information to long-term investors.[19]

Instead of calling stocks cheap it would be clearer at these times to say "stocks are expensive, but bonds are more expensive." Of course, this is a less catchy sales pitch than "stocks are cheap on the Fed model."

What are the consequences of this phraseology? Consider the small investor who might hear pundits say stocks are fair or cheap, according to faulty Fed model logic. It seems reasonable that this investor might take this to mean the stock market's long-term prospective real return is favorable when compared to historical returns. Someone who is retiring, assuming fair or cheap means equities will perform up to or exceeding their

Exhibit 6 S&P 500 E/P and Ten-Year Yields

historical standards going forward, and who budgets and saves accordingly is potentially in for real trouble.

How P/Es and Nominal Rates Move Together

Our evidence should make it clear that traditional valuation (P/E) is what matters in forecasting long-term real stock returns, not the Fed model. Yet recall that Exhibit 1 demonstrates that the Fed model indeed seems to have power to describe how investors actually go about setting P/Es. I now examine this descriptive power, showing it to be genuine, but robust only over the long term if investors' changing perceptions of stock and bond risk are also taken into account.

Exhibit 1 goes from 1965 through 2001. Exhibit 6 shows the same data over the longer 1926–2001 period.

What happened? Over this whole period, E/P and Y have been correlated at only +0.18. This is in stark contrast to 1965–2001 when the correlation was +0.81. Furthermore, over 1926–1965, E/Ps were almost uniformly substantially above ten-year Treasury yields, but over 1965–2001 they were generally a bit below interest rates. Clearly, if one is unwilling to simply dismiss the 1926–1965 data, the empirical support for the Fed model (in its descriptive role) is dealt a serious blow.

An answer comes from applying the models examined and discussed in Bernstein [1997b] and Asness [2000]. They argue that the simple Fed model, even used only as a tool to document investors' error of money illusion, leaves out a crucial variable: investors' changing perception of risk. Whether in error or not, if investors compare E/P to nominal Y, why would they always demand a constant E/P = Y? Should not investors demand more from stocks when they perceive stocks to be riskier versus bonds, and vice versa?

In Asness [2000] I specify a functional form for this relationship and fit parameters to this model. Equation (6) demonstrates a highly similar specification:

$$E/P = a + bY + c\sigma_{stocks} + d\sigma_{bonds} \tag{6}$$

The motivation for Equation (6) is as follows. Even if investors erroneously move E/Ps with nominal rates, it is arbitrary to assume E/P = Y. This can be generalized in two ways. First, E/P does not have to equal Y; rather it can be any linear function of Y, as in E/P = a + bY. Next, note that the simple equation E/P = a + bY is still missing an adjustment for risk. In Asness [2000] I proxy for perceptions of stock and bond market risk by adding two new terms to (6): the prior realized 20-year volatility of equities and bonds.

Essentially, when σ_{stocks} is high versus σ_{bonds}, investors have experienced more volatility in stocks versus bonds over the last generation.[20] The hypothesis for Equation (6) is now that b is positive, c is positive, and d is negative. With c positive and d negative, it means that the weighted difference of stock and bond volatility is relevant to the level of E/P. In other words, investors do in fact (through the mistake of money illusion) set E/P as a function of nominal interest rates (positive b), but they also require a higher E/P versus Y when their generation has experienced relatively more volatility in stocks as compared to bonds (positive c and negative d).

When Equation (6) is estimated over 1926–2001, the results are as follows:

$$\begin{aligned} E/P = {} & \underset{(1.1)}{0.3\%} + \underset{(34.7)}{0.96}\,Y + \underset{(32.5)}{0.37}\,\sigma_{stocks} - \underset{(-25.4)}{0.78}\,\sigma_{bonds} \\ R^2 = {} & 62.0\% \end{aligned} \tag{7}$$

The 0.18 correlation of E/P and Y over 1926–2001 corresponds to an adjusted R^2 of 3%. The addition of σ_{stocks} and σ_{bonds} raises this to 62%—considerable improvement.

E/P is strongly related to the difference between stock and bond volatility, and conditioning on this relationship returns the relationship between E/P and Y to almost exactly the level expected by the Fed model (a 0.96 coefficient) over the entire 1926–2001 period.[21] Once volatility is adjusted for, investors have empirically moved stock market E/Ps one-to-one with nominal interest rates.

I show in Asness [2000] that this relationship, although laced with econometric difficulties, survives all robustness tests with flying colors (including working back to 1871, and working better than all competing models for out-of-sample forecasting).

Exhibit 7 plots the actual and fitted P/E from Equation (7) (inverting fitted E/Ps to get fitted P/Es).[22] While the simple Fed model implicitly produces a horizontal line as a best fit (R^2 = 3%), Equation (7) produces quite an impressive fit. The most notable errors occur at the start in the mid-1920s and in the bubble of 1999–2000, although much of that spectacular rise is captured. The peak in the fitted series in 1999 is similar to the actual peak P/E, although the fitted series does not stay there as long.

In fact, 1999–2000 is a nice example of the difference between describing how P/Es are set versus justifying them. When the fitted series peaked in the 40s in 1999, it was not saying that this P/E is rational for the S&P 500 (it was not). It was saying that, assuming investors act the way they have in the past, and given how low equity volatility had been versus bond volatility, and how low interest rates were, *such an irrationally high P/E was*

Exhibit 7 S&P 500 P/E and P/E Fitted on Y and Volatility

to be expected. The Fed model, alone or modified for volatility, offers no solace to those buying the S&P 500 at a P/E of 44, but it does explain what tricked them into doing so.

In fact, this model very neatly resolves the conundrum of why E/P and Y are very highly correlated over 1965–2001, but very weakly correlated over 1926–2001, and why E/P is approximately equal to Y in magnitude over 1965–2001, but generally dwarfs Y over 1926–1965. While interest rates were low in the first half of 1926–2001, realized stock market volatility was very high versus bond market volatility (even after October 1929 rolls out of the sample). A simple model of E/P based on nominal interest rates cannot hope to capture the fact that investors, rightly or wrongly, demanded a very high E/P versus Y over this time, largely to compensate them for their perception of very high equity versus bond risk. Over 1965–2001, the ratio of stock and bond volatility was more stable and thus the model without the volatility adjustment fits well (Exhibit 1).

There is strong evidence that investors contemporaneously set stock market E/Ps (P/Es) as a function of nominal interest rates. All else equal, higher Y implies higher E/P (lower P/E). Over a long period like 1926–2001, however, changing perceptions of stock and bond market risk must be accounted for, or this missing variable obscures the relationship. Accounting for this properly, we see that for at least 75 years, while it may have all been because of the error of money illusion, investors have indeed been following the Fed model.

Conclusion

The very popular Fed model has the appearance but not the reality of common sense. Its lure has captured many a Wall Street strategist and media pundit. However, the common sense is largely misguided, most likely due to a confusion of real and nominal (money illusion). The empirical evidence tells us the Fed model has no power to forecast long-term real stock returns. To the contrary: Traditional methods, like examining the market's unadjusted P/E alone, are very effective.

In its practical recent use, the Fed model offers a toxic combination of comparing an often exaggerated E/P (using forecasted operating earnings) to an irrelevant benchmark (nominal Y). Effectively, the Fed model is a misleading sales tool for stocks. Its popularity is presumably driven by its simplicity; its flexibility (if you don't like the E/P, just call some expenses non-recurring); its superficial rigor (it looks like math); its false initial resemblance to common sense (pundit after pundit enjoys explaining to a presumably impressed audience how bonds really have a P/E too); and most assuredly the fact that it is now, and for some time has been, more bullish than the traditional model.

Now, as opposed to its failure for forecasting long-term stock returns, the Fed model seems to be a success at *describing* how investors actually set current market P/Es. There is strong evidence that investors set stock market E/Ps lower (P/Es higher) when nominal interest rates are lower (and vice versa). This relation is strong and clear over the last 30 to 40 years. Over the 1926–2001 time period, however, it is apparent only when we properly account for a missing variable, perceived stock versus bond risk.

Many market commentators confuse this *descriptive* power of the Fed model for a proof that one should use the Fed model to make investment decisions. These are different issues. It is a strange leap to observe that investors consistently make an error—and then recommend that error, citing precedent.

Endnotes

The author thanks Robert Arnott, Theodore Aronson, Michael Ashton, Gabriel Baracat, Peter Bernstein, Richard Bernstein, William Bernstein, John Brynjolfsson, James Clark, Jonathan Clements, Richard Cohn, Thomas Cowhey, Christopher Darnell, Frank Fabozzi, Craig French, Kenneth French, Mark Hulbert, Antti Ilmanen, Benjamin Inker, Robert Krail, Kevin Lansing, Martin Leibowitz, John Liew, Burton Malkiel, Matthew McLennan, James Montier, Lars Nielsen, Thomas Philips, Jay Ritter, Jeremy Siegel, Lawrence Siegel, and Katherine Welling for comments on this article or an earlier version.

[1]While stories vary, it is often claimed to have been first found in a 1997 Federal Reserve Monetary Policy Report to Congress.

[2]The Fed model is often presented in both the form of a difference (E/P − Y) and a ratio (E/P ÷ Y). I focus on E/P − Y. The logic and the statistical tests in this article differ little if differences are replaced with ratios. Also, ratios obviously get increasingly strange as interest rates fall.

[3]I do not promote the P/E ratio versus other reasonable measures of valuation like the dividend yield or Tobin's Q, rather only the concept of looking at raw versions of valuation (unadjusted for interest rates or inflation) when forecasting long-term real stock returns.

[4]Many use next year's forecasted stock market earnings for the E in the Fed model's E/P. I use long-term trailing earnings because forecasted earnings are available for only a small fraction of the time period studied, and are essentially unusable for tests of whether the Fed model forecasts long-term returns. While forecasted earnings may be a better or a worse measure than trailing earnings (depending on one's faith in Wall Street), it is difficult to imagine this choice of E mattering a great deal in tests of the viability of comparing any E/P to interest rates.

In addition, my E/P and the IBES forecast E/P are highly correlated time series (0.97 since 1976), and any level differences are irrelevant (e.g., forecasted P/Es are generally lower than trailing P/Es both because earnings grow over time and because Wall Street on average is overoptimistic), as level differences end up in the regression intercepts. Additionally, my U.S. time series results are essentially replicated in the cross-section of country returns, this time using forecasted E from IBES.

[5]Actually, as a portion of earnings must be reinvested, you get only the dividend yield plus other distributions, not the earnings yield. This distinction is quite important itself, rendering a comparison of E/P and Y a bit silly.

[6]Recent times have seen PAY values considerably lower than historical averages. The impact of this is unclear. When PAY is low, it is possible that firms are simply retaining earnings for productive use, or to give to shareholders through other means that are equivalent to dividends (e.g., share repurchases). Arnott and Asness [2003] and Bernstein [1997a, 1998] would argue that historically there is a strong tendency for low payouts to lead to lower than normal future earnings growth (low G_E), so the assumption in the text may be optimistic when payouts are low.

[7]One can argue with this assumption of a permanent instantaneous shift in expected inflation, but this argument goes against the Fed model. If one argues that inflation changes are transient and will regress to the mean, then the Fed model is complete gibberish, as a very long-dated asset like the stock market cannot have a radically different fair P/E based on a temporary blip in the CPI.

[8]All t-statistics are adjusted for overlapping observations where appropriate. All R^2 values are adjusted for degrees of freedom.

[9]Asikoglu and Ercan [1992], in a related study, find a 73% flow-through from inflation to nominal earnings for industrial stocks over 1974–1988, with considerable variation by industry. Leibowitz and Kogelman [1993] also discuss this issue in depth.

[10]Boudoukh and Richardson [1993] confirm that over the long term, unlike over the short term, stocks are a good inflation hedge. In fact, this ongoing conundrum—why aren't stocks a good short-term inflation hedge while they are a good long-term inflation hedge?—is in all likelihood related to the issue of the Fed model's predictive versus descriptive efficacy.

[11]See Polk, Thompson, and Vuolteenaho [2003] for another example of the Fed model's explanatory efficacy.

[12]Note that the backward-looking dark bars actually cover an extra decade of returns (1955–1964) versus the forward-looking light bars. The story of Exhibit 2 is robust to shifting either series forward or backward by a few years. However, if we stray far from the 1965–2001 period when E/P and Y track each other so well, Exhibit 2 would change appropriately.

[13]These points are not original. In particular, see Modigliani and Cohn [1979], Ritter and Warr [2002], and Siegel [2002b].

[14]Thanks to Matthew McLennan and Thomas Philips in particular for making this point clear to me.

[15]Ritter and Warr [2002] do argue that the net of the accounting effects is that P/Es should be higher not lower when inflation is high, and thus the Fed model is not simply wrong but backward.

[16]In Asness [2000], I show some short-term forecasting success for a modified Fed model that incorporates the information in the volatilities of stocks and bonds. Even if the Fed model is misspecified and followed in error, if investors make this error with great regularity, and often return to it when they diverge from its norms, some tactical efficacy may be achieved.

[17]See Siegel [2002a], among others, for evidence that not only is E/P a real quantity, but also it is itself a reasonable estimate of the complete expected real return on equities.

[18]Another alternative is to replace Y – I with the yield on long-term TIPS. Note that, when coincidentally RP is approximately equal to I, the Fed model will be a valid relative value tool by accident.

[19]Inker [2002] makes the interesting point that if stocks and bonds are equally overvalued, stocks are the more dangerous asset as they are "longer duration," meaning if both stock and bond expected returns revert to normal, stocks have further to fall.

[20]The relationship is quite robust to other reasonable time periods for measuring volatility.

[21]The coefficient is higher on bond volatility presumably because bond volatility itself varies less through time.

[22]Ignoring certain relatively small convexity issues that arise from inverting an estimate of E/P.

References

Arnott, Robert D., and Clifford S. Asness. "Surprise! Higher Dividends Lead to Higher Earnings Growth." *Financial Analysts Journal*, January/February 2003.

Arnott, Robert D., and Peter L. Bernstein. "What Risk Premium Is 'Normal'?" *Financial Analysts Journal*, March/April 2002.

Asikoglu, Yaman, and Metin R. Ercan. "Inflation Flow-Through and Stock Prices." *The Journal of Portfolio Management,* Spring 1992, pp. 63–68.

Asness, Clifford S. "Stocks vs. Bonds: Explaining the Equity Risk Premium." *Financial Analysts Journal*, March/April 2000.

Bernstein, Peter L. "A Further Look at the Matter of Dividends." Economics & Portfolio Strategy, Peter L. Bernstein, Inc., April 1, 1998.

_____. "Payouts and Payoffs and an Interesting Erratum." Economics & Portfolio Strategy, Peter L. Bernstein, Inc., March 15, 1997a.

_____. "Stock/Bond Risk Perceptions and Expected Returns." Economics & Portfolio Strategy, Peter L. Bernstein, Inc., February 1, 1997b.

Boudoukh, Jacob, and Matthew Richardson. "Stock Returns and Inflation: A Long Horizon Perspective." *American Economic Review*, 83 (1993), pp. 1346–1355.

Fama, Eugene E., and Kenneth R. French. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy*, 96, No. 2 (April 1988), pp. 246–273.

_____. "The Equity Risk Premium." *Journal of Finance,* 57 (April 2002), pp. 637–659.

Inker, Ben. "The Fed Model: Don't Believe the Hype." Working paper, GMO, October 2002.

Leibowitz, Martin L., and Stanley Kogelman. "The Effects of Inflation," Chapter 8 in *Franchise Value and the Price/Earnings Ratio.* Charlottesville, VA: The Research Foundation of the Institute of Chartered Financial Analysts, July 1993, pp. 119–137.

Modigliani, F., and R. Cohn. "Inflation, Rational Valuation, and the Market." *Financial Analysts Journal,* March/April 1979.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho. "New Forecasts of the Equity Premium." Working paper, Kellogg School at Northwestern University and Harvard Business School, July 2003.

Ritter, Jay R., and Richard S. Warr. "The Decline of Inflation and the Bull Market of 1982–1999." *Journal of Financial and Quantitative Analysis,* 37, No. 1 (March 2002), pp. 29–61.

Shiller, Robert J. *Irrational Exuberance.* Princeton: Princeton University Press, 2000.

Siegel, Jeremy J. "The Rise in Stock Valuations and Future Equity Returns." *Journal of Investment Consulting,* Summer 2002a.

_____. *Stocks for the Long Run,* 3rd ed. New York: McGraw-Hill, 2002b.

Style Timing: Value versus Growth

Is value dead?

Clifford S. Asness, Jacques A. Friedman, Robert J. Krail, and John M. Liew

A large body of both academic and industry research supports the efficacy of value strategies for choosing individual stocks. Fama and French [1992, 1993]; Lakonishok, Shleifer, and Vishny [1994]; and Capaul, Rowley, and Sharpe [1993] among others present evidence from both the U.S. and other countries that over the long term, value stocks outperform growth stocks. Yet value strategies are far from riskless. They can produce long periods of poor performance.

In an effort to improve upon value strategies, researchers have tried to forecast these returns, with mixed results. Arnott [1992], Fan [1995], Sorensen and Lazzara [1995], Bernstein [1995], and Kao and Shumaker [1999] investigate models that forecast differences between the returns to value and growth strategies according to measures of aggregate economic and financial conditions. These studies focus on variables like the earnings yield on the S&P 500, the slope of the yield curve, corporate credit spreads, corporate profits, and other macroeconomic measures. Some of these variables appear to have power to forecast value versus growth returns, and others do not.

One criticism of this approach is that it may be susceptible to uncovering spurious ex post relationships. Because all the variables may be economically meaningful, it becomes very difficult to determine which of the observed relations are real and which ones are artifacts of the data.

Reprinted from *The Journal of Portfolio Management*, Spring 2000, Vol. 26 (3), with permission.

We propose a different approach considering two simple and intuitive variables: 1) the spread in valuation multiples between a value portfolio and a growth portfolio (the value spread); and 2) the spread in expected earnings growth between a growth portfolio and a value portfolio (the earnings growth spread).[1]

The motivation for these variables follows from a version of the Gordon [1962] model. This model states that:

$$E(R) = E/P + g \tag{1}$$

where E(R) represents the expected return of a given stock, E/P represents the stock's earnings yield, and g represents the expected earnings growth in perpetuity.[2]

The Gordon model is a simplistic decomposition of the expected return of a stock, and it relies on some strong assumptions, but it is a useful heuristic nevertheless. According to this model, expected returns can be decomposed into two factors: 1) a simple valuation ratio, and 2) a forecast of future earnings growth. This decomposition motivates our work.

We can rewrite Equation (1) for both value stocks and growth stocks as follows:

$$E(R_{value}) = E/P_{value} + g_{value} \tag{2}$$

$$E(R_{growth}) = E/P_{growth} + g_{growth} \tag{3}$$

Taking the difference between these two equations, we arrive at a simple style timing model:

$$\begin{aligned} &E(R_{value} - R_{growth}) = \\ &(E/P_{value} - E/P_{growth}) - (g_{growth} - g_{value}) \end{aligned} \tag{4}$$

The first term $(E/P_{value} - E/P_{growth})$ represents the value spread. Since value stocks are often defined as stocks with high E/P (or other similar valuation ratio), and growth stocks are often defined as stocks with low E/P, the value spread should by construction be positive. This first style timing variable is simply motivated by the observation that, all else equal, when the difference between E/P for value stocks and E/P for growth stocks is abnormally great, the difference between the expected return to value stocks and growth stocks should be abnormally great. In other words, when the value spread is especially wide, cheap stocks are *really* cheap, and expensive stocks are *really* expensive.

The second term in Equation (4) is the growth spread, $(g_{growth} - g_{value})$. Note that in reversing the notation for the value spread, we express the growth spread as the expected earnings growth for growth stocks minus the expected earnings growth for value stocks. Since growth stocks tend to be strong earners, and value stocks tend to be relatively distressed, $(g_{growth} - g_{value})$ should be positive.

Equation (4) shows that both the value spread and the growth spread are important determinants of the expected return difference between value and growth. A high E/P stock does not necessarily have a high expected return if it also has low expected earnings growth. In other words, a high E/P can be justified if a stock's expected earnings growth is poor. Thus, a wider-than-normal value spread does not necessarily signal that the expected return to value is higher than growth. If the earnings growth spread is sufficiently wide, it

can justify a wider-than-normal value spread and hence no greater-than-normal expected return premium.

We find that both value spreads and earnings growth spreads are important indicators of the attractiveness of value versus growth. Using data from January 1982 through October 1999, we find that the combination of the value spread and the earnings growth spread forecasts the future returns of value versus growth. This relation is both statistically and economically significant. Moreover, while we motivate the analysis using a version of E/P, we find that the results are general to other common valuation ratios. The three measures of value that we focus on are: earnings-to-price, book-to-price, and sales-to-price.

At the time of this writing (November 1999), our model is forecasting near-historic highs for the expected return of value versus growth. Value spreads are near historic highs, and the level of these spreads cannot be explained by high expected earnings growth spreads. In fact, expected earnings growth spreads are actually relatively low. In other words, value stocks currently appear to be far cheaper than growth stocks compared to historical norms, but rather than giving up more expected earnings growth than normal, value stocks are actually giving up less expected earnings growth.

Data and Methodology

Our first goal is to form a simple but robust proxy for value. We believe that a composite of three accounting ratios that incorporates earnings, book value, and sales along with price captures the main characteristic of value, but is not overly sensitive to any one accounting item. We focus on the measures of value as follows: earnings-to-price (E/P), book-to-price (B/P), and sales-to-price (S/P). Each month-end, from December 1981 through September 1999, we form each of these three value indicators for each stock in our investable universe as follows:[3]

E/P_{it} = IBES forecasted next twelve months' earnings per share at time t divided by IBES stock price at time t.

B/P_{it} = Compustat annual balance sheet common equity divided by market value of equity at time t.

S/P_{it} = Compustat annual sales divided by the sum of market value of equity at time t plus Compustat annual balance sheet book value of long-term debt minus Compustat annual balance sheet cash and cash-equivalents. Note that we use a proxy for the value of the entire firm rather than the value of the firm's equity, because sales apply to the entire capital structure of the firm as opposed to book value and earnings, which apply only to equityholders.

The investable universe for each month is defined as the top 1,100 most liquid stocks, as measured by the trailing quarter's total dollar trading volume, that also rank among the top 1,500 stocks by market capitalization.[4] We believe this restricted universe makes our analysis relevant to institutional money managers.

Asness, Porter, and Stevens [1999] find that trading strategies based on industry-adjusted value factors significantly outperform those based on non-industry-adjusted valuation factors. In other words, value is a better strategy for choosing stocks within an industry than for choosing industries. Following the Asness, Porter, and Stevens methodology, we form industry-adjusted versions of each of the three value measures. This approach

compares each stock's accounting ratio to its industry average rather than to the entire universe. This differs from the approach reflected in traditional value and growth indexes, such as the S&P/BARRA and Frank Russell indexes, which use non-industry-adjusted measures of value and growth and hence entail significant industry biases.[5]

For each of the three raw (i.e., non-industry-adjusted) value measures, we form an industry-adjusted version by subtracting the industry average as follows:

$$\text{Industry} - \text{Adjusted } E/P_{it} = E/P_{it} - \overline{E/P_{it}} \tag{5}$$

$$\text{Industry} - \text{Adjusted } B/P_{it} = B/P_{it} - \overline{B/P_{it}} \tag{6}$$

$$\text{Industry} - \text{Adjusted } S/P_{it} = S/P_{it} - \overline{S/P_{it}} \tag{7}$$

where $\overline{E/P_{it}}$, $\overline{B/P_{it}}$, and $\overline{S/P_{it}}$ are stock i's industry average valuation ratio at time t.[6]

We believe that each of these three measures represents a noisy proxy for value. To obtain a robust aggregate industry-adjusted value measure, we form a composite of industry-adjusted E/P, B/P, and S/P. To combine the three value indicators into one overall measure of value, every month we first rank each stock in the universe on each variable. Then, for each stock we compute its average rank across the three industry-adjusted value measures as follows:

$$\text{Value Composite} = \text{Average}\left[\text{Rank}\left(E/P\right), \text{Rank}\left(B/P\right), \text{Rank}\left(S/P\right)\right]$$

We then rerank our universe of stocks on this average rank measure. Thus, our overall measure of value for each stock is a rank based on a one-third weight for each of our three industry-adjusted value indicators.

The second goal is to construct a measure of expected earnings growth. We use analysts' long-term earnings growth estimates to form a proxy for expected growth for each stock i as follows:

$$\text{Growth}_{it} = \text{IBES Median Long} - \text{Term EPS Growth Forecast at Time t}$$

According to Equation (1), we are interested in a proxy for current expected earnings growth in perpetuity. There are some potential problems with our proxy. First, analysts may not update their long-term forecasts in one leap. Rather, they may move slowly (psychologists call this anchoring and adjustment). Thus, the current IBES data may lag actual estimates of long-term earnings growth.

Second, IBES uses a five-year horizon to define long-term growth. Some of the expected earnings growth priced into stocks (rationally or irrationally) may be forecasted to occur beyond this five-year horizon.

Third, to the extent analysts overextrapolate past earnings growth in their forecasts of future earnings growth, the IBES forecasts may be biased. Fourth, this expected earnings growth measure is available only from 1982–1999 and serves as the limit in testing our model farther back in time.

To the extent all these problems are material, they could weaken the forecasting power of this variable or the statistical power of the tests.

Exhibit 1 shows that over the last eighteen years the expected real earnings growth of the S&P 500 (market capitalization-weighted $growth_{it}$ minus one-year trailing inflation) has been steadily increasing.[7] Analysts are currently exceptionally bullish about future earnings growth for the overall market.

With a composite measure of industry-adjusted valuation and a measure of expected earnings growth for each stock, we are armed to proceed with the style timing analysis.

Performance of Value Strategies

Exhibit 2 presents the performance for each of the three industry-adjusted value measures (E/P, B/P, and S/P), as well as the composite version. We first form four sets of decile portfolios by sorting all stocks in our universe separately on each of the three individual intra-industry value measures and on the composite value measure. We then form zero-investment portfolios by subtracting the equal-weight returns on decile 1 (growth stocks) from the equal-weight returns on decile 10 (value stocks). The portfolios are rebalanced quarterly.[8] In addition, the zero-investment portfolio returns are conditionally beta-adjusted each month by subtracting the product of the net beta (calculated using BARRA's betas) of the portfolio and the monthly excess return on the S&P 500.

As many other authors have found, value outperformed growth over this period. From 1982 to 1999, a zero-investment portfolio based on our composite indicator that is long decile 10 (value stocks) and short decile 1 (growth stocks) produces excess returns of 6.25% per year. This equates to an annual Sharpe ratio of 0.58, which is statistically significant with a t-statistic of 2.46.

Exhibit 3 presents the rolling twelve-month excess returns of our composite value strategy. While value stocks on average beat growth stocks, there are sustained periods of

Exhibit 1 Expected Real Long-Term Earnings Growth — S&P 500 Index, January 1982–October 1999

Exhibit 2 Value Strategy Performance Decile 10 (Value) Minus Decile 1 (Growth) January 1982–October 1999

	E/P	B/P	S/P	Comp.
Average (annualized) (%)	2.74	3.38	6.79	6.25
Std. Dev. (annualized) (%)	9.83	9.52	8.93	10.75
Annual Sharpe Ratio	0.28	0.36	0.76	0.58
t-Statistic	1.18	1.50	3.21	2.46
Beta with S&P 500	0.00	–0.03	–0.02	–0.03
Best 12 Months (%)	34.7	25.5	34.6	46.1
Worst 12 Months (%)	–36.1	–26.8	–15.8	–28.1

Exhibit 3 Rolling Twelve-Month Performance of the Value Strategy, Decile 10 (Value) Minus Decile 1 (Growth), January 1982–October 1999

poor performance. Since 1982 there have been three major bear markets for value: 1989–1990, 1995–1996, and, most recently, 1998–1999.

Forecasting Value versus Growth

Given that value strategies can produce prolonged droughts, any indication of when these tough periods will occur would be, to say the least, rather useful. We propose two variables to predict the returns of the value strategy. The first is the composite value spread, which is calculated as follows:

1. Again, form decile portfolios each month by sorting on the composite industry-adjusted value measure from low to high. We define the value stock portfolio as

decile 10 (highest composite value) and the growth stock portfolio as decile 1 (lowest composite value).

2. Compute the median of each of the three raw accounting ratios (E/P, B/P, and S/P) for both decile 10 (value stocks) and decile 1 (growth stocks), and then compute the ratio as follows:

$$\text{E / P Spread}_t = \frac{\left(\text{Value Portfolio Median E / } P_{it}\right)}{\left(\text{Growth Portfolio Median E / } P_{it}\right)} \tag{9}$$

$$\text{B / P Spread}_t = \frac{\left(\text{Value Portfolio Median B / } P_{it}\right)}{\left(\text{Growth Portfolio Median B / } P_{it}\right)} \tag{10}$$

$$\text{S / P Spread}_t = \frac{\left(\text{Value Portfolio Median S / } P_{it}\right)}{\left(\text{Growth Portfolio Median S / } P_{it}\right)} \tag{11}$$

E/P spread_t, for example, represents the multiple that growth stocks are selling for versus value stocks at time t. Although E/P spread_t should always be greater than 1.0, a wider-than-average E/P spread_t might represent a time when value stocks are abnormally cheap relative to growth stocks.[9]

3. While we present evidence for each of the three spreads individually, we focus on a composite valuation spread that equally weights each of the three spreads. As before, we believe this composite value spread represents a more robust measure of the current relative pricing of value versus growth stocks than a spread based on any single valuation measure.

To form the composite value spread, we first standardize separately each of the three value spread variables formed in Equations (9)–(11) so that their time series average is zero and standard deviation is 1.0.

We next average each of the three standardized value spread variables, and restandardize to form the composite value spread measure. When the composite value spread is zero, value stocks are cheaper than growth stocks by their historical average amount. A positive composite value spread indicates value stocks are cheaper than normal (value spread is wider than normal), and a negative composite value spread indicates value stocks are not as cheap as normal (value spread is narrower than normal).

Note that in computing the spread we use ratios rather than differences. This removes the influence of the overall market price. Following a period of strong market performance, valuation ratios (E/P, B/P, and S/P) for all stocks generally decrease. Hence, spreads based on differences will be compressed. Ratios make movements in the valuation spread less sensitive to overall market moves and more representative of relative performance between value versus growth stocks.

Exhibit 4 presents summary statistics for the time series of the spreads of E/P, B/P, and S/P between decile 10 (value stocks) and decile 1 (growth stocks). For example, our median value stock's E/P is on average 2.0× our median growth stock's E/P. To put it another way, on average our growth portfolio sells for double the P/E of our value portfolio.

Exhibit 4 Summary Statistics for Value Spreads Value (Decile 10) ÷ Growth (Decile 1) For E/P, B/P, And S/P, January 1982–October 1999

	E/P Spread	B/P Spread	S/P Spread
Average	2.0×	4.1×	5.4×
Std. Dev.	0.3×	0.8×	1.1×
Maximum	3.0×	6.4×	8.5×
Minimum	1.5×	2.7×	3.4×
Current (11/99)	2.9×	5.8×	5.6×

Similarly, our median value stock's B/P is on average 4.1× our median growth stock's B/P, and our median value stock's S/P is on average 5.4× our median growth stock's S/P.

Note, as one would expect, the minimum spread demonstrates that on each measure the median value stock is always cheaper than the median growth stock. The degree to which the median value stock is more attractively priced varies substantially through time, and it is this variation that we seek to capitalize on to forecast the conditional attractiveness of value versus growth.

Exhibit 5 graphs the E/P, B/P, and S/P spreads through time. While each spread exhibits some degree of idiosyncratic behavior, they all exhibit common trends. Currently, all the spreads are above their long-term averages, and the E/P and B/P spreads are near historic highs.

The second variable we propose to predict the returns of value versus growth is the earnings growth spread. The earnings growth spread is calculated by computing the difference between the expected earnings growth for the median stock in the growth portfolio and the expected earnings growth for the median stock in the value portfolio. Since the expected earnings growth spread is unaffected by the level of market prices, and since it is plausible that the expected earnings growth on the value portfolio is occasionally near zero, we use the difference instead of the ratio.

On average, our median growth stock's expected long-term earnings growth is 8.4% higher than the median value stock's expected earnings growth. As we discuss above, this is not surprising, as value stocks tend to be distressed companies, while growth stocks tend to be market darlings.

Exhibit 6 shows that, as expected, the median value stock's expected earnings growth never exceeds the median growth stock's expected earnings growth. Yet the spread varies significantly over time with a maximum of 15.5% and a minimum of 3.9%.

Notice the sharp drop in the earnings growth spread in October 1998. Following the market turmoil from July through September (the portfolios are rebalanced each calendar quarter-end), there was very large turnover in decile 10 and decile 1. In other words, there was a large change in the composition of value and growth stocks as defined by our aggregate industry-adjusted value measure.[10] Furthermore, the sharp decline in the earnings growth spread is mainly due to this significant portfolio turnover rather than changes in analyst growth estimates.

The result is that the median consensus estimate of long-term earnings growth for growth stocks drops by 4% while that for value stocks actually increases by 1.5%.

Exhibit 5 Value (Decile 10) ÷ Growth (Decile 1), January 1982–October 1999

Exhibit 6 Expected Long-Term Earnings Growth Spread — Growth (Decile 1) Minus Value (Decile 10), January 1982–October 1999

Exhibit 7 graphs the standardized composite value spread and standardized earnings growth spread (i.e., both series are standardized to have a mean of zero and a standard deviation of one through time). Recall that our composite value spread is a standardized average of the value spreads for each of E/P, B/P, and S/P. The correlation between the two spreads is 0.62. When value spreads are wider than normal (i.e., value stocks look abnormally cheap relative to growth stocks), growth spreads tend to be wider than normal (i.e., growth stocks have abnormally high expected growth relative to value stocks).

An important implication of this strong positive correlation is that the value spread alone is not a sufficient indicator of the attractiveness of value strategies. In general, when value stocks are priced more cheaply than average compared to growth stocks, they are also giving up more expected earnings growth than normal. Going back to the Gordon model, if wide value spreads are driven only by big differences between the expected earnings growth of growth stocks versus value stocks, then there could be no abnormal expected return advantage to value versus growth.

Exhibit 8 presents predictive regressions of the next twelve-month rolling return of the composite value strategy (return on value stocks minus return on growth stocks from Exhibit 3) on each of the three individual value spreads (E/P, B/P, and S/P) and on the composite value spread, each with and without the standardized earnings growth spread. The fact that each explanatory variable is standardized makes the regression coefficients directly comparable. Each row represents a separate linear regression. We present the R-squares and coefficients (with the t-statistics in parentheses).[11]

Regressions (1)–(4) show that all four of the value spreads alone do a pretty good job of predicting future returns to the value strategy. Regardless of the earnings growth spread, when value spreads are wider than normal, the expected return to value stocks versus growth stock is generally higher than normal. In particular, the composite value spread

Exhibit 7 Standardized Composite Value Spread and Expected Earnings Growth Spread, January 1982–October 1999

Exhibit 8 Predictive Regressions of Annual Value Strategy Returns on Value Spreads and Earnings Growth Spreads, January 1982–October 1999

	Constant	Value Spread	Earnings Growth Spread	Adjusted R-Square
1. B/P Only	7.12	7.57		
	(2.19)	(2.72)		24.6%
2. E/P Only	7.23	7.36		
	(2.18)	(2.28)		19.5%
3. S/P Only	6.98	6.44		
	(2.09)	(2.04)		17.4%
4. Composite Value Only	7.18	7.58		
	(2.17)	(2.47)		22.8%
5. Earnings Growth Only	6.45		1.84	
	(1.80)		(0.45)	1.1%
6. B/P with Earnings Growth	7.92	13.43	–7.57	
	(3.03)	(4.17)	(–2.67)	35.4%
7. E/P with Earnings Growth	7.87	11.57	–5.37	
	(2.79)	(5.77)	(–2.37)	26.0%
8. S/P with Earnings Growth	7.69	12.07	–7.05	
	(2.79)	(4.24)	(–3.01)	26.6%
9. Composite Value with Earnings Growth	8.29	15.52	–9.48	
	(3.26)	(5.13)	(–4.49)	38.7%

alone is significantly positively related to future value versus growth returns (t-statistic of 2.47), and explains almost 23% of the future annual return variation.

Regression (5) shows that the earnings growth spread alone predicts very little of next year's value returns. The coefficient on the earnings growth spread is actually positive. That is, when the earnings growth spread is wider, the expected return for value is higher. This result appears to contradict the prediction in Equation (4), which suggests a negative relation. Exhibit 7, however, shows that the earnings growth spread is strongly positively related to the value spread. When the earnings growth spread is wide, this tends to be when the value spread is wide. Thus, the positive relation between the earnings growth spread and next year's value versus growth return is spurious and due to the strong positive relation between the value spread and next year's returns to value versus growth.

Regressions (6)-(9) show that when the earnings growth spread is combined with any of the value spreads, there is a great improvement in the ability to forecast next year's return to value compared to using the value spread alone. In particular, the composite value spread combined with the earnings growth spread explains about 39% of the variability in next year's value versus growth returns. The coefficient on the earnings growth spread is now strongly negative as expected, and the coefficients on the value spreads are higher than those in regressions (1)-(4). Thus, when the value spread is wide (i.e., value stocks look abnormally cheap compared to growth stocks), and the earnings growth spread is small (i.e., the expected earnings growth advantage for growth stocks versus value stocks is smaller than normal), this is a very good time for value.

How good is a 38.7% adjusted R-square model? Exhibit 9 plots the fitted values from regression (9) against the actual twelve-month future return to value versus growth from January 1982 to November 1999.[12] The graph of the actual next twelve-month returns ends in November 1998 (which represents the twelve-month value return from November 1998 through October 1999).

Exhibit 9 Actual and Predicted Twelve-Month Returns of Value versus Growth, January 1982–November 1999

The model is clearly not perfect, and the forecast errors can be great (as they have been for the last year), but the model does correctly forecast most of the major moves.

Current Forecasts

If we have built a good model, it is of course particularly interesting to observe its current forecast. At the time of this writing (end of October 1999), the model forecasts a 52% return spread (3.6 standard deviations above the historical average) between value and growth over the coming year. Clearly, this is near historic highs for value versus growth.

Historically, the median growth stock on average sells for 2.0×, 4.1×, and 5.4× the median value stock respectively on P/E, P/B, and P/S. Yet the current situation is far from average. The median growth stock now sells for 2.9×, 5.8×, and 5.6× the median value stock. This leads to a composite value spread that is quite high by historical standards (2.1 standard deviations above the average).

Typically, wide value spreads are accompanied by wide growth spreads (recall the two variables have a 0.62 historical correlation). This is not the case now. Historically, the median growth stock is expected to outgrow the median value stock by an average of 8.4% per year. Currently, it is expected to outgrow by only 5.0% per year (1.2 standard deviations below the average). While growth stocks still have higher expected earnings growth than value stocks, the differential is now significantly less than average. Thus, according to our simple model, we have the best of all worlds for value stocks going forward.

We should point out again that the version of value defined here is not necessarily the same as some popular conceptions of value (e.g., S&P/BARRA and Frank Russell). For instance, the beta-adjusted correlation between the returns on our long/short value minus growth portfolio and our proxy for the S&P/BARRA value minus growth portfolio is 0.48.[13] The performance difference is also significant, considering that the beta-adjusted S&P/BARRA value minus growth portfolio has realized only a 0.08 Sharpe ratio over the January 1982 through October 1999 period, while our value strategy realizes a 0.58 Sharpe ratio. Clearly they are related, but they do not capture the exact same phenomenon.

As we say earlier, we feel that using a composite of industry-adjusted valuation indicators creates a far more robust, higher Sharpe ratio strategy. Obviously our model and our currently very optimistic forecast apply most strongly to a value manager who uses a similar approach (i.e., seeking value without strong industry biases and using a diversified set of indicators).

Conclusion

Expected return premiums can vary through time as a consequence of rational or irrational forces. Rational forces can be either variation through time in the risk of value stocks versus growth stocks, or variation through time in the amount people must be paid to bear this risk. Certainly, the tough recent performance of value strategies has squeezed out many weak hands, and only the strongest advocates may be left. Anecdotally, it would not be surprising if at times like this investors would require greater compensation for bearing value versus growth risk.

Irrational forces could be simply a mispricing between value and growth stocks and time variation in the relative degree of mispricing. Our model's current positive forecast for value could simply reflect an irrational mania — growth at a reasonable price has become growth at any price. Some have conjectured that value strategies have been harmed by a disconnect in the LBO process in that low-grade bond yields are currently very high relative to Treasuries, making it difficult to "unlock" the value in undervalued companies. While this might be plausible, it begs the question of why value was such an effective strategy for years before the low-grade bond market became important.

The recent performance of value strategies, and other historical bear markets for value, clearly shows that value can lose to growth for prolonged periods of time. "The world has changed!" is a common cry heard from those skeptical of value strategies, especially after these rough periods. Today, new technology, globalization, and newly established franchise values, among many other factors, may allow some companies to grow earnings far in excess of and longer than what the market has seen in the past.

We propose a simple model that explicitly seeks to capture this sentiment through Wall Street analysts' forecasts of long-term earnings. We find evidence that suggests that this model can forecast the returns to value versus growth. Moreover, its forecasting power is strong from both a statistical and an economic perspective; it currently forecasts near-historic highs in the expected one-year return of value stocks versus growth stocks.

According to this model, value is very far from dead.

Endnotes

The authors thank Richard Bernstein, Adam Blitz, John Bogle, Thomas Dunn, Kenneth French, Ronald Gutfleish, Takehiro Hamada, Brian Hurst, Antti Ilmanen, Robert Jones, David Kabiller, Oktay Kurbanov, Josef Lakonishok, Michael Patchen, Thomas Philips, James Poterba, Narayan Ramachandran, Jeremy Siegel, Rex Sinquefield, Andre Stern, Todd Tibbetts, Ingrid Tierens, and Geoffrey Verter for valuable comments and suggestions. The opinions expressed are the authors', and should not be taken to represent those of their company.

[1]Fan [1995] proposes the P/E spread as a predictor of value versus growth in tests based on the S&P/BARRA value and growth indexes. He finds only a weak predictive relation between the P/E spread and future returns. Our work continues along this line, but also incorporates the earnings growth spread and uses spreads in valuation measures based on ratios rather than straight differences to remove the influence of the overall market's valuation. Additionally, we focus on value strategies that are industry-adjusted and diversified across several indicators. Asness, Porter, and Stevens [1999] show that value strategies are significantly more effective on an industry-adjusted basis. The S&P/BARRA indexes are constructed using raw book-to-price and thus incorporate significant industry biases. Our results suggest that industry adjustment, using ratios of valuation indicators rather than differences, and incorporation of the earnings growth spread produce significantly stronger predictive relations.

[2]The Gordon model refers to earnings (or free cash flow) that are not reinvested to generate future earnings (i.e., dividends or potential dividends). We employ a more general version of valuation that includes earnings-to-price, book-to-price, and sales-to-price. The results are not sensitive to differences in the specification of value.

[3]Note that all Compustat accounting data are lagged a minimum of six months to minimize the possible effect of look-ahead bias on our value measures. Lagging the data by six months actually generates between six- and eighteen-month lags, depending on the time of year and the month of the fiscal year-end. For example, the value indicators for a stock with a December fiscal year-end will

use the December fiscal year-end data (time t) for the twelve months July (time t + 6) through June (time t + 18). Thus, the accounting data lag varies from six to eighteen months.

[4]Additionally, for each stock we require a return to exist in the subsequent month; we exclude ADRs and Internet stocks; and we include only U.S. domiciled stocks (including financials). Internet stocks are excluded for reasons of severe data limitations (such as negative earnings and book value) and to avoid potential biases in the current forecast, as Internet stocks are a very new phenomenon.

If we restrict the universe to the S&P 500, the style timing results, while somewhat weaker (as would be expected due to the much smaller cross-section of stocks), are still statistically and economically significant.

[5]Our industry classification is based on the historical BARRA USE3 risk model, which provides industry classifications for each stock in five different industries. We assign each stock in our universe to its most important BARRA industry.

[6]The industry averages are calculated using a 50/50 blend of the equal-weighted and market capitalization-weighted average. This combination reduces the problem that a few very large stocks can overly influence the industry averages when using market capitalization weights, and mitigates the problem of overweighting small stocks when using equal weights. In addition, to limit the weight of outliers, the lowest and highest 0.5% of the observations of each raw value variable are set to the next lowest or highest value prior to calculating the industry averages. The results are not sensitive to these choices.

[7]An equal-weight index also shows a very pronounced, but smaller, recent increase.

[8]If, instead of equal weights within decile 1 and 10, we weight by market capitalization, the style timing results are somewhat weaker, but are all still statistically and economically significant.

Note that quarterly rebalancing mitigates the effects of short-term return reversal that can upwardly bias the value strategy returns. Since the price at time t is used in the denominator for each accounting ratio, recent bad performance will bias a stock's valuation measure upward, and to the extent that returns reverse, will bias the value strategy returns upward. Some of the observed short-term return reversal can be attributed to bid-ask effects. Thus, the achievable performance of value strategies can be overstated if one employs a monthly rebalancing strategy. We find that monthly rebalancing produces annual Sharpe ratios that can be as much as 50% higher than quarterly rebalancing.

[9]Technically, the E/P ratio can be less than one since we form our value portfolio using a combination of valuation factors. Practically, because of the high correlation among the three value variables, this never occurs.

[10]The sharp drop is not specific to our choice of using decile sorts to create our value and growth portfolios. If we use quintiles instead of deciles, we see a similar decline.

[11]Since the regressions employ overlapping annual returns, the residuals will be serially correlated. Hence, the t-statistics are adjusted for serial autocorrelation of a general MA (11) form.

[12]The forecast for November 1999 is based on the regression coefficients and updated composite value and earnings growth spreads.

[13]We form proxies for the S&P/BARRA value and growth indexes by employing a very similar methodology to that used for the actual index construction. Over the overlapping period (October 1993 to October 1999), our versions of S&P/BARRA value and growth are 0.992 and 0.990 correlated to the actual indexes, respectively.

References

Arnott, R. "Style Management: The Missing Element in Equity Portfolios." *Journal of Investing,* Summer 1992.

Asness, C., B. Porter, and R. Stevens. "Intra-Industry and Inter-Industry Variation in the Cross-Section of Expected Stock Returns." Working paper, AQR Capital Management, 1999.

Bernstein, R. *Style Investing: Unique Insight into Equity Management*. New York: John Wiley & Sons, 1995.

Capaul, C., I. Rowley, and W. Sharpe. "International Value and Growth Stock Returns." *Financial Analysts Journal,* 50 (1993), pp. 27–36.

Fama, E., and K. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), pp. 3–56.

_____. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), pp. 427–465.

Fan, S. "Equity Style Timing and Allocation." In R. Klein and J. Lederman, eds., *Equity Style Management*. Chicago: Irwin, 1995.

Gordon, M. *The Investment, Financing, and Valuation of the Corporation*. Homewood, IL: Irwin Publishing, 1962, Ch. 4.

Kao, D., and R. Shumaker. "Equity Style Timing." *Financial Analysts Journal*, 55 (1999), pp. 37–48.

Lakonishok, J., A. Shleifer, and R. Vishny. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance*, 49 (1994), pp. 1541–1578.

Sorensen, E., and C. Lazzara. "Equity Style Management: The Case of Growth and Value." In R. Klein and J. Lederman, eds., *Equity Style Management*. Chicago: Irwin, 1995.

PART IV

Demystifying Hedge Fund Strategies

Do Hedge Funds Hedge?

Characteristics of Risk and Return in Risk Arbitrage

An Alternative Future: Part I

This page intentionally left blank

Do Hedge Funds Hedge?

Be cautious in analyzing monthly returns

Clifford Asness, Robert Krail, and John Liew

In a poor year for traditional investments, 2000 was a good year for hedge funds in aggregate and a great year for some particular hedge fund styles. The year 2000 represents a textbook example of the diversification that investors desire from hedge funds. Unlike traditional investments, hedge funds can take long and short positions, and therefore have the ability to isolate a manager's security selection or timing skill from the asset class in which the manager trades. In this way, hedge funds can offer an investment not only with potentially attractive returns, but also with low to zero correlation with most traditional portfolios.[1]

In recent years, hedge funds seem to have delivered. Over 1994–2000 the CSFB/Tremont index of hedge funds produced compound annual net returns of 13.2% with 10.0% annualized monthly volatility. Agarwal and Naik [2000b], Liang [2001], and Peskin et al. [2000], among others, examine monthly hedge fund returns and find only moderate correlation with most traditional asset class indexes. After adjusting for equity market exposure and other sources of systematic risk, these studies find that hedge funds still produce significant excess returns or alphas.

We argue that these results based on monthly data may be misleading. Many hedge funds hold, to various degrees and combinations, illiquid exchange-traded securities or difficult-to-price over-the-counter securities, which can lead to non-synchronous price reactions. Illiquid exchange-traded securities often do not trade at, or even near, the end of every month (even small- and medium-capitalization stocks may be subject to thin trading). Moreover, publicly available traded prices often do not exist for hard-to-price over-the-counter securities. The lack of prices may leave hedge funds with "flexibility" in how they mark these positions for month-end reporting. A cynic might argue that hedge funds can use this flexibility to manage their reported monthly returns, a practice that Weisman [2000] wryly calls "marketing supportive accounting."

The presence of stale prices due to either illiquidity or managed pricing can artificially reduce estimates of volatility and correlation with traditional indexes. This type of non-synchronous price reaction has been the subject of research for empirically estimating betas of small-capitalization stocks and other illiquid securities. Dimson [1979]

Reprinted from *The Journal of Portfolio Management*, Fall 2001, Vol. 28 (1), with permission.
Recipient of the 2003 Bernstein Fabozzi/Jacobs Levy Best Article Award.

and Scholes and Williams [1977] propose simple techniques to measure market betas by using summed betas from regressions of returns on both contemporaneous and lagged market returns.

We apply these and other techniques to hedge fund returns, and find that simple monthly beta and correlation estimates greatly understate hedge fund equity market exposure. Similarly, simple estimates of volatility using monthly returns seem to understate actual hedge fund volatility. Furthermore, when we account for a more accurate level of market exposure, we find that the broad index of hedge funds and most hedge fund subcategories do not add value over this period compared to what would be expected, given their average market exposure. In other words, according to our tests, the positive aggregate hedge fund returns over this period might be due to market exposure rather than to alpha or manager skill.

Because we focus only on aggregate hedge fund indexes, these results may or may not apply to any individual hedge fund. A carefully chosen portfolio of hedge funds may still provide both the return and the diversification benefits investors seek. In light of our results, however, we argue that at a minimum investors should be cautious analyzing monthly returns and, where possible, should use the techniques we discuss.

We first discuss the CSFB/Tremont hedge fund indexes, and then examine the risk and return characteristics of hedge funds using monthly returns. We propose some methods to account for the problem of non-synchronous pricing in monthly hedge fund returns, and finally present new results on the return and diversification benefits of hedge funds using these methods.

Data

In order to examine the return and risk characteristics of hedge funds, we use returns for the CSFB/Tremont hedge fund indexes from January 1994 through September 2000. These indexes consist of an aggregate hedge fund index designed to represent the industry as a whole and nine subindexes designed to track the primary hedge fund investment styles. CSFB/Tremont constructs all ten indexes monthly by asset-weighting net-of-fee returns for selected funds.[2]

To determine index constituents, CSFB/Tremont begins with funds in the TASS+ database that: have at least $10 million in assets, have provided audited financial statements, and have met CSFB/Tremont's reporting requirements. The indexes exclude funds of funds, but include funds closed to new investment. This set of funds constitutes the CSFB/Tremont universe (as of November 15, 2000, there were 656 funds in the universe).

Each fund in the universe is then assigned to one of nine subcategories according to the fund's investment style. From this universe, CSFB/Tremont selects a subset of funds for inclusion in the subindexes so that each subindex represents at least 85% of the assets under management in that subcategory.

At least three potential biases may impact studies that use hedge fund indexes to proxy for the unobservable market portfolio of all hedge funds:

1. Survivorship bias: Survivorship bias occurs when indexes exclude all or part of the returns for dissolved or defunct funds from the index calculation. Since defunct funds typically have had very poor returns, excluding them from the index calculation will produce

an unrealistically high estimate of a truly investible hedge fund portfolio. While CSFB/Tremont includes some defunct funds in the index calculations, it does not include every fund that ceased operations over the period. Moreover, TASS (and Hedge Fund Research) began collecting data on dead funds only in 1994, so hedge fund data prior to 1994 will entail significant survivorship bias and would not be suitable for accurate estimation of hedge fund risk and return.[3]

2. Backfill bias: Backfill bias occurs if database vendors backfill returns when a new fund is added instead of including its returns only on a going-forward basis. This will overstate index performance, since inclusion in the index is voluntary, and thus funds will generally be added only after very good past performance. CSFB/Tremont includes funds on a going-forward basis only, and therefore avoids any backfill bias.
3. Self-selection bias: Self-selection bias may occur if top- or bottom-performing funds lack the same incentive as other funds to report to data vendors, and thus are excluded from index calculations. This bias is generally slight, since most funds report, and even if some funds are excluded they must exhibit very strong persistence in their performance for any significant downward or upward bias to occur in index performance.

Although CSFB/Tremont attempts to minimize these biases, they are notoriously difficult to eliminate fully, even with the best intentions. Therefore, performance numbers should be evaluated in the context of these biases. In all likelihood, the estimates of average returns based on these indexes will be biased upward, primarily because of survivorship bias. Additionally, since in our particular sample period the stock market was exceptionally strong, survivorship bias in the indexes could bias our estimates of hedge fund market exposure upward since funds with high betas were more likely to survive.

Initial Results

Hedge Fund Summary Statistics

Exhibit 1 presents summary statistics for monthly returns on the aggregate hedge fund index, the nine hedge fund subindexes, and the S&P 500 from January 1994 through September 2000. The aggregate index of hedge funds produces solid results over this period. Hedge funds posted 8.0% average annual excess returns over cash (13.2% compound total returns) net of fees for the period, with 10.0% annualized monthly volatility, resulting in a 0.80 Sharpe ratio. This period coincides with a strong bull market, however, and some hedge funds appear to swim with the tide.

Overall, the aggregate hedge fund index has a monthly correlation of 0.52 with the S&P 500. Two of the most correlated styles are event-driven and long/short equity, which produce Sharpe ratios of 1.05 and 0.94, respectively. Not every style seems to benefit from market exposure. Convertible arbitrage funds produce impressive returns and a Sharpe ratio of 1.07, with a very low 0.13 monthly correlation with the S&P 500.

As expected, dedicated short biased funds produce returns that are highly negatively correlated with the market. Not surprisingly, given the strength of the market over this period, this style posts very poor absolute returns. Its best month is August 1998, the worst month for the market and for many of the other hedge fund styles. Managed futures, which

Exhibit 1 Hedge Fund Returns—Monthly Data January 1994–September 2000

Portfolio	Annualized Excess Return	Annualized Standard Deviation	Annualized Sharpe Ratio	Correlation With S&P 500	Maximum Monthly Return	Month of Maximum Return	Minimum Monthly Return	Month of Minimum Return
Aggregate Hedge Fund Index	8.0%	10.0%	0.80	0.52	8.1%	12/99	–8.0%	08/98
Convertible Arbitrage	5.4%	5.1%	1.07	0.13	3.1%	04/00	–5.1%	08/98
Event-Driven	7.0%	6.7%	1.05	0.60	3.4%	01/94	–12.2%	08/98
Equity Market-Neutral	6.4%	3.5%	1.85	0.48	2.8%	07/97	–1.6%	03/97
Fixed-Income Arbitrage	1.6%	4.4%	0.36	0.08	1.5%	04/95	–7.3%	10/98
Long/Short Equity	11.8%	12.6%	0.94	0.62	12.6%	12/99	–11.9%	08/98
Emerging Markets	2.3%	20.8%	0.11	0.50	16.1%	08/94	–23.4%	08/98
Global Macro	7.7%	14.4%	0.54	0.36	10.1%	08/95	–11.9%	10/98
Managed Futures	–1.2%	11.1%	–0.10	0.01	9.5%	08/98	–9.8%	09/95
Dedicated Short Bias	–7.1%	18.6%	–0.38	–0.76	22.3%	08/98	–9.1%	02/00
S&P 500	14.6%	14.2%	1.03	1.00	9.3%	03/00	–14.9%	08/98

All returns are excess of the one-month T-bill return. Annualized excess return calculated by multiplying monthly excess returns by 12. Annualized Sharpe ratio equals ratio of annualized excess return and annualized standard deviation.

realize close to zero correlation with the market over this period, also have their best month in August 1998. This result is consistent with previous studies that find that CTA returns exhibit a non-linear "long volatility" characteristic, where beta increases in big up markets and decreases in big down markets (Fung and Hsieh [2000]).

Finally, note that volatility varies considerably across hedge fund indexes. Using monthly data, realized volatility ranges from 3.5% per year for equity market-neutral funds to 20.8% per year for emerging markets funds. This clearly demonstrates the need for the simple, yet crucial concept of risk-adjusting returns when analyzing and comparing hedge fund performance.

Adjusting Hedge Fund Returns for Market Exposure

Given the level of market exposure (both positive and negative) evident in hedge fund returns, it behooves investors to determine whether hedge funds are adding value beyond the returns that they derive from this exposure. How a particular hedge fund makes money is an important issue for at least two reasons:

- Diversification: If a hedge fund has passive market exposure, and generates positive returns because the market goes up, it will possibly lose money when the market goes down, and thus will not provide the diversification benefits some hedge fund investors seek.
- Fees: Investors should not pay hedge fund fees for exposure they can get from index funds at a fraction of the cost.

Many hedge funds admit explicitly that they have market exposure, but claim that they add excess returns above and beyond that exposure. *Caveat emptor.* Investors should understand what this means about the pricing of these funds. These hedge funds can be viewed as selling a bundled package that includes (in proportions that vary across managers) 1) an S&P 500 index fund (or some other equity index fund) and 2) some manager skill. The skill needs to be good because the manager usually charges at least a 1% management fee plus a 20% performance fee on the S&P 500 index fund portion when the going rate is about 20 basis points from Vanguard (and even lower for institutional clients).

Unfortunately, investors cannot easily disentangle skill from market exposure. In part due to less stringent disclosure requirements, investors are rarely privy to hedge fund holdings and investment strategies. Therefore, in most cases, investors must base evaluations only on superficial strategy descriptions and historical returns.

A common approach to estimating hedge fund market exposure is to run regressions of monthly hedge fund returns on S&P 500 returns of the form:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t} \tag{1}$$

where $R_{i,t}$ represents the net-of-fees return on hedge fund i in excess of cash, and $R_{m,t}$ represents the return on the S&P 500 in excess of cash and in excess of an annual fee of 20 basis points (as the hedge fund returns we study are net of fees, we also remove fees from our market proxy).

For simplicity, we take the S&P 500 as a proxy for equity market risk and ignore other potential measures of systematic risk such as value, size, momentum, international and emerging equity, credit, and liquidity. While these other risk factors represent an important extension of our study, as a first step we focus only on equity exposure, which generally embodies the most important risk in most investors' portfolios (even many fixed-income portfolios have equity exposure through spread, liquidity, or credit risk). Additionally, we consider only linear relations between hedge fund returns and the equity market.[5]

We can rewrite the regression in Equation (1) as:

$$R_{i,t} - \beta_i R_{m,t} = \alpha_i + \varepsilon_{i,t} \tag{2}$$

$R_{i,t} - \beta_i R_{m,t}$ can be interpreted as the return on a hedged strategy, where we short β_i units of the S&P 500 against our purchase of hedge fund *i*. Thus, α_i represents the average return for the "hedged" hedge fund (since $\varepsilon_{i,t}$ has a zero mean) and can be interpreted as the manager's realized skill. In other words, the regression intercept represents an estimate of the added value or alpha that hedge funds produce after accounting for their average market exposure or beta.

Exhibit 2 presents the results of these regressions. The results suggest that, despite possessing market exposure, hedge funds do in general appear to add value.[6] The aggregate index returns 2.63% per year net of fees above and beyond what would be expected from its estimated beta of 0.37.[7]

The alphas for convertible arbitrage and equity market-neutral are statistically significant, which is quite impressive, given the short time period. These funds return 4.78% (t-statistic of 2.35) and 4.69% (t-statistic of 3.84) per year above and beyond what is

Exhibit 2 Monthly Regressions of Excess Hedge Fund Returns on Contemporaneous S&P 500 Excess Returns January 1994–September 2000

	Monthly Regressions		
Portfolio	**Alpha (Annualized %)**	**Beta vs. S&P 500**	**Adjusted R^2**
Aggregate Hedge Fund Index	2.63 (0.76)	0.37 (5.46)	26.5%
Convertible Arbitrage	4.78 (2.35)	0.04 (1.12)	0.3%
Event-Driven	2.93 (1.35)	0.28 (6.62)	34.9%
Equity Market-Neutral	4.69 (3.84)	0.12 (4.89)	22.2%
Fixed-Income Arbitrage	1.24 (0.70)	0.02 (0.71)	–0.6%
Long/Short Equity	3.82 (0.95)	0.55 (6.98)	37.4%
Emerging Markets	–8.38 (–1.15)	0.74 (5.15)	24.2%
Global Macro	2.41 (0.44)	0.37 (3.43)	11.8%
Managed Futures	–1.30 (–0.29)	0.01 (0.12)	–1.2%
Dedicated Short Bias	7.34 (1.50)	–0.99 (–10.34)	57.0%

T-statistics in parentheses. Annualized alpha calculated by multiplying regression intercept by 12. Hedge fund and S&P 500 returns in the regressions are excess of one-month T-bill return.

expected from their respective market exposure, which in the case of convertible arbitrage is near zero.

Interestingly, dedicated short biased funds actually *add* 7.34% per year after accounting for their market exposure, although the alpha is not statistically significant due to the high return volatility. While these managers have had rough performance in absolute terms, it is solely due to their negative market exposure. If hedge funds should not get credit for average positive market exposure, to be fair, they should not be penalized for average negative market exposure. A hedge fund with negative market exposure provides tremendous risk reduction for most portfolios. If on top of that risk reduction it can add alpha, this makes it a valuable asset.

One potential explanation for the strong positive alpha for dedicated short biased funds is survivorship bias. Given the strength of the market over this period, survivorship bias in the dedicated short biased index may be significantly greater than in the other categories. Additionally, dedicated short biased funds probably charged much lower performance fees than other funds over this period. Therefore, our use of net-of-fee returns will upwardly bias our estimates of skill for dedicated short biased funds compared to the other hedge fund styles.

A caveat is in order. Our methodology and interpretation of the results in Exhibit 2 implicitly ignores the possibility that a hedge fund's average market exposure over this period may be tactical (i.e., intentionally more than normal). Part of a hedge fund's skill and added value may come from the ability to forecast the direction of the market over long horizons. Since we examine data for 1994–2000, it is possible that some of the hedge funds that have maintained positive market exposure over this entire period have done so tactically, and thus correctly forecasted the bull market. To the extent that this is true, those hedge funds should get credit for their market-timing ability.

Few hedge funds, however, claim to seek significant added value from market timing. Obviously, the empirical analysis above and the analysis that we present below cannot determine whether the observed market exposure was passive or active, and this can have important implications for assessing manager skill.

Mark-to-Market Problems in Hedge Fund Returns

Stale or Managed Prices

Both academics and practitioners have used the type of analysis represented in Exhibit 2 to evaluate hedge fund returns. Agarwal and Naik [2000b], Liang [2001], and Peskin et al. [2000], among others, find similar results. Yet we argue that the regressions presented in Exhibit 2 may produce misleading results.

The reason is that many hedge funds hold illiquid exchange-traded securities or difficult-to-price over-the-counter securities, and these holdings can lead to non-synchronous movements in returns. If securities do not trade near the end of every month, or if there are no publicly available traded prices, hedge funds have flexibility in how they mark their positions for month-end reporting (managers typically estimate prices using their own models along with broker- dealer input).[8]

Given the widespread practice of computing Sharpe ratios, correlations, betas, and other summary statistics using monthly data, hedge fund managers have a strong incentive to show monthly returns that are both consistent and uncorrelated with the market. This

creates a potentially serious conflict of interest, and can lead to non-synchronous price reactions as managers use this flexibility to smooth their returns.

The presence of stale prices due to either illiquidity or managed pricing can artificially reduce estimates of volatility and correlation with traditional indexes. If hedge funds have positive market exposure, in the event the market trades off near the end of the month, and an illiquid security does not trade or is not accurately marked, the drop in price will not show up until the following month when presumably the security trades or is marked correctly. In some cases, a security may be so illiquid, or the "managing" of pricing so extreme, that it does not get accurately marked for several months. Thus, there may be significant lagged relations between market returns and reported hedge fund returns, rendering simple monthly regression betas understated, perhaps severely.[9]

Longer-Horizon Returns

Using longer-horizon returns is one simple way to alleviate the effects of non-synchronous price reactions on estimates of volatility and correlation. While longer-horizon returns are still affected by stale or managed pricing, the impact will represent a smaller component of these returns.

Exhibit 3 presents a comparison of volatility and correlations computed using monthly returns versus nonoverlapping calendar quarterly returns. In the absence of any month-end pricing problems and if monthly returns are identically and independently distributed, annualized monthly volatility should equal annualized quarterly volatility. In the presence of stale or managed prices, however, annualized monthly volatility should be lower than annualized quarterly volatility.[10]

Exhibit 3 shows that in all categories except managed futures and global macro, quarterly volatility is higher than monthly volatility, and thus Sharpe ratios based on quarterly data are lower than those based on monthly data. Moreover, the difference is considerable for certain styles and appears related to the general liquidity of the underlying assets. Convertible arbitrage funds, which tend to trade in hard-to-price over-the-counter fixed-income securities, experience the greatest increase in volatility (41.5%) while managed futures funds, which tend to trade in highly liquid exchange-traded securities, experience a small drop in volatility going from monthly to quarterly data.

Correlations also rise when quarterly data are used. The aggregate hedge fund index correlation with the S&P 500 rises to 0.64 from 0.52, and most of the subindexes also experience increases.

The evidence in Exhibit 3 supports our hypothesis that both hedge fund volatility and market risk are understated when they are estimated using monthly data.

Lagged Betas

In the presence of stale or managed prices, simple linear regressions of the form we conduct in Exhibit 2 may produce estimates of beta that are biased downward. This is a common problem when estimating betas for small firms, which because of their illiquidity suffer from a similar bias.[11]

Exhibit 3 Monthly versus Quarterly Volatility and Market Correlations January 1994–September 2000

Portfolio	Annualized Excess Return	Monthly Annualized Standard Deviation	Quarterly Annualized Standard Deviation	Quarterly vs. Monthly % Difference	Monthly Correlation with S&P 500	Quarterly Correlation with S&P 500	Quarterly vs. Monthly Difference
Aggregate Hedge Fund Index	8.0%	10.0%	10.1%	0.7%	0.52	0.64	0.12
Convertible Arbitrage	5.4%	5.1%	7.2%	41.5%	0.13	0.23	0.11
Event-Driven	7.0%	6.7%	8.6%	28.3%	0.60	0.64	0.05
Equity Market-Neutral	6.4%	3.5%	4.2%	22.9%	0.48	0.50	0.02
Fixed-Income Arbitrage	1.6%	4.4%	4.8%	9.5%	0.08	0.26	0.18
Long/Short Equity	11.8%	12.6%	13.4%	6.3%	0.62	0.76	0.15
Emerging Markets	2.3%	20.8%	26.6%	28.0%	0.50	0.44	–0.06
Global Macro	7.7%	14.4%	20.8%	–0.2%	0.36	0.41	0.05
Managed Futures	–1.2%	11.1%	14.4%	–5.4%	0.01	–0.29	–0.30
Dedicated Short Bias	–7.1%	18.6%	18.7%	0.6%	–0.76	–0.78	–0.02

All returns are excess of the one-month T-bill return. Quarterly returns calculated by summing monthly returns. Annualized excess return is calculated by multiplying quarterly excess returns by 4.

Dimson [1979] and Scholes and Williams [1977] propose a very simple technique to measure market betas by running regressions of returns on both contemporaneous and lagged market returns of the form:

$$R_{i,t} = \alpha_i + \beta_{0i} R_{m,t} + \beta_{1i} R_{m,t-1} + \beta_{2i} R_{m,t-2} + \beta 3_i R_{m,t-3} + \ldots + \varepsilon_{i,t} \tag{3}$$

If hedge fund returns are not fully synchronous with market returns due to stale or managed prices, then lagged market returns should also be correlated with current hedge fund returns. In this case, the summed beta (i.e., $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \ldots +$) represents a more accurate measure of a hedge fund's true beta with the market.

In other words, we are trying to capture the fact that, assuming there is a real relation between the market and a hedge fund, when the market moves, the hedge fund should also move (e.g., if managers actually tried to sell their securities they would see this return immediately). Yet stale or managed pricing may prevent the move from fully showing up in the hedge fund's reported returns in the same month. Instead, the move may show up slowly as securities are priced correctly in subsequent months. Thus, the hedge fund appears to move with the market at a lag.

The regression with lagged market returns measures the magnitude and statistical significance of this effect and provides a potentially more accurate beta estimate. Exhibit 4A presents the details of these regressions using three months of lagged market returns, and Exhibit 4B summarizes the results.

Exhibits 4A and 4B show some dramatic increases in beta over Exhibit 2. Lagged betas enter the regression strongly for the aggregate hedge fund index and for almost every hedge fund style. For the aggregate index, the beta more than doubles from 0.37 (column (1) in Exhibit 4B) in the simple monthly regressions to 0.84 (column (4) in Exhibit 4B) when we account for lagged relations. Perhaps most surprising, convertible arbitrage betas increase dramatically from 0.04 to 0.43. Other large increases include event-driven, which increases from 0.28 to 0.61, and fixed-income arbitrage, which increases from 0.02 to 0.36.[12]

In fact, in every category save managed futures, the betas are magnified. The styles with positive betas produce even larger positive betas, and the styles with negative betas produce even more negative betas.

While it is clear that the lagged betas are economically significant, we also report tests of statistical significance in Exhibit 4A. We separately test the null hypotheses that the sum of all the betas (both contemporaneous and lagged) equals zero (column labeled Sum All Betas) and the sum of just the lagged betas equals zero (column labeled Sum Lagged Betas.)

For every category except managed futures, these F-tests strongly reject the null hypothesis that the summed beta is zero. As expected, the statistical results for just the lagged betas are not as strong, but we still reject the null hypothesis for all categories except equity market-neutral, emerging markets, managed futures, and dedicated short biased. In particular, the tests strongly reject the null hypothesis that the lagged betas are zero for convertible arbitrage, event-driven (which includes distressed debt), and fixed-income arbitrage, which represent the categories most likely to include significant illiquid exchange-traded securities and hard-to-price over-the-counter securities.

Exhibit 4A Monthly Regressions of Excess Hedge Fund Returns on Contemporaneous and Lagged Excess S&P 500 Returns January 1994–September 2000

	Regression Coefficients and t-Statistics						Hypothesis Testing	
Portfolio	Alpha (annualized %)	Beta with S&P 500 (t)	Beta with S&P 500 (t – 1)	Beta with S&P 500 (t – 2)	Beta with S&P 500 (t – 3)	Adjusted R^2	Sum All Betas (= 0)	Sum Lagged Betas (= 0)
Aggregate Hedge Fund Index	–4.45	0.40	0.12	0.22	0.10	35.3%	0.84	0.44
	(–1.16)	(6.21)	(1.85)	(3.37)	(1.45)		(0.0%)	(0.1%)
Convertible Arbitrage	–0.98	0.08	0.16	0.13	0.07	23.8%	0.43	0.35
	(–0.46)	(2.16)	(4.31)	(3.46)	(1.82)		(0.0%)	(0.0%)
Event-Driven	–2.12	0.31	0.18	0.08	0.05	47.0%	0.61	0.30
	(–0.91)	(8.04)	(4.39)	(1.89)	(1.19)		(0.0%)	(0.0%)
Equity Market-Neutral	3.36	0.13	0.05	0.01	0.02	23.4%	0.20	0.08
	(2.32)	(5.18)	(1.95)	(0.39)	(0.84)		(0.1%)	(10.8%)
Fixed-Income Arbitrage	–3.78	0.05	0.10	0.15	0.06	25.2%	0.36	0.31
	(–2.08)	(1.61)	(3.23)	(4.84)	(1.83)		(0.0%)	(0.0%)
Long/Short Equity	–2.83	0.57	0.10	0.18	0.14	40.9%	0.99	0.42
	(–0.61)	(7.39)	(1.25)	(2.24)	(1.76)		(0.0%)	(0.9%)
Emerging Markets	–16.20	0.79	0.30	0.10	0.06	25.3%	1.25	0.46
	(–1.88)	(5.47)	(2.02)	(0.68)	(0.39)		(0.0%)	(11.8%)
Global Macro	–6.64	0.41	0.12	0.37	0.09	21.1%	0.98	0.57
	(–1.08)	(3.94)	(1.12)	(3.45)	(0.83)		(0.0%)	(0.7%)
Managed Futures	1.72	–0.01	–0.15	–0.01	–0.02	–1.9%	–0.19	–0.17
	(0.32)	(–0.15)	(–1.58)	(–0.10)	(–0.19)		(38.3%)	(34.1%)
Dedicated Short Bias	11.59	–1.01	–0.15	0.02	–0.13	57.5%	–1.27	–0.25
	(2.00)	(–10.45)	(–1.51)	(0.22)	(–1.26)		(0.0%)	(19.7%)

T-statistics in parentheses. The last two columns report the sum of the contemporaneous and lagged betas (Sum All Betas) and the separate sum of the lagged betas (Sum Lagged Betas); p-values for the F-test versus zero shown in parentheses. Hedge fund and S&P 500 returns used in the regressions are excess of the one-month T-bill return.

Exhibit 4B Summary of Monthly Regressions of Excess Hedge Fund Returns on S&P 500 Returns January 1994–September 2000

	(1)	(2)	(3)	(4)	(4) – (1)
		Betas from Lagged S&P 500 Regressions (Exhibit 4A)			
Portfolio	**Simple Monthly Regression Beta (Exhibit 2)**	**Contemporaneous Beta (β_0)**	**Sum of Lagged Betas ($\beta_1 + \beta_2 + \beta_3$)**	**Total Summed Beta ($\beta_0 + \beta_1 + \beta_2 + \beta_3$)**	**Difference in Beta**
Aggregate Hedge Fund Index	0.37	0.40	0.44	0.84	0.47
Convertible Arbitrage	0.04	0.08	0.35	0.43	0.38
Event-Driven	0.28	0.31	0.30	0.61	0.33
Equity Market-Neutral	0.12	0.13	0.08	0.20	0.09
Fixed-Income Arbitrage	0.02	0.05	0.31	0.36	0.33
Long/Short Equity	0.55	0.57	0.42	0.99	0.45
Emerging Markets	0.74	0.79	0.46	1.25	0.51
Global Macro	0.37	0.41	0.57	0.98	0.61
Managed Futures	0.01	–0.01	–0.17	–0.19	–0.20
Dedicated Short Bias	–0.99	–1.01	–0.25	–1.27	–0.28

Overall, Exhibits 4A and 4B suggest that when we account for non-synchronous pricing, hedge funds seem to do a lot less hedging than simple estimates might suggest.

One quite surprising result is the large increase in beta for long/short equity, as these funds are generally perceived to trade liquid exchange-traded stocks. The beta of long/short equity funds jumps from 0.55 measured using traditional techniques, to almost double that at 0.99 using our lagged technique (i.e., we find long/short equity funds as exposed to stock market risk as an S&P 500 index fund).

There are at least three potential explanations for this jump. First, long/short funds may take significant positions in small-capitalization stocks. Again, a security does not have to be traded over-the-counter or highly illiquid for our lagged technique to be applicable. In fact, the technique was originally developed, not for illiquid fixed-income securities, but for small-capitalization stocks.[13]

Second, funds in the long/short equity index, and in fact funds in other categories, may have some portion of their assets invested in highly illiquid private securities. Third, any statistical result is subject to sampling error.

While we strongly statistically reject the hypothesis that there is no lagged effect for long/short equity funds, it is certainly possible that our results are on the high side of reality due to random coincidental fluctuation over this period.

Clearly, the LTCM/Russian debt crisis during the fall of 1998 represents an influential data point in our analysis. Excluding the extreme four-month period of August–November 1998 from the regressions in Exhibit 4A, however, only mildly weakens our results. The aggregate hedge fund index produces a summed beta of 0.69 when we exclude this period versus 0.84 for the full sample. The summed beta of 0.69 is still significantly higher than the comparable simple monthly regression beta of 0.41, with the lagged portion statistically significant. Thus, while weaker, our basic results are robust to excluding this volatile, although perhaps informative, period.[14]

While our results are consistent with significant non-synchronous pricing problems in monthly hedge fund returns, they are also consistent with a real lead-lag relation between hedge fund returns and market returns. In other words, our findings might be due to actual reaction of hedge fund returns to moves in the market at a lag (not a lag in marking). For example, Mitchell and Pulvino [2000] find that the probability of deal failure for pending mergers is negatively related to lagged market returns, which could induce a positive relation between merger arbitrage strategy returns and lagged market returns.

While it is difficult to disentangle these two competing hypotheses, we find that across hedge fund styles, the significance of the lagged betas is roughly correlated with the underlying illiquidity of the securities they trade. Furthermore, if these are really true predictive effects, it means that the market and some very savvy hedge fund managers are ignoring this source of potential predictability and profits. In addition, if hedge fund investors cannot trade on this information (given lock-ups or limited liquidity for redemptions), it does not matter which explanation is correct, as both explanations imply more market risk for the hedge fund buyer.[15]

Asymmetric Betas

Assuming that the lagged betas we find are due to non-synchronous pricing, it remains unclear whether the non-synchronous pricing effects are caused by unintentional stale pricing due to illiquidity or from intentional managed pricing (or some of both). To shed some light on this topic, we examine the lagged betas separately for up and down markets.

Presumably, if unintentional stale prices produce the lagged betas, the effect would be symmetrical for up and down markets. If intentional managed pricing drives the results, however, the effect might be asymmetrical, as managers may be more concerned with smoothing downside returns than upside returns. If this is the case, the lagged betas in negative markets would be more significant than those in positive markets.

Exhibit 5 presents a summary of regressions on the CSFB/Tremont indexes that allow for different lagged betas for positive and negative market returns. Instead of three lagged terms as in Exhibit 4A, there are six, each multiplied by a dummy variable representing whether the market was up or down that month.

The regressions show that the summed lagged beta in up markets for the aggregate hedge fund index is 0.17, which is not statistically significant, while the summed lagged beta in down markets is 0.79, which is highly statistically significant. The test of the difference in these lagged betas is also highly statistically significant.

While it is certainly not conclusive, we think this is circumstantial evidence that intentional manager smoothing is at least a part of the effect we document.

Exhibit 5 Monthly Regressions of Excess Hedge Fund Returns on Contemporaneous and Lagged Excess S&P 500 Returns in Up and Down Markets—January 1994–September 2000

	Betas in Up Markets			Betas in Down Markets		
Portfolio	**Contemporaneous Beta (β_0)**	**Sum of Lagged Betas ($\beta_1 + \beta_2 + \beta_3$)**	**Total Summed Beta ($\beta_0 + \beta_1 + \beta_2 + \beta_3$)**	**Contemporaneous Beta (β_0)**	**Sum of Lagged Betas ($\beta_1 + \beta_2 + \beta_3$)**	**Total Summed Beta ($\beta_0 + \beta_1 + \beta_2 + \beta_3$)**
Aggregate Hedge Fund Index	0.27	0.17	0.44	0.59	0.79	1.38
Convertible Arbitrage	0.02	0.22	0.24	0.17	0.47	0.64
Event-Driven	0.03	0.24	0.27	0.61	0.32	0.93
Equity Market-Neutral	0.13	0.18	0.31	0.11	–0.05	0.06
Fixed-Income Arbitrage	0.02	0.06	0.08	0.13	0.57	0.70
Long/Short Equity	0.41	0.51	0.92	0.75	0.36	1.11
Emerging Markets	0.45	0.01	0.46	1.21	1.01	2.22
Global Macro	0.33	–0.06	0.27	0.60	1.38	1.98
Managed Futures	0.29	–0.20	0.09	–0.34	–0.06	–0.40
Dedicated Short Bias	–0.72	–0.60	–1.32	–1.28	0.15	–1.13

Hedge Fund Alphas

Notice what happens to the intercepts from the regressions in Exhibit 4A. The alphas that are generally positive in Exhibit 2 are now generally negative, although not statistically significant. In Exhibit 2, the overall index of hedge funds adds about 2.6% annually when we account for market exposure using simple monthly regressions; when we account for the lagged betas, the hedge fund index subtracts about 4.5% annually. Since our summed betas are consistently higher, we require higher hedge fund returns over this strong period for the market. Since this did not occur, the intercepts are lower. This result is especially surprising, given the likely presence of survivorship bias in the data, which would most likely bias alphas upward.[16]

Exhibit 6 summarizes the Sharpe ratios of three different strategies:

1. The simple unhedged returns to each hedge fund index (from Exhibit 1): $R_{i,t}$
2. A monthly beta-hedged hedge fund index (using betas from Exhibit 2): $R_{i,t} - \beta_i R_{m,t}$
3. A summed beta-hedged hedge fund index (using betas from Exhibit 4A): $R_{i,t} - (\beta_{0i} + \beta_{1i} + \beta_{2i} + \beta_{3i})\, R_{m,t}$

Strategies (2) and (3) represent portfolios invested in a particular hedge fund index and short enough of the S&P 500 to reduce the in-sample beta to zero (using either the simple monthly beta from Exhibit 2 or the summed beta from Exhibit 4A).

The Sharpe ratio of the overall index drops from a respectable 0.80 to a disappointing –0.40 when we account for the summed beta.[17] In fact, for every category except managed futures and dedicated short bias, Sharpe ratios drop when we adjust for market exposure according to summed betas. The only styles with positive excess returns versus our summed beta estimates are equity market-neutral, which goes from a very high, unhedged Sharpe ratio of 1.85 to a still excellent 1.05, and managed futures and dedicated short biased, which benefit significantly when their negative betas are taken into account.

Adjusting for summed betas does not always make hedge fund managers appear less skillful. For example, for the year 2000 (January–September) hedge fund performance actually improves when adjusted correctly for market exposure. Over this period the aggregate hedge fund index is up 4.9% in absolute terms or a mediocre 0.50% above cash. The S&P 500 was down over this period, and, given their positive market exposure, we expect hedge funds to also be down. When we account for their market exposure using our simple beta estimate of 0.37, hedge funds actually added 2.7% above cash. Because our results indicate that the 0.37 estimated beta may be understated, when we adjust for market exposure using our summed beta estimate of 0.84, hedge funds actually add a very respectable 5.5%. In other words, this was a very good period for hedge funds in aggregate when viewed relative to our summed beta estimate.

In fact, it was a banner period for some categories, in particular, convertible arbitrage. Convertible arbitrage funds returned 20% in excess of cash, which is obviously excellent, given their low volatility. This performance improves to 22.5% when adjusted for the summed beta of 0.43. Perhaps the recent strong hedge fund performance represents

Exhibit 6 Annual Sharpe Ratios of Unhedged and Hedged Hedge Fund Returns January 1994–September 2000

	Unhedged and Hedged Performance		
Portfolio	**Monthly Unhedged Sharpe Ratio**	**Monthly Beta-Hedged Sharpe Ratio**	**Summed Beta-Hedged Sharpe Ratio**
Aggregate Hedge Fund Index	0.80	0.31	–0.40
Convertible Arbitrage	1.07	0.95	–0.11
Event-Driven	1.05	0.55	–0.27
Equity Market-Neutral	1.85	1.55	1.05
Fixed-Income Arbitrage	0.36	0.28	–0.56
Long/Short Equity	0.94	0.39	–0.23
Emerging Markets	0.11	–0.47	–0.82
Global Macro	0.54	0.18	–0.40
Managed Futures	–0.10	–0.12	0.14
Dedicated Short Bias	–0.38	0.61	0.89

support for the conjecture that hedge funds are better able to add value through slowly falling markets versus sharply crashing markets, or perhaps it is a harbinger that hedge funds are starting to hedge (maybe due to the painful 1998 experience), or perhaps it represents simple random fluctuation. Only time will tell.

Some additional caveats are in order. First, while our point estimate for the aggregate hedge fund index intercept or manager alpha is negative, the result is not statistically significant. Hedge fund returns are just too volatile and available for too short a time period to make any definitive statements about alpha.

Second, it is likely that the potentially important biases in hedge fund returns will affect our estimates of alpha much more than our estimates of beta. While these biases (especially survivorship bias) probably lead to overstated alphas, they still add to our uncertainty.

Third, the alphas we report are based on the implicit assumption that the equity market is the only relevant risk factor for hedge funds in aggregate and for each particular hedge fund style. While we believe market beta is important, it is clearly not the only relevant risk factor in hedge fund returns, and future research should address multiple risk factors in more depth.

Robustness to Hedge Fund Research Indexes

As a final robustness check, we run analogous tests on the Hedge Fund Research (HFR) hedge fund index return series. While there are many differences between the HFR and CSFB/Tremont data, the most notable is that the CSFB/Tremont indexes are asset-weighted, while the HFR indexes are equally weighted. Thus, the HFR data will give more emphasis to smaller hedge funds. Peskin et al. [2000] find that smaller and newer hedge funds generally have higher returns, but note that these funds are the most susceptible to reporting biases and the problem of not being fully investible (i.e., a large mandate could not be invested in these funds and run in a comparable style to the one that has produced the current track record).

The results for the HFR data support our hypothesis that hedge fund market risk, when estimated using monthly returns, is significantly understated. The simple regression beta of the HFR aggregate hedge fund index versus the S&P 500 over this period is 0.38, but jumps to 0.65 when the statistically significant lagged betas are taken into account.

The "unhedged" monthly Sharpe ratio of the HFR aggregate index is 1.23, but drops to 0.03 when measured as excess returns versus the summed beta. Although both the pre- and post-hedging risk-adjusted returns are better for the HFR series, the 1.20 fall in the Sharpe ratio is exactly the same magnitude as the fall in Sharpe ratio of the CSFB/Tremont aggregate hedge fund index (which fell from 0.80 to –0.40).

In addition, results for the comparable HFR subcategories are generally similar to those we report, with the exception of merger arbitrage. HFR provides a separate merger arbitrage index, while CSFB/Tremont aggregates merger arbitrage along with distressed securities, Regulation D, and high-yield into its event-driven index. The HFR merger arbitrage index does show strong and statistically significant positive risk-adjusted returns over the 1994–2000 period, even after accounting for lagged betas, which are positive but not statistically significant.

Conclusion

The recent strong bull market has produced enormous wealth. It has also been a democratic bargain, available for the participation of all investors through index funds costing only a handful of basis points per year. Hedge funds, on the other hand, are restricted investments generally available only to the few, and come at a high price tag. For this high cost, hedge funds generally claim to offer attractive returns that cannot be obtained by investing in index funds. To the extent that hedge funds can achieve these goals, they offer an investment that can be an important source of expected return and diversification for most investors' portfolios.

At first pass, cursory examination using monthly returns from 1994–2000 suggests that hedge fund investors have in fact received a great deal both in terms of return and diversification. Intentionally or unintentionally, though, hedge funds appear to price their securities at a lag. These marking problems can downwardly bias simple risk estimates based on monthly returns. When we account for this effect, we find that the return and diversification benefits vanish for the broad hedge fund universe and many subcategories.[18]

We propose that instead of examining simple Sharpe ratios calculated using monthly returns, investors focus on *hedged Sharpe ratios* that take into account market exposure, and use the lagged beta techniques we propose to estimate more accurate betas.[19] For hedge funds that are truly uncorrelated with the market, both simple and hedged Sharpe ratios will produce the same results, but for funds with large average market exposure the results may differ dramatically.

While the short 1994–2000 period makes drawing definitive conclusions about alpha difficult, this is not the case regarding beta. The evidence we present strongly suggests that non-synchronous pricing problems, whether due to stale or managed prices, are a significant issue in monthly hedge fund data and can lead to severely understated estimates of hedge fund risk. To the extent this is true, it follows directly that estimates of hedge fund value-added (alpha) will be overstated over any period with a rising equity market.

Our results suggest that constructing a successful hedge fund portfolio is not as simple as other studies indicate. Since we study only indexes, we certainly cannot conclude that attractive uncorrelated hedge funds (or hedge funds that add value net of their market exposure) do not exist. It may be possible to eliminate funds whose risk-adjusted returns are being inflated by the illiquidity we study, and be left with a portfolio of funds that provide attractive expected returns and diversification.

Many hedge funds make bold claims, and some produce monthly returns that seem to support those claims. Our results prompt us to warn that careful scrutiny of these claims is important and—at a minimum—that researchers and investors should use the techniques we discuss or similar methods when they evaluate hedge fund returns.

Appendix
CSFB/Tremont Subindex Categories*

Convertible Arbitrage

This strategy is identified by hedge investing in the convertible securities of a company. A typical investment is to be long the convertible bond and short the common stock of the same company. Positions are designed to generate profits from the fixed-income security as well as the short sale of stock, while protecting principal from market moves.

Event-Driven

This strategy is defined as equity-oriented investing designed to capture price movement generated by an anticipated corporate event. There are four popular subcategories in event-driven strategies: risk arbitrage, distressed securities, Regulation D, and high-yield investing.

Risk Arbitrage. Specialists invest simultaneously in long and short positions in both companies involved in a merger or acquisition. Risk arbitrageurs are typically long the stock of the company being acquired and short the stock of the acquirer. The principal risk is deal risk, should the deal fail to close.

Distressed Securities. Fund managers invest in the debt, equity, or trade claims of companies in financial distress and generally bankruptcy. The securities of companies in need of legal action or restructuring to revive financial stability typically trade at substantial discounts to par value, and thereby attract investments when managers perceive a turnaround will materialize.

Regulation D, or Reg D. This subset refers to investments in micro and small-capitalization public companies that are raising money in private capital markets. Investments usually take the form of a convertible security with an exercise price that floats or is subject to a look-back provision that insulates the investor from a decline in the price of the underlying stock.

High-Yield. Often called junk bonds, high-yield refers to investment in low-grade fixed-income securities of companies that show significant upside potential. Managers generally buy and hold high-yield debt.

Equity Market-Neutral

This investment strategy is designed to exploit equity market inefficiencies and usually involves being simultaneously long and short matched equity portfolios of the same size within a country. Market-neutral portfolios are designed to be either beta or currency-neutral, or both. Well-designed portfolios typically control for industry, sector, market capitalization, and other exposures. Leverage is often applied to enhance returns.

Fixed-Income Arbitrage

The fixed-income arbitrageur aims to profit from price anomalies between related interest rate securities. Most managers trade globally with a goal of generating steady returns with low volatility. This category includes interest rate swap arbitrage, U.S. and non-U.S. government bond arbitrage, forward yield curve arbitrage, and mortgage-backed securities arbitrage. The mortgage-backed market is primarily U.S.-based, over-the-counter, and particularly complex.

Long/Short Equity

This directional strategy involves equity-oriented investing on both the long and the short sides of the market. The objective is not to be market-neutral. Managers have the ability to shift from value to growth, from small- to medium- to large-capitalization stocks, and from a net long position to a net short position. Managers may use futures and options to hedge. The focus may be regional, such as long/short U.S. or European equity, or sector-specific, such as long and short technology or health care stocks. Long/short equity funds tend to build and hold portfolios that are substantially more concentrated than those of traditional stock funds.

Emerging Markets

This strategy involves equity or fixed-income investing in emerging markets around the world. Because many emerging markets do not allow short-selling, or offer viable futures or other derivative products with which to hedge, emerging market investing often employs a long-only strategy.

Global Macro

Global macro managers carry long and short positions in any of the world's major capital or derivative markets. These positions reflect their views on overall market direction as influenced by major economic trends and/or events. The portfolios of these funds can include stocks, bonds, currencies, and commodities in the form of cash or derivatives instruments. Most funds invest globally in both developed and emerging markets.

Managed Futures

This strategy invests in listed financial and commodity futures markets and currency markets around the world. The managers are usually referred to as commodity trading advisors, or CTAs. Trading disciplines are generally systematic or discretionary. Systematic traders tend to use price and market-specific information (often technical) to make trading decisions, while discretionary managers use a less structured approach.

Dedicated Short Bias

Dedicated short-sellers were once a robust category of hedge funds before the long bull market rendered the strategy difficult to implement. A new category, short-biased, has emerged. The strategy is to maintain net short as opposed to pure short exposure. Short-biased managers take short positions in mostly equities and derivatives. The short bias of a manager's portfolio must be constantly greater than zero to be classified in this category.

*Source: CSFB/Tremont web site (http://www.hedgeindex.com).

Endnotes

The authors thank Alan Au, Richard Anderson, Mark Anson, Robert Arnott, Brad Barber, Franco Baseggio, Gib Bassett, Giovanni Beliossi, Peter Bernstein, William Bernstein, Adam Blitz, Kenneth Brown, William Brown, Henk Brulleman, Arthur Cinader, Kent Clark, Roger Clarke, Chris Darnell, Thomas Dunn, Claude Erb, Cedric Fan, Kenneth French, Jacques Friedman, Steven Galbraith, Ibrahim Gharghour, Ira Handler, Brian Hurst, Antti Ilmanen, Raymond Iwanowski, David Kabiller, Oktay Kurbanov, Diana Lieberman, Hal Lux, George Main, David Patterson, David Pelleg, Thomas Philips, Robert Pinchuk, Burt Porter, Todd Pulvino, John Rowsell, Rudi Schadt, Laurence Siegel, Robert Snigaroff, Ross Stevens, Steven Thorley, Todd Tibbets, Ingrid Tierens, and Geoff Verter for valuable comments and suggestions.

[1]Asness [1998] discusses this topic in more detail.

[2]See the CFSB/Tremont website for more detailed information (www.hedgeindex.com). The constituents of each index are publicly available.

[3]Please see Liang [2000] for an in-depth comparison of the TASS and Hedge Fund Research databases and an analysis of survivorship bias in hedge fund returns.

[4]Additionally, since these indexes represent portfolios of hedge funds, styles that represent 1) more funds and 2) more heterogeneous funds will tend to produce indexes with lower volatility.

[5]Agarwal and Naik [2000a], Fung and Hsieh [2000], and Mitchell and Pulvino [2000] study the performance of hedge funds using non-linear models. Additionally, we use net-of-fee returns, and managers typically charge a performance fee of 20% of any positive returns. Therefore, even if gross returns are linear in the market, net returns will not be. Since the beta on net returns will be 80% of the beta on gross returns on the upside and 100% on the downside, our net-of-fee betas may underestimate the betas on gross returns.

[6]The market correlations reported in Exhibit 1 are invariant to volatility and thus they may be compared directly across funds with differing volatilities. Betas are not invariant to volatility; the higher an asset's volatility, the higher the asset's beta (for the same level of correlation). For example, equity market-neutral has an estimated beta of only 0.12, but since its volatility is so low (3.5% annualized) this is actually a very significant beta both economically (annual returns are

reduced from an unadjusted 6.4% to an alpha of 4.7%) and statistically (t-statistic of 4.89). At the other end of the spectrum, emerging markets has a 0.74 beta, but its t-statistic of 5.15 is comparable to equity market-neutral.

In the simple regression framework of Exhibit 2, the relation between beta and correlation is as follows:

$$\beta_{i,m} = \rho_{i,m}(\sigma_i/\sigma_m)$$

Therefore, the beta for an index is equal to its correlation with the S&P 500 times the ratio of the hedge fund's volatility to the S&P 500's volatility. For example, the beta of 0.12 for equity market-neutral can be computed from the data in Exhibit 1 as follows: 0.48(3.5%/14.2%) = 0.12.

[7]Since volatility varies considerably across the hedge fund styles, naive comparisons of the magnitude of the alphas can be quite misleading. In contrast, both the t-statistic testing the hypothesis that the alpha is equal to zero and the Sharpe ratio of the "hedged" hedge fund in Exhibit 6 (defined as α/volatility $[R_{i,t} - \beta R_{m,t}]$) are comparable across portfolios with varying volatilities.

[8]An extreme example of stale or managed pricing can be found in the "private equity" category of alternative investments. Private equity suffers from a lag in mark-to-market returns that is so severe that the techniques in this article do not apply—but this does not mean the effect does not exist. On the contrary, it is likely that most empirical estimates of private equity risk (beta and volatility) are severely understated. It appears that traditional hedge fund managers have been crossing over more and now own some private equity-like investments. If this is the case, our techniques will not fully capture the understatement of hedge fund risk. Finally, some private equity managers and investors seem to believe these investments provide tremendous diversification for traditional portfolios, sometimes going so far as to call them "uncorrelated." Well, the ostrich thinks he's uncorrelated also, but that does not fool the lion.

[9]Non-synchronous pricing does not necessarily always reduce estimates of beta and volatility. For example, monthly volatility can be overstated for truly market-neutral strategies that are long and short securities that trade in different time zones (or are illiquid). If a hedge fund is long Japan and short the U.S., then non-synchronous closing prices can lead to overstated estimates of volatility. For example, if on the last day of the month the U.S. trades off, since Japan is closed, the portfolio is marked with the U.S. move, but no concurrent Japan move. Of course, the next day Japan would most likely gap down, but that would show up only in the following month's return.

[10]There are other potential reasons, besides simple sampling error, for annualized quarterly volatility to exceed annualized monthly volatility. For example, if true hedge fund returns (marked timely and accurately) are positively autocorrelated at the monthly level, we will observe an increase in volatility. In the absence of such autocorrelation in the S&P 500, we do not expect a corresponding increase in market correlation from this effect. Over our sample period, the S&P 500 monthly returns exhibit mildly negative, but statistically insignificant, autocorrelation at each of the first three lags.

[11]A downward bias applies only for the more common case of positive market exposure. In the case of negative market exposure, as in dedicated short biased funds, non-synchronous price reactions actually bias estimates of beta upward in simple regressions (i.e., estimates are less negative than actual betas).

[12]Event-driven includes a fairly heterogeneous combination of risk arbitrage, distressed debt, high-yield, and Regulation D funds.

[13]For example, Ibbotson, Kaplan, and Peterson [1997] apply this summed beta technique to empirical tests of the capital asset pricing model. They argue that simple measures of small-capitalization stock betas are understated due to stale pricing, and thus factors like size tend to empirically dominate beta. When they account for lagged effects, the beta estimates for small-capitalization stocks increase substantially, and they find that beta does a good job in explaining cross-sectional differences in average returns.

[14]As an additional robustness check, we run the regressions in Exhibit 4A with the three lagged beta coefficients constrained to equal the same value as opposed to letting them vary. This has no material effect on our results. In particular, the restricted regression produces a summed beta for the aggregate hedge fund index of 0.80, as opposed to 0.84 for the unrestricted regression reported in Exhibit 4A.

[15]Hedge funds often require 30 days or more advance notification for fund withdrawals.

[16]While we focus on market risk, we also conduct initial tests using other potential explanatory risk factors. If in addition to the S&P 500, we include in our lagged regression tests the excess return of small-capitalization stocks over large stocks, we find that the aggregate hedge fund index and many hedge fund styles have a small-capitalization stock bias over this period. Since small stocks underperformed large stocks over the 1994–2000 period, this bias helps explain the negative alphas we find. In fact, when accounting for the small-capitalization stock factor, the aggregate hedge fund index has neither added nor subtracted value over this period. Our main result that standard monthly regression betas understate true beta, and thus simple estimates of performance overstate skill, remains unchanged.

We also conduct initial tests on the non-linear effects other authors have identified in hedge fund returns. In particular, adding in a squared market return shows that managed futures load positively on this "long volatility" factor, while the aggregate hedge fund index and most other individual categories load negatively on this factor. Unlike our lagged market betas, the non-linear effects are highly sensitive to the inclusion of the volatile months in the fall of 1998. Most important, inclusion of the squared market return in our lagged beta regressions has little impact on our results.

[17]Volatility estimates used in the denominator of the Sharpe ratios for the summed beta-hedged strategies will be overestimated due to non-synchronous pricing differences between the hedge fund returns and the S&P 500 returns. Average return estimates are not biased, however, and thus the Sharpe ratio will generally be biased toward zero (i.e., less negative for negative Sharpe ratio strategies and less positive for positive Sharpe ratio strategies).

[18]While we do not address this issue here, it is easy to surmise that for a taxable investor the situation is even worse, as taxes would have much more of a negative impact on hedge funds' returns than on an S&P 500 index fund. See Arnott and Jeffrey [1993] for an excellent discussion of the effects of portfolio turnover on after-tax performance of equity funds.

[19]This is akin to using information ratios for measuring active returns against a benchmark.

References

Agarwal, Vikas, and Narayan Naik. "Performance Evaluation of Hedge Funds with Option-Based and Buy-and-Hold Strategies." Working paper, London School of Business, 2000a.

_____. "On Taking the 'Alternative' Route: Risks, Rewards, Style and Performance Persistence of Hedge Funds." *Journal of Alternative Investments*, 2 (2000b), pp. 6–23.

Arnott, Robert, and Robert H. Jeffrey. "Is Your Alpha Big Enough to Cover Its Taxes?" *The Journal of Portfolio Management*, Spring 1993.

Asness, Clifford. "Market-Neutral Investing: Putting the 'Hedge' in Hedge Funds." Working paper, AQR Capital Management, 1998.

Dimson, E. "Risk Measurement when Shares Are Subject to Infrequent Trading." *Journal of Financial Economics,* 7 (1979).

Fung, William, and David A. Hsieh. "The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers." Working paper, Paradigm Financial Products International LLC, and Fuqua School of Business, Duke University, 2000.

Ibbotson, Roger, Paul Kaplan, and James Peterson. "Estimates of Small Stock Betas are Much Too Low." *The Journal of Portfolio Management,* 24, No. 2 (1997), pp. 104–111.

Liang, Bing. "Hedge Fund Performance: 1990–1999." *Financial Analysts Journal*, January/February 2001.

Mitchell, Mark, and Todd Pulvino. "Characteristics of Risk and Return in Risk Arbitrage." Working paper, Harvard Business School and Northwestern University, 2000.

Peskin, Michael, Michael Urias, Satish Anjilvel, and Brian Boudreau. "Why Hedge Funds Make Sense." Working paper, Morgan Stanley Dean Witter, 2000.

Scholes, M., and J.T. Williams. "Estimating Betas from Non-synchronous Data." *Journal of Financial Economics,* 5 (1997).

Weisman, Andrew. "The Dangers of Historical Hedge Fund Data." Working paper, Nikko Securities, 2000.

This page intentionally left blank

Characteristics of Risk and Return in Risk Arbitrage

*Mark Mitchell and Todd Pulvino**

This paper analyzes 4,750 mergers from 1963 to 1998 to characterize the risk and return in risk arbitrage. Results indicate that risk arbitrage returns are positively correlated with market returns in severely depreciating markets but uncorrelated with market returns in flat and appreciating markets. This suggests that returns to risk arbitrage are similar to those obtained from selling uncovered index put options. Using a contingent claims analysis that controls for the nonlinear relationship with market returns, and after controlling for transaction costs, we find that risk arbitrage generates excess returns of four percent per year.

After the announcement of a merger or acquisition, the target company's stock typically trades at a discount to the price offered by the acquiring company. The difference between the target's stock price and the offer price is known as the arbitrage spread. Risk arbitrage, also called merger arbitrage, refers to an investment strategy that attempts to profit from this spread. If the merger is successful, the arbitrageur captures the arbitrage spread. However, if the merger fails, the arbitrageur incurs a loss, usually much greater than the profits obtained if the deal succeeds. In this paper, we provide estimates of the returns to risk arbitrage investments, and we also describe the risks associated with these returns.

Risk arbitrage commonly invokes images of extraordinary profits and incredible implosions. Numerous articles in the popular press detail large profits generated by famous

* Harvard Business School and Kellogg School of Management, respectively. We are grateful to seminar participants at the Cornell Summer Finance Conference, Duke University, Harvard University, the University of Chicago, the University of Kansas, the New York Federal Reserve Bank, the University of North Carolina, Northwestern University, the University of Rochester, and the University of Wisconsin-Madison for helpful comments, and to three anonymous referees, Malcolm Baker, Bill Breen, Emil Dabora, Kent Daniel, Bob Korajczyk, Mitchell Petersen, Judy Posnikoff, Mark Seasholes, Andrei Shleifer, Erik Stafford, René Stulz, Vefa Tarhan, and especially Ravi Jagannathan for helpful discussions. We would also like to thank the many active arbitrageurs who have advanced our understanding of risk arbitrage.

Reprinted from *The Journal of Finance*, December 2001, Vol. 56 (6) 2135-2175, with permission.

arbitrageurs such as Ivan Boesky and even larger losses by hedge funds such as Long Term Capital Management. Overall, existing academic studies find that risk arbitrage generates substantial excess returns. For example, Dukes, Frohlich, and Ma (1992) and Jindra and Walkling (1999) focus on cash tender offers and document annual excess returns that far exceed 100 percent. Karolyi and Shannon (1998) conclude that a portfolio of Canadian stock and cash merger targets announced in 1997 has a beta of 0.39 and an annualized return of 26 percent, almost twice that of the TSE 300. In a similar study using a much larger sample of U.S. cash and stock mergers, Baker and Savasoglu (2002) conclude that risk arbitrage generates annual excess returns of 12.5 percent.

These findings suggest that financial markets exhibit systematic inefficiency in the pricing of firms involved in mergers and acquisitions. However, there are two other possible explanations for the large excess returns documented in previous studies. The first explanation is that transaction costs and other practical limitations prevent investors from realizing these extraordinary returns. The second explanation is that risk arbitrageurs receive a risk premium to compensate for the risk of deal failure. In this paper, we attempt to empirically distinguish between these three alternative explanations.

To assess the effect of transaction costs, we use a sample of 4,750 mergers and acquisitions between 1963 and 1998 to construct two different series of risk arbitrage portfolio returns.[1] The first portfolio return series is a calendar-time value-weighted average of returns to individual mergers, ignoring transaction costs and other practical limitations (value-weighted risk arbitrage returns are subsequently referred to as VWRA returns). The second portfolio return series mimics the returns from a hypothetical risk arbitrage index manager (subsequently referred to as RAIM returns). RAIM returns include transaction costs, consisting of both brokerage commissions and the price impact associated with trading less than perfectly liquid securities. RAIM returns also reflect practical constraints faced by most risk arbitrage hedge funds. However, unlike actively managed hedge funds, no attempt to discriminate between anticipated successful and unsuccessful deals is made when generating RAIM returns. Comparing the VWRA and RAIM return series indicates that transaction costs have a substantial effect on risk arbitrage returns. Ignoring transaction costs results in a statistically significant alpha (assuming linear asset pricing models are valid) of 74 basis points per month (9.25 percent annually). However, when we account for transaction costs, the alpha declines to 29 basis points per month (3.54 percent annually).

The second possible explanation for the extraordinary returns to risk arbitrage documented in previous studies is that they simply reflect compensation for bearing extraordinary risk. Although previous studies that report excess returns attempt to control for risk, they make the implicit assumption that linear asset pricing models are applicable to risk arbitrage investments. However, this assumption is problematic if the returns to risk arbitrage are related to market returns in a nonlinear way. Building on Merton's (1981) work on the ability of fund managers to time the market, Glosten and Jagannathan (1994) show how to evaluate the performance of investment strategies that exhibit nonlinear relationships with market returns. They argue that these types of strategies must be evaluated using a contingent claims approach that explicitly values the nonlinearity. As an example, Fung and Hsieh (2001) demonstrate the presence of extraordinary types of risk in a potential hedge fund strategy referred to as "trend following." They show that the payoff to the trend following strategy is related to the payoff from an investment in a lookback straddle. Glosten and Jagannathan's (1994) argument is also supported by Bhagat, Brickley, and

Loewenstein's (1987) analysis of interfirm cash tender offers. They argue that investing in the target company's stock after the tender offer announcement is like owning the stock plus a put option on the target's stock. Results from their analysis indicate that, when using the Capital Asset Pricing Model (CAPM) to control for risk, there are significant excess returns to investing in tender offers. They conclude that the CAPM does not fully capture the risk associated with tender offer investments.

In this paper, we investigate whether the reason that linear asset pricing models fail to capture the risk from investing in merger stocks is that the returns to merger stock investments are correlated with market returns in a nonlinear way. Results from our analysis indicate that, in flat and appreciating markets, risk arbitrage generates returns 50 basis points per month (6.2 percent annually) greater than the risk-free rate with essentially a zero market beta. However, in months where the stock market experiences a decrease of 4 percent or more, the market beta of the risk arbitrage portfolio increases to 0.50. Thus, our RAIM portfolio generates moderate returns in most environments but, in rare cases, generates large negative returns. This pattern is robust across time periods and is invariant to changes in assumptions used to estimate transaction costs. We conclude that risk arbitrage is akin to writing uncovered index put options. Given this optionlike feature, standard empirical asset pricing models cannot be used to assess the risk/reward performance associated with risk arbitrage, and the alphas reported in previous studies do not represent excess returns. Instead, the risk/reward profile of risk arbitrage must be evaluated using a contingent claims approach similar to the one suggested by Glosten and Jagannathan (1994). The contingent claims approach, rather than linear models, should also be used when generating benchmarks for evaluating active risk arbitrage hedge fund managers.

This paper is the first to document the high correlation between risk arbitrage returns and market returns in down markets. However, the highly nonlinear relationship between risk arbitrage returns and market returns does not explain the excess returns found in previous studies. Using a contingent claims analysis and assuming that there are no transaction costs, we estimate excess returns of 10.3 percent. This is greater than, not less than, the 9.25 percent estimate obtained using CAPM to measure the excess return generated by risk arbitrage investments. When returns that incorporate transaction costs and other practical limitations are used, the contingent claims analysis implies excess returns of 4 percent annually, far less than the excess returns reported in previous studies. These results suggest that not accounting for transaction costs and other practical limitations is the primary explanation for the large excess returns reported in previous studies. This does not mean that the nonlinear relationship between risk arbitrage returns and market returns is inconsequential. Risk arbitrage is appropriate only for those investors that are willing to incur negative returns in severely depreciating markets and limited positive returns in flat and appreciating markets.

To confirm that our findings are not an artifact of the assumptions that we use to estimate transaction costs, we repeat our nonlinear analysis using returns from active risk arbitrage hedge funds over the 1990 to 1998 time period. Results using this sample of hedge fund returns are remarkably similar to those obtained using returns from our RAIM portfolio.

The remainder of this paper is organized as follows. Section I describes typical arbitrage investments. Section II provides a brief overview of existing risk arbitrage research and outlines three explanations of the returns from this strategy. The data used in this paper

are described in Section III. Section IV describes the construction of the time series of risk arbitrage returns. Results are presented in Sections V and VI. Section VII concludes.

I. Description of Typical Investments

There are two primary types of mergers, cash mergers and stock mergers. In a cash merger, the acquiring company offers to exchange cash for the target company's equity or assets. In a stock merger, the acquirer offers its common stock to the target shareholders in lieu of cash. The arbitrageur's investment depends on the form of payment to the target shareholders. In a cash merger, the arbitrageur simply buys the target company's stock. Because the target's stock typically sells at a discount to the payment promised by the acquirer, profits can be made by buying the target's stock and holding it until merger consummation. At that time, the arbitrageur sells the target's common stock to the acquiring firm for the offer price. There are two sources for the return from this investment. The primary source of profit is the difference between the purchase price of the target's stock and the ultimate offer price. The secondary source of profit is the dividend paid by the target company.

In a stock merger, the arbitrageur sells short the acquiring firm's stock in addition to buying the target's stock. In this case, there are three sources of the arbitrageur's profit. The primary source of profit is the difference between the price obtained from the short sale of the acquirer's stock and the price paid for the target's stock. The second source of profit is the dividend paid on the investment in the target's stock. However, this is offset by dividends that must be paid on the acquirer's stock, since it was borrowed and sold short. The third source of profits in a stock deal comes from interest paid by the arbitrageur's

Figure 1. This figure plots the median arbitrage spread versus time until deal resolution. The arbitrage spread is defined to be the offer price minus the target price divided by the target price. For failed deals, the deal resolution date is defined as the date of the merger termination announcement. For successful deals, the resolution date is the consummation date.

broker on the proceeds from the short sale of the acquirer's stock. For individual investors, the interest rate is typically zero. However, for institutions and hedge funds, short proceeds earn interest at a rate close to the risk-free rate.

More complicated deal structures involving preferred stock, warrants, debentures, and other securities are common. From the arbitrageur's perspective, the important feature of all of these deal structures is that returns depend on mergers being successfully completed. Thus, the primary risk borne by the arbitrageur is that of deal failure. Figure 1 displays a representative picture of the losses and gains from risk arbitrage. This figure tracks the median arbitrage spread (the percentage difference between the target's stock price and the offer price) over time, measured from the deal resolution date. For unsuccessful deals, the spread remains relatively wide during the life of the merger. When a merger deal fails, the median spread widens dramatically, increasing from 15 percent to more than 30 percent on the termination announcement day. A much different pattern exists for risk arbitrage investments in successful merger transactions. In successful deals, the arbitrage spread decreases continuously as the deal resolution date approaches. Upon successful consummation of the merger, the spread collapses to zero. The fact that spreads are much wider for unsuccessful transactions suggests that the probability of deal failure is incorporated into the stock prices of target firms.

II. Possible Explanations of Risk Arbitrage Returns

Most of the previous studies that attempt to assess the profitability of risk arbitrage conclude that it generates substantial risk-adjusted returns. Excess returns are greatest in those studies that focus on cash tender offers. Using a sample of 761 cash tender offers between 1971 and 1985, Dukes et al. (1992) conclude that an investor who purchased the target's stock on the day of the tender offer announcement and sold the stock subsequent to the tender offer resolution would have earned a daily return of 0.47 percent. This corresponds to an annualized return well over 100 percent, although the authors concede that it would be difficult for an investor to repeat these returns on a continuing basis. Jindra and Walkling (1999) report similar results. Using a sample of 361 cash tender offers between 1981 and 1995, they find that an arbitrageur who purchased the target stock one day after the acquisition announcement and sold one week later would have generated an annualized excess return between 102 percent and 115 percent. Bhagat et al. (1987) document tender period excess returns of 2.0 percent (18 percent annually, based on an average tender period of 29 days) obtained by buying the target's stock the day after the tender offer announcement and selling one day prior to the offer's expiration.

Studies that use transactions other than cash tender offers also document high returns from risk arbitrage investments. Larcker and Lys (1987) study returns to target stocks that were the subject of SEC 13D filings. Although their sample includes both cash transactions and stock swap mergers, their analysis focuses on the returns obtained from buying the target. They do not examine the typical arbitrage investment that also involves short selling the acquirer's stock. Nevertheless, they find excess returns of 5.3 percent and raw returns of 20.08 percent over the transaction period. Based on the median transaction period of 31 trading days, these numbers correspond to an annualized excess return of 51.9 percent and an annualized raw return of 337 percent. Like Larcker and Lys, Karolyi and Shannon (1998) also study both cash and stock mergers. From a sample of 37 Canadian mergers in 1997, they conclude that a risk arbitrage

portfolio would have generated a beta of 0.39 and an annualized return of 26 percent, almost twice the return achieved by the TSE 300 in 1997. Baker and Savasoglu (2002) use a much larger sample over the 1978 to 1996 time period and conclude that risk arbitrage generates excess returns of 1 percent per month (approximately 12.5 percent annualized).

Results presented in previous risk arbitrage studies are consistent with more recent papers that examine hedge fund returns. Both Aragwal and Naik (1999) and Ackermann et al. (1999) find that risk arbitrage hedge funds generate return0risk profiles that are superior to other hedge fund strategies. However, as Fung and Hsieh (2000) point out, survival biases present in existing hedge fund databases make it difficult to obtain accurate measurements of performance and risk characteristics of specific strategies.

The magnitudes of the returns reported in previous studies suggest that there exists a severe market inefficiency in the pricing of merger stocks. Yet two studies that use a different approach to examine risk arbitrage returns reach the opposite conclusion. Brown and Raymond (1986) use 89 takeover attempts to examine the ability of the arbitrage spread to distinguish between those deals that will ultimately succeed and those that will ultimately fail. Although they report neither returns nor estimates of risk, they find, consistent with market efficiency, that deal failure probabilities are accurately reflected in the target's and acquirer's stock prices. Samuelson and Rosenthal (1986) perform a similar analysis using a sample of cash tender offers. They conclude that the target's stock price measured well before resolution of the tender offer is a good predictor of the stock price at the conclusion of the tender offer. Based on this, they argue that there are few opportunities to earn excess returns by investing in tender offer targets.

In this paper, we use a long time series of risk arbitrage portfolio returns to attempt to distinguish between market inefficiency and two alternative explanations of returns to risk arbitrage investments. The first alternative explanation is that transaction costs and other practical limitations prevent the average investor from realizing the extraordinary gains documented in previous studies. Of the practical limitations, one of the most important stems from the use of event time, rather than calendar time, to calculate risk arbitrage returns. The event-time approach involves calculating the rate of return obtained from investing after the merger announcement and selling after deal resolution. Returns from individual deals are first "annualized" and then averaged across deals. The problem with this approach is that it assumes that the risk arbitrage portfolio can earn event-time returns continuously. Particularly for transactions that are consummated quickly, this assumption can lead to large annualized returns. For example, on December 2, 1996, Zycon Corporation agreed to be acquired by the buyout firm Hicks, Muse, Tate, and Furst for $16.25 per share. Three days later, Hadco Corporation entered a competing bid of $18 per share. An arbitrageur that purchased Zycon stock one day after the original bid would have made a three-day return of 9.5 percent and an annualized return of 1,903 percent. Large returns such as this weigh heavily in the averaging process used to calculate event-time returns. Yet, as some authors point out, it is not realistic to assume that an investor could achieve these returns on a continuing basis (Dukes et al. (1992), Karolyi and Shannon (1998)). To address this issue, we calculate average risk arbitrage returns based on calendar time rather than event time. That is, we simulate a hypothetical risk arbitrage portfolio (discussed in detail below) that correctly models the investment holding period.

The second possible explanation for the extraordinarily large documented returns to risk arbitrage is that they represent compensation to investors for bearing extraordinary

amounts of systematic risk. Because most announced mergers are successfully consummated, risk arbitrage investments usually generate small positive returns. Conditional on successful consummation, these returns depend on the initial arbitrage spread and not on overall stock market returns. Therefore, returns to risk arbitrage should contain very little systematic risk. However, risk arbitrage returns may be positively correlated with market returns during severe market downturns. This will be true if the probability of deal failure increases in depreciating markets. For example, an acquirer that agrees to pay $50 per share for a target company when the S&P 500 index is 1300 may be willing to pay only $30 if the S&P 500 falls to 800. If the acquirer reneges on the deal, the risk arbitrage investment is likely to generate a negative return. This effect will be compounded if investments were made under the belief that risk arbitrage investments are "market neutral." Shleifer and Vishny (1997) argue that even though the hedge fund managers that typically invest in merger situations may understand the risk/return profile associated with risk arbitrage, their investors may not. Consequently, investors may redeem their capital at precisely the wrong time, forcing risk arbitrage hedge fund managers to "bail out of the market when their participation is most needed."[2]

To distinguish between the market inefficiency story and the risk story, we perform two analyses. First, we estimate the CAPM and the Fama and French (1993) three-factor asset pricing model:

$$(R_{Risk\ Arb} - R_f) = \alpha + \beta_{Mkt}\,(R_{Mkt} - R_f) \tag{1}$$
$$(R_{Risk\ Arb} - R_f) = \alpha + \beta_{Mkt}\,(R_{Mkt} - R_f) + \beta_{SMB}SMB + \beta_{HML}HML,$$

where $R_{Risk\ Arb}$ is the monthly return to a portfolio of risk arbitrage investments, R_f is the risk-free rate, R_{Mkt} is the return to the value-weighted CRSP index, *SMB* is the difference in returns between a portfolio of small stocks and a portfolio of big stocks, and *HML* is the difference in returns between a portfolio of high book-to-market stocks and a portfolio of low book-to-market stocks. The intercept, α, measures the average monthly abnormal return to the risk arbitrage portfolio, which is zero under the null of market efficiency, given the model. If the estimated α is significantly positive, this suggests that the risk arbitrageur earns excess returns, assuming that the model is correct.

The second analysis consists of estimating the following piecewise linear CAPM-type model:

$$R_{Risk\ Arb} - R_f = (1-\delta)\,[\alpha_{Mkt\ Low} + \beta_{Mkt\ Low}\,(R_{Mkt} - R_f)] \tag{2}$$
$$+\,\delta\,[\alpha_{Mkt\ High} + \beta_{Mkt\ High}\,(R_{Mkt} - R_f)],$$

where δ is a dummy variable if the excess return on the value-weighted CRSP index is above a threshold level and zero otherwise. To ensure continuity, we impose the following restriction on the model:

$$\alpha_{Mkt\ Low} + \beta_{Mkt\ Low}(Threshold) = \alpha_{Mkt\ High} + \beta_{Mkt\ High}(Threshold). \tag{3}$$

If risk arbitrage is akin to writing uncovered index put options, we should observe an optionlike feature in risk arbitrage returns. During flat and appreciating markets, $\alpha_{Mkt\ High}$ estimated from the above regression should be positive (the put premium) and the estimate of $\beta_{Mkt\ High}$ should be close to zero. However, during market downturns, risk arbitrage returns should be negative, implying that $\beta_{Mkt\ Low}$ should be greater than zero.

Figure 2. This figure depicts the piecewise linear model specified in equations (2) and (3). $R_{Risk\ Arb}$ is the monthly return obtained from the risk arbitrage portfolio, R_f is the monthly risk-free rate, and R_{Mkt} is the monthly return obtained from the CRSP value-weighted index. The market beta is allowed to vary depending on market returns. $\beta_{Mkt\ Low}$ is the slope coefficient when the difference between the market return and the risk-free rate is less than the threshold. $\beta_{Mkt\ High}$ is the slope coefficient when the difference between the market return and the risk-free rate is greater than the threshold.

Figure 2 provides a graphical depiction of the model specified by equations (2) and (3), assuming a negative threshold.

III. Data Description

Unlike many previous studies that focus on specific types of transactions such as cash tenders, we study arbitrage returns to cash tenders, cash mergers, and stock swap mergers. There are two advantages to including multiple types of mergers in the sample. First, it allows us to simulate a realistic investment strategy that is similar to strategies pursued by risk arbitrage hedge funds. To keep investors' money employed, these hedge funds typically invest in a broad range of merger situations, not just cash deals. Second, it provides a sample that is large enough to study the time-series characteristics of risk arbitrage returns, especially returns realized during severe market downturns. This is necessary to accurately measure the systematic risk inherent in risk arbitrage.

The data set for this study includes all CRSP firms that were delisted during the period 1963 to 1998 because of a merger or acquisition, and also includes all CRSP firms that received unsuccessful merger and acquisition bids that were covered by the Dow Jones News Service or the *Wall Street Journal*. Critical transaction information such as announcement dates, preliminary agreement dates, termination dates, entry of a second

bidder, and transaction terms was obtained by reading Dow Jones News Service and *Wall Street Journal* articles relating to each merger transaction. The final sample consists of 9,026 transactions.

Of these 9,026 transactions, we exclude 4,276 transactions. There are two reasons for dropping observations. First, many of the 9,026 transactions contain complicated terms. For example, the merger agreement might call for the target's shareholders to exchange their shares for a combination of cash, preferred stock, and warrants. Determining the value of the "hedge" in such a transaction is not possible since market values of hybrid securities are generally unavailable. Because our goal is to simulate the returns to a diversified risk arbitrage portfolio, we limit the sample to those transactions where the arbitrageur's investment is straightforward. The resulting sample includes cash mergers, cash tenders, and simple stock swap transactions.[3] The second reason for omitting transactions is because of lack of accurate data. In many cases, the exact terms of the transaction cannot be determined from our reading of the Dow Jones News Service and *Wall Street Journal* articles. In other cases, the terms reported in the *Wall Street Journal* or Dow Jones News Service imply wildly unrealistic returns.

Table I contains a summary of the 4,750 mergers used in this study, broken down by announcement year and transaction type. The sample contains relatively few mergers in the

Table I Sample Summary

This table includes a summary of the mergers used in this paper. Only those mergers that used 100 percent cash or 100 percent stock are included in the sample. Transactions that used a combination of securities (e.g., cash plus warrants, preferred plus common stock) are omitted. Transaction duration is measured as the number of trading days from the date of the merger announcement to the date that the merger is either consummated or canceled. Target and acquirer equity market values are measured on the day after the merger announcement. Standard deviations are in parentheses.

Year	Number of Mergers Announced	Number of Cash Transactions as Percent of Total	Average Transaction Duration		Average Target Market Equity Value ($ Millions)		Average Acquirer Market Equity Value ($ Millions)	
1963	30	47%	70	(48)	55.9	(52.2)	585.9	(1,091.5)
1964	25	52%	58	(37)	80.2	(147.8)	357.5	(509.5)
1965	29	72%	53	(45)	66.1	(113.6)	279.7	(527.3)
1966	31	55%	70	(89)	88.2	(86.8)	583.7	(856.3)
1967	40	50%	54	(48)	132.4	(168.9)	466.4	(559.8)
1968	58	40%	79	(197)	147.2	(272.6)	426.6	(574.9)
1969	31	26%	89	(86)	107.1	(163.4)	563.8	(1,116.6)
1970	32	22%	70	(35)	86.3	(116.7)	581.9	(948.2)
1971	24	38%	65	(48)	111.1	(108.9)	725.2	(1,128.3)
1972	28	32%	94	(131)	82.2	(99.7)	853.3	(1,092.1)
1973	89	57%	69	(71)	44.9	(48.6)	374.5	(643.1)
1974	99	68%	50	(66)	58.2	(93.3)	411.5	(716.1)
1975	82	61%	65	(66)	71.2	(178.2)	513.7	(1,227.8)

(*Continued*)

Table I (***Continued***)

Year	Number of Mergers Announced	Number of Cash Transactions as Percent of Total	Average Transaction Duration		Average Target Market Equity Value ($ Millions)		Average Acquirer Market Equity Value ($ Millions)	
1976	110	54%	63	(49)	60.6	(134.8)	679.3	(1,367.8)
1977	182	76%	59	(55)	83.7	(142.2)	409.9	(824.5)
1978	191	82%	63	(65)	81.3	(132.5)	473.4	(1,228.1)
1979	214	88%	59	(62)	98.3	(144.4)	501.9	(1,970.5)
1980	158	82%	68	(64)	143.0	(222.5)	1,047.8	(3,702.8)
1981	151	83%	56	(56)	513.8	(1,676.2)	713.9	(2,164.4)
1982	147	86%	56	(47)	167.0	(413.8)	383.9	(757.9)
1983	168	82%	66	(67)	151.2	(248.4)	450.2	(991.1)
1984	249	90%	48	(45)	529.4	(2,006.4)	494.9	(1,363.4)
1985	221	89%	59	(58)	596.9	(1,600.8)	1,071.9	(2,955.0)
1986	333	89%	50	(68)	354.5	(703.5)	755.4	(1,798.5)
1987	306	86%	50	(54)	404.9	(1,157.5)	1,199.3	(4,327.2)
1988	428	94%	40	(44)	696.6	(1,741.1)	900.5	(3,651.4)
1989	284	90%	50	(49)	543.3	(1,299.7)	1,031.6	(3,389.4)
1990	115	82%	66	(69)	349.0	(860.4)	1,888.7	(5,285.4)
1991	91	45%	87	(63)	437.7	(966.1)	2,728.8	(5,947.3)
1992	80	41%	101	(65)	261.2	(414.7)	2,000.0	(2,902.2)
1993	89	49%	94	(58)	335.9	(859.8)	1,831.1	(2,047.7)
1994	119	49%	80	(64)	521.5	(1,380.6)	3,116.9	(5,257.6)
1995	157	58%	70	(55)	734.6	(1,299.9)	5,139.3	(13,900.0)
1996	129	46%	65	(50)	808.7	(2,653.0)	7,278.9	(20,000.0)
1997	114	62%	52	(40)	801.4	(1,813.5)	5,696.1	(16,800.0)
1998	116	58%	47	(41)	1,175.8	(2,087.4)	9,504.7	(31,000.0)
Complete Sample	4,750	73%	59.3	(62.4)	390.7	(1,236.1)	1,548.7	(7,627.7)

1960s; however, the number of mergers increases substantially beginning in the late 1970s. The percentage of transactions that use cash as the medium of exchange also increases substantially in the late 1970s. There is no apparent pattern over time in the average duration of transactions. For the entire sample, the average time from bid announcement to transaction resolution is 59.3 trading days. However, for deals that ultimately fail, the average transaction time is 39.2 days, whereas it is 64.2 days for deals that ultimately succeed. A final feature worth noting is that target companies are significantly smaller than acquiring companies. Measured one day after the merger announcement, the average market equity value of target firms is $391 million and the average market equity value of acquiring firms is $1.55 billion.

IV. Risk Arbitrage Return Series

The analyses reported in this paper are based on monthly risk arbitrage returns. Monthly returns are obtained by compounding daily returns using two approaches, each of which is described below. In both approaches, we begin by calculating daily returns at the close of market on the day after the merger announcement. Daily returns are calculated for every transaction-day up to and including the "resolution day." For successful deals, the resolution day is defined to be the day on which the target's stock is delisted from CRSP. For failed deals, the resolution day is the day after deal failure is publicly announced. Using the day after the announcement as the beginning date ensures that arbitrage returns are not inadvertently biased upward by the takeover premium. Similarly, using the day after deal failure is announced as the resolution date for failed transactions ensures that the arbitrage returns are not biased upward by inadvertently exiting failed deals before the failure is announced.

Transactions in which the terms of the deal are revised before deal consummation are treated as multiple transactions. An investment in the transaction under the original terms is made at the close of market on the day following the announcement. This position is closed at the close of market on the day following the announcement of the bid revision. At the same time, an investment is made in the revised transaction and is held until the transaction resolution date. Transactions in which there are multiple bidders are handled in a similar manner. That is, one target can generate multiple transactions. Positions in a given transaction are held until the bidder announces that it is terminating its pursuit of the target, or when the target is delisted from CRSP, whichever occurs earlier.

For transactions where cash is used as the method of payment, the following equation is used to calculate daily returns:

$$R_{it} = \frac{P_{it}^{T} + D_{it}^{T} - P_{it-1}^{T}}{P_{it-1}^{T}}, \tag{4}$$

where R_{it} is the daily return, P_{it}^{T} is the target's stock price at the close of the market on day t, D_{it}^{T} is the dividend paid by the target on day t, and P_{it-1}^{T} is the target's closing stock price on day $t-1$ (subscript i refers to transaction number, t refers to transaction time in days, and T refers to "target").

Because the risk arbitrage position for stock deals consists of a long position in the target and a short position in the acquirer, calculating daily returns is more complicated for stock deals than for cash deals. The return for stock deals consists of the sum of the returns from the long position in the target's stock and the short position in the acquirer's stock. In addition to appreciation (or depreciation) of the stock prices and dividends for both the target and the acquirer, the interest earned on short proceeds must be accounted for.[4] To determine the daily return, the change in the value of the position in a particular deal is divided by the position value on the previous day. The calculation below assumes that short proceeds earn interest at the risk-free rate.

$$R_{it} = \frac{P_{it}^{T} + D_{it}^{T} - P_{it-1}^{T} - \Delta(P_{it}^{A} + D_{it}^{A} - P_{it-1}^{A} - r_f P_{i1}^{A}}{Position\ Value_{t-1}}, \tag{5}$$

where superscript T refers to the target, superscript A refers to the acquirer, Δ is the hedge ratio (equal to the number of acquirer shares to be paid for each outstanding target share), r_f is the daily risk-free rate, and P_{i1}^{A} is the acquirer's stock price at the close of market on the day following the merger announcement.

Monthly return time series are calculated from daily returns using two different methodologies, described in detail below. The first method is similar to that used in studies that use a calendar-time (not event-time) approach (e.g., Baker and Savasoglu (2002)). It consists of the average return across all merger deals at a given point in time, but ignores transaction costs and other practical aspects associated with risk arbitrage investments (VWRA returns). The second approach generates the return time series from a hypothetical risk arbitrage index manager (RAIM returns). Because they include transaction costs, and because capital is invested in cash when there is not enough merger activity to employ the simulated fund's capital, RAIM returns are lower than VWRA returns.

A. Value-weighted Average Return Series (VWRA)

For every active transaction month in the sample period, monthly returns are calculated by compounding daily returns. An active transaction month is defined for every transaction to be any month that contains a trading day between the transaction's beginning date and its resolution date (defined above). If a transaction is active for only part of a month, the partial-month return is used. This effectively assumes that capital is invested in a zero-return account for that portion of the month that the transaction is not active.[5] Portfolio monthly returns are obtained by calculating a weighted average of transaction-month returns for each month, where the total market equity value of the target company is used as the weighting factor. This approach mitigates the bias that is induced by calculating monthly returns by compounding equal-weighted daily returns (Canina et al. (1998)). The equation below specifies the monthly return calculation procedure:

$$R_{month\,j} = \sum_{i=1}^{N_j} \frac{V_i \left[\prod_{t=m}^{M} (1 + R_{it} - 1) \right]}{\sum_{i=1}^{N_j} V_i}, \tag{6}$$

where j indexes months between 1963 and 1998, i indexes active deals in a month (there are N_j active deals in month j), and t indexes trading days in a transaction month. Because the target's market equity is used as the weighting factor, a greater proportion of the portfolio is invested in larger, and presumably more liquid, targets. However, this approach in no way controls for illiquidity in the acquirer's stock. Thus, returns calculated using the weighted averaging procedure may be unrealistic in that they assume that there is an ample supply of the acquirer's stock available to be shorted. Of course, this is only a problem with stock-for-stock mergers where the acquirer's stock is difficult to borrow. In cash tenders and mergers, the typical risk arbitrage investment does not involve trading in the acquiring firm's equity, and therefore, the liquidity of the acquirer's stock is inconsequential.

There are two other features of the VWRA approach that are worth noting. First, this method effectively assumes that the arbitrage portfolio is invested in every transaction. Because of the fixed costs associated with investing in a transaction, this is a feature that large risk arbitrage hedge funds are unable to implement. Second, it assumes that there are no transaction costs associated with investing in a transaction.[6] Both of these assumptions are clearly unrealistic. However, the time series of returns generated from this approach

provide a benchmark that is useful for comparing results from this study to those documented in other papers.

B. Risk Arbitrage Index Manager Returns (RAIM)

The second time series of risk arbitrage returns used in this paper attempts to correct for the unrealistic assumptions embedded in the first method by simulating a risk arbitrage portfolio. Note that in this portfolio, the hypothetical arbitrageur does not attempt to discriminate between anticipated successful and unsuccessful deals. To generate this time series of returns, the portfolio is seeded with $1 million of capital at the beginning of 1963.[7] As mergers are announced, the $1 million is invested subject to two constraints. The first constraint is that no investment can represent more than 10 percent of the total portfolio's value at the time the investment is made. This is a standard rule of thumb followed by risk arbitrage hedge funds and is intended to insulate the fund from a catastrophic loss caused by failure of a single deal. The second constraint limits the fund's investments in illiquid securities. It does this by restricting the amount invested in any single deal such that the price impact on both the target and acquirer's stock is less than 5 percent. To implement this constraint, the following price impact model developed by Breen, Hodrick, and Korajczyk (1999, Equation 1) is used:

$$\frac{\Delta P}{P} = \beta(NTO), \tag{7}$$

where $\Delta P/P$ is the price impact equal to the percentage change in price resulting from a trade with net turnover equal to *NTO*. Net turnover is defined as one-tenth of the buyer initiated volume minus seller-initiated volume divided by shares outstanding. β is the illiquidity coefficient, obtained by calculating predicted values using regression results presented in Table 5 of Breen et al. A detailed description of their procedure is provided in the Appendix of this paper. It is, however, worth noting here that their results may not accurately reflect the true costs of trading over the time period studied in this paper. Breen et al. use the period from January 1993 through May 1997 to estimate their price impact model. To the degree that financial markets have become more liquid over time, their results may understate the true price impact of trading in earlier time periods. Their results also focus on "typical" event periods, not merger situations. If merger events substantially increase or decrease the price impact associated with trading merger stocks, using the Breen et al. results will, respectively, understate or overstate the true price impact. With these caveats in mind, we use their analysis both to restrict position sizes in illiquid securities and to calculate transaction costs associated with price impact. To calculate the allowable size of every investment, we invert equation (7) and perform the following calculation for both the target and the acquirer:

$$\text{Maximum Number of Shares} = N = \frac{\Delta P}{P\beta}(10)(\textit{Shares Outstanding}), \tag{8}$$

where price impact, $\Delta P/P$, is set equal to 5 percent and β equals the predicted value from the Breen et al. model. To determine the size of an investment, the most restrictive stock (e.g., target or acquirer) is used as long as the resulting position is less than 10 percent of the simulated fund's total capital. If both the target's stock and the acquirer's stock are

extremely liquid, the 10 percent diversification constraint binds. In this case, as long as the simulated fund has sufficient cash, it invests 10 percent of total capital in the deal.

In addition to limiting the magnitudes of investments, the cost associated with the price impact (which we refer to as indirect transaction costs) predicted by the Breen et al. (1999) model are subtracted from total capital. However, because the total cost can be reduced by splitting an order for N shares into n_{trade} transactions, the following cost model is used:

$$\text{Indirect Transaction Cost} = \frac{(N)(\Delta P)}{n_{trade}} = \frac{N^2 P\beta}{(10)(Shares\ Outstanding)(n_{trade})}, \tag{9}$$

where N is the total number of shares traded, P is the stock price, and n_{trade} is the number of individual trades used to trade N shares of stock. Based on conversations with practicing risk arbitrageurs, we use n_{trade} equal to 10 as an estimate of the typical number of trades used to make an investment.

In addition to indirect transaction costs associated with price impact, we also model direct transaction costs consisting of brokerage fees, transaction taxes, and other surcharges. Prior to 1975, direct trading costs, which were regulated by the NYSE and enforced by the SEC, were substantial. They are described in detail in the Appendix. Because risk arbitrage requires frequent trading, these fees turn out to be important components of risk arbitrage returns. Based on conversations with investment professionals that traded in the mid-1970s, brokerage fees dropped substantially after deregulation and continue to drop, albeit at a decreasing rate. Jarrell (1984) estimates that, for institutions, per share direct transaction costs decreased by 50 percent between 1975 and 1980. Because of the relatively high turnover associated with risk arbitrage investments, risk arbitrageurs probably experienced even more substantial reductions in trading costs. To calculate returns for the index portfolio after 1975, we assume per-share transaction costs (outlined in Table AII of the Appendix) that decrease to \$0.10 per share between 1975 and 1979, to \$0.05 per share between 1980 and 1989, and to \$0.04 per share between 1990 and 1998.

Table II presents the annualized time series of monthly returns for both the VWRA and RAIM portfolios. As expected, the VWRA portfolio significantly outperforms the RAIM portfolio. Whereas the VWRA portfolio generates a compound annual return of 16.05 percent, the RAIM portfolio generates a compound annual return of 10.64 percent. Of the 5.41 percent difference, approximately 1.5 percent can be attributed to direct transaction costs (e.g., brokerage commissions, surcharges, and taxes), and 1.5 percent can be attributed to indirect transaction costs (e.g., price impact). The remaining 2.5 percent can be attributed to limitations in position sizes caused by illiquidity in the merging firms' stocks. Thus, ignoring transaction costs and the price impact associated with investing in thinly traded equities imposes a substantial upward bias to calculated returns.[8]

Also shown in Table II are the annualized CRSP value-weighted average return and the risk-free rate of return. Over the 1963–1998 time period, the CRSP value-weighted index had a compound annual return of 12.24 percent, almost 400 basis points *less* than the VWRA average and only 160 basis points greater than the RAIM average. Annual standard deviations and Sharpe ratios are also presented in Table II. Even though the compound annual return of the RAIM portfolio is lower than the market return, the low volatility associated with risk arbitrage returns results in a Sharpe ratio that exceeds that of the market.

Table II Annual Risk Arbitrage Return Series

This table presents the annual return series for the value-weighted risk arbitrage (VWRA) portfolio, the risk arbitrage index manager (RAIM) portfolio, the annual CRSP value-weighted index, and the annual risk-free rate. VWRA portfolio returns are obtained by taking the weighted average of returns from all active merger deals, ignoring transaction costs. RAIM returns include transaction costs and other practical limitations associated with risk arbitrage investments. The ratio of the sum of target firms' equity values and the end-of-year total market value is also presented. All annual returns are obtained by compounding monthly returns. Annual standard deviations are obtained by multiplying the standard deviation of monthly returns by the square root of 12.

Year	Value-weighted Risk Arbitrage (VWRA) Return	Risk Arbitrage Index Manager (RAIM) Return	CRSP Value-weighted Average Return	Risk-free Rate of Return	$ Value of Announced Deals/Total Market Value
1963	14.51%	6.64%	20.89%	3.13%	0.40%
1964	10.27%	4.44%	16.30%	3.48%	0.35%
1965	9.09%	3.30%	14.38%	3.94%	0.47%
1966	11.46%	–4.03%	–8.68%	4.69%	0.69%
1967	14.45%	9.06%	28.56%	4.05%	1.16%
1968	–8.65%	–2.88%	14.17%	4.75%	1.72%
1969	22.10%	3.18%	–10.84%	6.49%	1.10%
1970	14.18%	5.70%	0.08%	6.17%	0.30%
1971	19.93%	5.79%	16.20%	4.15%	0.15%
1972	16.65%	3.52%	17.34%	3.93%	0.13%
1973	20.38%	–7.45%	–18.77%	7.17%	0.39%
1974	12.95%	12.93%	–27.86%	7.97%	0.42%
1975	12.83%	12.29%	37.37%	5.63%	0.29%
1976	19.93%	19.20%	26.77%	4.91%	0.36%
1977	28.56%	8.27%	–2.98%	5.25%	0.72%
1978	20.40%	18.03%	8.54%	7.41%	0.93%
1979	17.15%	13.85%	24.40%	10.42%	0.82%
1980	29.30%	38.54%	33.23%	11.33%	0.47%
1981	38.44%	35.15%	–3.97%	14.50%	0.68%
1982	38.41%	31.99%	20.42%	10.38%	0.42%
1983	17.35%	12.67%	22.70%	8.86%	0.45%
1984	21.45%	8.13%	3.28%	9.62%	0.63%
1985	15.65%	15.00%	31.46%	7.38%	0.50%
1986	13.32%	20.61%	15.60%	5.93%	0.68%
1987	13.81%	3.81%	1.76%	5.17%	0.63%
1988	27.23%	27.63%	17.62%	6.50%	0.61%
1989	6.83%	5.36%	28.44%	8.16%	0.32%
1990	6.69%	4.38%	–6.02%	7.53%	0.11%

(*Continued*)

Table II (*Continued*)

Year	Value-weighted Risk Arbitrage (VWRA) Return	Risk Arbitrage Index Manager (RAIM) Return	CRSP Value-weighted Average Return	Risk-free Rate of Return	$ Value of Announced Deals/Total Market Value
1991	18.19%	12.13%	33.59%	5.32%	0.07%
1992	9.12%	4.48%	9.03%	3.36%	0.07%
1993	14.16%	12.31%	11.49%	2.90%	0.09%
1994	17.07%	12.58%	–0.62%	3.98%	0.12%
1995	12.57%	10.96%	35.73%	5.47%	0.11%
1996	11.32%	15.39%	21.26%	5.14%	0.06%
1997	9.48%	11.64%	30.46%	5.11%	0.06%
1998	12.64%	4.09%	22.49%	4.70%	0.06%
Compound annual rate of return	16.05%	10.64%	12.24%	6.22%	
Annual standard deviation of returns	9.29%	7.74%	15.08%	0.73%	
Sharpe ratio (annual)	1.06	0.57	0.40	0.0	

Figure 3. This figure shows the value, over the 1963 to 1998 time period, of $1 invested at the beginning of 1963 for four different investments: (1) value-weighted risk arbitrage (VWRA), (2) value weighted CRSP index, (3) risk arbitrage index manager (RAIM), (4) Treasury bills. Because of transaction costs and other practical issues, the VWRA returns would not have been obtainable; they are included for comparison purposes. The RAIM returns take transaction costs and other practical issues into account and are representative of the returns that could have been obtained from an index of merger arbitrage investments. The horizontal axis labels correspond to months (i.e., 9812 is December, 1998).

Returns summarized in Table II are shown graphically in Figure 3. This figure shows the value of $1 invested at the beginning of 1963 in various strategies, including treasuries, equities, and risk arbitrage. The effect of ignoring transaction costs on risk arbitrage returns is obvious by comparing the returns from the VWRA portfolio to the returns from the RAIM port folio. It is also evident from Figure 3 that returns to risk arbitrage are much less volatile than market returns.

V. Results

To determine whether the returns to risk arbitrage reflect market inefficiencies or rewards for bearing rare-event risk, we estimate equations (1) through (3) over the 1963 to 1998 time period. Because the RAIM portfolio return series is more realistic, we focus our discussion on results obtained using this return series as the dependent variable.

A. Risk Factors

Panel A of Table III presents results for the entire 432-month (36-year) sample. The first regression presents results from estimating the CAPM. Results from this regression indicate that the alpha is positive 29 basis points per month and is significantly different from zero. Furthermore, the estimated market beta is only 0.12. This result indicates that over a broad range of market environments, risk arbitrage returns are independent of overall market returns.

Similar results are obtained when the Fama and French (1993) three-factor model is used. The alpha is 27 basis points per month and the market beta is 0.11, both significantly different from zero. The SMB coefficient is also statistically different from zero in this regression. Because the arbitrage trade in a stock transaction consists of a long position in a relatively small target and a short position in a relatively large acquirer, the correlation between RAIM returns and SMB is not surprising.

Panels B and C of Table III report results from estimating equation (1) after limiting the sample to months where the market return minus the risk-free rate is less than −3.0 percent and −5.0 percent, respectively. The estimated alphas using these subsamples of data increase dramatically. As shown in Panel B, when the excess market return is the only independent variable, the estimated alpha is 260 basis points per month (36.1 percent annualized) and the beta is 0.51. The adjusted R^2 increases dramatically when the sample is limited to months with negative market returns (from 0.057 to 0.306) suggesting that the systematic risk in risk arbitrage is driven by time periods where market returns are negative. Including the Fama–French factors reduces the alpha to 206 basis points per month (27.72 percent annualized), but it is still statistically different from zero at the 1 percent level.

The coefficient estimates in Table III suggest that the relationship between risk arbitrage returns and market returns is nonlinear. To further assess the degree of nonlinearity in risk arbitrage returns, we estimate the piecewise linear model specified in equations (2) and (3) and depicted in Figure 2.[9] The piecewise linear analysis is performed only for the market model; nonlinearities associated with the Fama and French (1993) SMB and HML factors are not assessed.

Table III Time Series Regressions of Risk Arbitrage Returns on Common Risk Factors

This table presents results from the following two regressions of risk arbitrage returns on common risk factors:

$$R_{Risk\ Arb} - R_f = \alpha + \beta_{Mkt}(R_{Mkt} - R_f)$$
$$R_{Risk\ Arb} - R_f = \alpha + \beta_{Mkt}(R_{Mkt} - R_f) + \beta_{SMB} R_{SMB} + \beta_{HML} R_{HML},$$

where $R_{Risk\ Arb}$ is the monthly return on a portfolio of risk arbitrage transactions, R_f is the monthly risk-free rate, R_{Mkt} is the monthly return on the value-weighted CRSP index, R_{SMB} is the Fama–French small minus big monthly return series, and R_{HML} is the Fama–French high book-to-market minus low book-to-market return series. Two different time series of risk arbitrage returns are used. The first is based on a risk arbitrage index manager (RAIM) portfolio beginning in 1963 and ending in 1998. This return series is net of transaction costs. The second, which ignores transaction costs, is the value weighted average of returns to individual merger investments (VWRA), averaged across transactions in each month. The target firm's market capitalization is used as the weighting factor. Panel A of the table presents results for the entire time period. Panel B presents results after restricting the sample to those months with market returns more than three percent less than the risk-free rate. Panel C presents results after restricting the sample to those months with market returns more than five percent less than the risk-free rate. Standard errors are in parentheses.

Dependent Variable	α	β_{Mkt}	β_{SMB}	β_{HML}	Adj. R^2	Sample Size
Panel A: Complete Sample						
RAIM portfolio returns	0.0029	0.1232			0.057	432
	(0.0010)**	(0.0236)***				
RAIM portfolio returns	0.0027	0.1052	0.1221	0.0357	0.076	432
	(0.0011)*	(0.0265)***	(0.0380)**	(0.0434)		
VWRA portfolio returns	0.0074	0.0540			0.006	432
	(0.0013)***	(0.0293)				
VWRA portfolio returns	0.0079	0.0176	0.0774	–0.0904	0.014	432
	(0.0013)***	(0.0331)	(0.0475)	(0.0542)		
Panel B: Market Return $R_f < -3\%$						
RAIM portfolio returns	0.0260	0.5074			0.306	76
	(0.0059)***	(0.0869)***				
RAIM portfolio returns	0.0206	0.4041	0.2996	0.1824	0.396	76
	(0.0058)***	(0.1035)***	(0.1063)**	(0.1258)		
VWRA portfolio returns	0.0368	0.5194			0.219	76
	(0.0076)***	(0.1107)***				
VWRA portfolio returns	0.0356	0.5532	0.0219	0.1994	0.214	76
	(0.0079)***	(0.1417)***	(0.1456)	(0.1723)		

(*Continued*)

Table III (*Continued*)

Dependent Variable	α	β_{Mkt}	β_{SMB}	β_{HML}	Adj. R^2	Sample Size
			Panel C: Market Return $R_f < -5\%$			
RAIM portfolio returns	0.0232 (0.0134)	0.4830 (0.1479)**			0.222	35
RAIM portfolio returns	0.0116 (0.0120)	0.2884 (0.1588)*	0.4761 (0.1722)**	0.2774 (0.2035)	0.424	35
VWRA portfolio returns	0.0354 (0.0132)	0.5103 (0.1450)			0.251	35
VWRA portfolio returns	0.0298 (0.0137)	0.5000 (0.1804)	0.0934 (0.1956)	0.2735 (0.2311)	0.257	35

*, **, *** indicate significance at the 0.05, 0.01, and 0.001 levels, respectively.

One problem with implementing the piecewise linear model is determining the location of the threshold (i.e., the kink point). To avoid using a completely ad hoc method of determining the threshold, we present results obtained by setting the threshold equal to –4.0 percent, the value that minimizes the sum of squared residuals.

Panel A of Table IV presents results using the complete sample covering the 1963–1998 time period. Results from this panel indicate that in most market environments, risk arbitrage produces a return that is 53 basis points per month (6.5 percent annually) greater than the risk-free rate and a beta that is close to zero. However, when the market return is more than 4 percent below the risk-free rate, the risk arbitrage market beta increases to 0.49.

Panels B, C, and D of Table IV show that in all subperiods, market betas are significantly different in up and down markets. Furthermore, except for the 1963 to 1979 time period when merger activity was relatively low, the intercept terms are large and significantly different from zero. Note however, that because of the nonlinear relationship between risk arbitrage returns and market returns, these intercepts cannot be interpreted as excess returns. Scatter plots of RAIM returns versus market returns for various subperiods are shown in Figure 4. This figure shows that the nonlinear relationship between risk arbitrage returns and market returns is not time-period dependent.

The increase in market beta in depreciating markets is caused, at least in part, by the increased probability of deal failure following a severe market downturn. Table V shows results from a probit regression that estimates the probability of deal failure. For purposes of this analysis, deal failure is defined to be any deal where the arbitrageur lost money. Thus, mergers where the terms were revised downward but that were ultimately consummated are treated as failed deals. As shown in this table, the probability that a merger will fail is a decreasing function of market returns in the previous two months. That is, deals are more likely to fail following market downturns. Based on the coefficient estimates in Table V, a 5 percent decrease in either the contemporaneous market return or the lagged market return increases the probability of deal failure by 2.25 percent. Table V also shows that hostile deals have a 12.8 percent greater probability of failure than friendly deals. In our data set, "hostile" refers to deals in which articles in the Dow Jones News Service or *Wall Street Journal* report that target management rejected the bid in question.[10] Leveraged buyouts also have higher failure probabilities.[11]

Table IV Piecewise Linear Regressions: Risk Arbitrage Returns Versus Market Returns

This table presents results from the following piecewise linear regression relating risk arbitrage returns to market returns:

$$R_{Risk\ Arb} - R_f = (1-\delta)[\alpha_{Mkt\ Low} + \beta_{Mkt\ Low}(R_{Mkt} - R_f)] + \delta[\alpha_{Mkt\ High} - \beta_{Mkt\ High}(R_{Mkt} - R_f)],$$

where $R_{Risk\ Arb}$ is the monthly return on a portfolio of risk arbitrage transactions, R_f is the risk-free rate, R_{Mkt} is the monthly return on the value-weighted CRSP index, and δ is a dummy variable equal to one if the market return is greater than a threshold and zero otherwise. To ensure continuity, the following restriction is imposed:

$$\alpha_{Mkt\ Low} + \beta_{Mkt\ Low}(Threshold) = \alpha_{Mkt\ High} + \beta_{Mkt\ High}(Threshold).$$

Results are presented for a threshold equal to –4 percent, that being the threshold that maximizes the adjusted R^2 for the complete sample. Panel A presents results using the entire 432-month sample between 1963 and 1998. Panels B, C, and D present results for various subperiods. Standard errors are in parentheses.

Dependent Variable	$\alpha_{Mkt\ High}$	$\beta_{Mkt\ Low}$	$\beta_{Mkt\ High}$	Adj. R^2	Sample Size
		Panel A: Complete Sample			
RAIM portfolio returns	0.0053 (0.0011)***	0.4920 (0.0673)***	0.0167 (0.0292)	0.124	432
VWRA portfolio returns	0.0101 (0.0013)***	0.4757 (0.0840)***	–0.0678 (0.0364)	0.065	432
		Panel B: 1963–1979			
RAIM portfolio returns	0.0025 (0.0016)	0.3849 (0.1175)***	–0.0206 (0.0435)	0.044	204
		Panel C: 1980–1989			
RAIM portfolio returns	0.0095 (0.0024)***	0.5825 (0.1115)***	0.0987 (0.0589)	0.233	120
		Panel D: 1990–1998			
RAIM portfolio returns	0.0054 (0.0016)***	0.4685 (0.1134)***	–0.0287 (0.0467)	0.127	108

*** indicates significance at the 0.001 level.

As previously described, capturing the arbitrage spread in a cash deal requires the arbitrageur to buy the target's stock. However, in a stock merger, capturing the arbitrage spread requires the arbitrageur to purchase the target's stock and simultaneously short sell the acquirer's stock. As long as the target's value and the acquirer's value are equally affected by the decrease in overall market value, the market decrease will not cause the acquirer to overpay for the target. However, in a cash deal, the decrease in the target's value is not offset by a commensurate decrease in the price paid by the acquirer. Thus, the arbitrageur is long market risk in cash deals and is market neutral in stock deals.

Figure 4. This figure plots risk arbitrage index manager (RAIM) returns against market returns for three subsamples of data. Panel A presents returns from 1975 to 1998, Panel B presents returns from 1975 to 1986, and Panel C presents returns from 1987 to 1998. Data labels correspond to months (i.e., 9808 is August, 1998). Fitted lines from a piecewise linear regression are also shown.

Table V Effect of Market Returns on the Probability of Deal Failure

This table presents results from the following probit model:

$$Fail = \alpha + \beta_1 R_{Mkt} + \beta_2 R_{Mkt-1} + \beta_3 R_{Mkt-2} + \beta_4 LBO + \beta_5 Cash\ Dummy + \beta_6 Premium + \beta_7 Size + \beta_8 Tender + \beta_9 Hostile,$$

where *Fail* is a dummy variable equal to one if the arbitrage return is negative and zero otherwise; R_{Mkt} is the monthly return on the value-weighted CRSP index for the month corresponding to the deal resolution date; R_{Mkt-1} is the monthly return on the value weighted CRSP index for the month prior to the deal resolution date; R_{Mkt-2} is the monthly return on the value weighted CRSP index two months prior to the deal resolution date; *LBO* is a dummy variable if the acquirer was private; *Cash Dummy* is a dummy variable if the acquirer offered to pay 100 percent cash for the target; *Premium* is the takeover premium equal to the target stock price one day after the announcement of the merger divided by the target stock price 30 days prior to the merger announcement; *Size* is the logarithm of the target's market equity value; *Tender* is a dummy variable equal to one if the offer was a cash tender; and *Hostile* is a dummy variable equal to one if articles in the Dow Jones News Service or *Wall Street Journal* report that target management rejected the bid in question. Standard errors (in parentheses) are calculated assuming independence across years. No assumptions are made regarding the independence of transactions that terminate in the same year.

Independent Variable	Coefficient Estimate	Marginal Effect
R_{Mkt}	–1.6481 (0.6493)**	–0.4444
R_{Mkt-1}	–1.7034 (0.3402)***	–0.4593
R_{Mkt-2}	–0.6164 (0.5649)	–0.1662
LBO	0.1748 (0.0656)**	0.0485
Cash dummy	0.1797 (0.0914)*	0.0465
Takeover premium	0.0086 (0.0405)	0.0023
Size	–0.0554 (0.0179)**	–0.0149
Tender dummy	–0.2360 (0.0796)**	0.0651
Hostile dummy	0.4221 (0.0629)***	0.1286
Constant	–0.5543 (0.1957)**	
R^2	0.040	
Number of observations	4,740	

*, **, *** indicate significance at the 0.05, 0.01, and 0.001 levels, respectively.

Table VI Piecewise Linear Regressions: Cash versus Stock Transactions

This table presents results from the following piecewise linear regression relating risk arbitrage returns to market returns:

$$R_{Risk\ Arb} - R_f = (1 - \delta[\alpha_{Mkt\ Low} + \beta_{MktLow}(R_{Mkt} - R_f)] + \delta[\alpha_{Mkt\ High} + \beta_{Mkt\ High}(R_{Mkt} - R_f)],$$

where $R_{Risk\ Arb}$ is the monthly return on a portfolio of risk arbitrage transactions, R_f is the risk-free rate, R_{Mkt} is the monthly return on the value-weighted CRSP index, and δ is a dummy variable equal to one if the market return is greater than a threshold and zero otherwise. To ensure continuity, the following restriction is imposed:

$$\alpha_{Mkt\ Low} + \beta_{Mkt\ Low}(Threshold) = \alpha_{Mkt\ High} + \beta_{Mkt\ High}(Threshold).$$

Results are presented for a threshold equal to –4 percent, that being the threshold that maximizes the adjusted R^2 for the complete sample. Panel A presents results obtained after restricting the sample to cash transactions; Panel B presents results for stock transactions. Standard errors are in parentheses.

Dependent Variable	$\alpha_{Mkt\ High}$	$\beta_{Mkt\ Low}$	$\beta_{Mkt\ High}$	Adj. R^2	Sample Size
		Panel A: Cash Transactions 1975–1998			
RAIM portfolio returns	0.0046 (0.0014)***	0.7745 (0.0822)***	0.1024 (0.0371)**	0.295	288
		Panel B: Stock Transactions 1975–1998			
RAIM portfolio returns	0.0051 (0.0008)***	0.1528 (0.0477)**	–0.0766 (0.0215)***	0.052	288

** and *** indicate significance at the 0.01 and 0.001 levels, respectively.

Given the increase in the probability of deal failure associated with cash deals and depreciating markets, the "down-market" beta in our piecewise linear regressions should be greater when the sample is limited to cash transactions. Table VI presents results from estimating the piecewise linear regressions after segmenting the data by means of payment. Panel A of Table VI presents results for cash transactions and Panel B presents results for stock transactions. This table confirms that the down-market beta is much greater when the sample is limited to cash transactions (0.77) than when it is limited to stock transactions (0.15).

B. Deal Flow

In addition to the systematic risk factors specified in equations (1) through (3), factors specific to the mergers and acquisitions market may also affect returns to risk arbitrage. In particular, institutional rigidities may restrict the flow of capital into risk arbitrage investments resulting in periodic imbalances between the supply of mergers and the demand for investments in merger stocks. The resulting imbalances would be greatest in periods when the volume of announced deals is high.

To determine whether the supply of transactions affects risk arbitrage returns, we constructed two variables that measure merger and acquisition activity. The first variable is

the number of announced transactions in the month; the second variable is the total market value of transactions (measured by target market value) announced in the month, divided by the total market value (NYSE, Nasdaq, AMEX). We included each of these variables in our piecewise linear regressions.

Results (available on request) from these regressions suggest that the link between risk arbitrage returns and merger activity is weak. Although there is a positive correlation between the number of mergers and risk arbitrage returns, the significance (both economic and statistical) of the relationship varies across time periods. The same is true of the relationship between risk arbitrage returns and the dollar volume of announced transactions. This lack of robustness leads us to conclude that "deal flow" is not a strong determinant of risk arbitrage returns.

C. Sensitivity Analysis

Calculating the RAIM returns used in this paper requires numerous assumptions regarding transaction costs and limitations associated with implementing the merger arbitrage strategy. It is possible that the results described thus far are an artifact of these assumptions. To test whether our assumptions are generating the nonlinear relationship between RAIM and market returns, we performed the analysis using alternative assumptions for diversification constraints, initial capital, and transaction costs. Results from these analyses are presented in Table VII. Scenarios 1 through 4 in Table VII present results for various levels of transaction costs. Comparing scenarios 1 and 2 indicates that direct transaction costs (e.g., brokerage commissions) decrease returns by approximately 1.37 percent annually. The effect of indirect transaction costs (price impact) can be estimated by comparing returns from scenarios 2 and 3. Scenario 2 includes indirect transaction costs estimated using the Breen et al. (1999) price impact model and scenario 3 assumes that indirect transaction costs are zero. Based on this comparison, indirect transaction costs decrease annual returns by 1.49 percent. If instead of eliminating indirect transaction costs we double them (scenario 4), returns are reduced by 2.51 percent. Comparing the 13.50 percent return in scenario 3 (no transaction costs) with the VWRA annual return of 16.05 percent reported in Table II (no transaction costs or practical limitations) indicates that practical limitations reduce annual returns by 2.5 percent per year.

Scenario 5 provides an estimate of the return generated by the interest paid on short proceeds. Whereas scenario 1 presents returns assuming that the risk-free rate of return is obtained on short proceeds, scenario 5 presents returns assuming that no interest is paid on short proceeds. The 2.02 percent difference in annual returns represents the portion of risk arbitrage returns that are generated by interest payments on short proceeds.

Results from scenarios 6 and 7 show the effect of altering the diversification constraint. It is common for merger arbitrage hedge funds to limit the maximum percentage of the portfolio's total value that can be invested in a single transaction. Results presented in Tables III, IV, and VI assume that the limit is 10 percent of the portfolio. Scenarios 6 and 7 allow this limit to vary from 20 percent down to 5 percent. Increasing the limit to 20 percent has a negligible effect on annual returns whereas decreasing it to 5 percent has a substantial effect; it decreases returns by 1.49 percent annually. Much of this decrease can

Table VII RAIM Sensitivity Analysis

This table presents results for alternative assumptions in the risk arbitrage index manager (RAIM) portfolio. Diversification constraint refers to the maximum percentage of the portfolio's total value that can be invested in a single transaction. Beginning capital is the amount of capital that the fund is seeded with at the beginning of 1963. Direct transaction costs are brokerage commissions and surcharges; indirect transaction costs are costs associated with price impact. Compounded annual returns for the 1963–1998 time period, and results from a piecewise linear regression of risk arbitrage returns on market returns are presented. The piecewise linear regression equation is

$$R_{Risk\ Arb} - R_f = (1-\delta)[\alpha_{Mkt\ Low} + \beta_{MktLow}(R_{Mkt} - R_f)] + \delta[\alpha_{Mkt\ High} + \beta_{Mkt\ High}(R_{Mkt} - R_f)],$$

where $R_{Risk\ Arb}$ is the monthly return on a portfolio of risk arbitrage transactions, R_f is the risk-free rate, R_{Mkt} is the monthly return on the value-weighted CRSP index, and δ is a dummy variable equal to one if the market return is greater than a threshold and zero otherwise. To ensure continuity, the following restriction is imposed:

$$\alpha_{Mkt\ Low} - \beta_{Mkt\ Low}(Threshold) = \alpha_{Mkt\ High} - \beta_{Mkt\ High}(Threshold).$$

Results are presented for a threshold equal to –4 percent, that being the threshold that maximizes the adjusted R^2 for the complete sample. Standard errors are in parentheses.

	Scenario							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Diversification constraint (%)	10	10	10	10	10	20	5	10
Beginning capital	$1 million	$1 million	$1 million	$1 million	$1 million	$1 million	$1 million	$10 million
Interest rate on short proceeds	Risk-free rate	Risk-free rate	Risk-free rate	Risk-free rate	Zero	Risk-free rate	Risk-free rate	Risk-free rate
Transaction costs	Direct and indirect	Indirect only	None	Direct plus 2 × indirect	Direct and indirect	Direct and indirect	Direct and indirect	Direct and indirect
Annual return, 1963–1998	10.64%	12.01%	13.50%	8.13%	8.62%	10.05%	9.15%	6.85%
$\alpha_{Mkt\ High}$	0.0053 (0.0011)***	0.0064 (0.0011)***	0.0076 (0.0011)***	0.0035 (0.0011)**	0.0036 (0.0011)**	0.0053 (0.0013)***	0.0043 (0.0009)***	0.0021 (0.0007)***
$\beta_{Mkt\ Low}$	0.4920 (0.0673)***	0.5108 (0.0673)***	0.5241 (0.0660)***	0.5057 (0.0697)***	0.5017 (0.0711)***	0.5823 (0.0789)***	0.5758 (0.0565)***	0.4342 (0.0453)***
$\beta_{Mkt\ High}$	0.0167 (0.0292)	0.0144 (0.0292)	0.0067 (0.0286)	0.0089 (0.0302)	0.0402 (0.0309)	0.0144 (0.0343)	0.0239 (0.0245)	0.0004 (0.0196)
Adj. R^2	0.124	0.132	0.140	0.120	0.126	0.125	0.125	0.190

** and *** indicate significance at the 0.01 and 0.001 levels, respectively.

be attributed to the lack of transactions in the 1960s and early 1970s. When the diversification constraint is very strict and there are few available deals, the RAIM portfolio is heavily invested in cash. This results in a decrease in overall returns. However, regardless of the level of the diversification constraint, the basic finding that betas are high in depreciating markets and close to zero in flat and appreciating markets remains.

The final analysis presented in Table VII involves the size of the initial capital base. Scenario 8 presents results obtained when initial capital is \$10 million instead of \$1 million. This causes annual returns to decrease by 3.79 percent. As is the case when the diversification constraint is tightened, much of this decrease is caused by the lack of merger activity in the 1960s and early 1970s. Nevertheless, this result suggests that the merger arbitrage strategy may be capacity constrained.

D. Contingent Claims Analysis Using Black–Scholes

Overall, results presented in Tables III through VII provide strong evidence supporting the notion that risk arbitrage is analogous to writing uncovered index put options. This suggests that standard measures of performance such as Jensen's alpha and the Sharpe ratio may not be appropriate for analyzing risk arbitrage returns. Rather than using a linear asset pricing model, the risk and reward associated with risk arbitrage would be better assessed using a contingent claims analysis. For example, the one-month return to a \$100 investment in a risk arbitrage portfolio can be replicated by a portfolio consisting of a long position in a risk-free bond and a short position in index put options. The face value of the bond is equal to (\$100)$(1 + r_f + \alpha_{Mkt\ High})$ and the number of put options is determined by the market beta in depreciating markets ($\beta_{Mkt\ Low}$ in Figure 2 and Table IV). The put option strike price is equal to (\$100)$(1 + Threshold + r_f)$. Thus, for a threshold of –4 percent and a risk free rate of 51 basis points per month (the sample average), the strike price is \$96.51.

To determine whether risk arbitrage generates excess returns, the cost of the replicating portfolio can be compared to the \$100 investment in risk arbitrage. If the cost of the replicating portfolio exceeds \$100, then risk arbitrage generates excess returns. Using coefficient estimates from Panel A of Table IV, and assuming Black–Scholes applies, the cost of the replicating portfolio is equal to the present value of the risk-free bond minus the put premium received from selling 0.492 index put options:

$$\text{Cost of Replicating Portfolio} = \frac{\$100 + \$0.51 + \$0.53}{1.0051} - (0.492)\, P(X, S, r_f, \sigma, T - t), \tag{10}$$

where $P(X, S, r_f, \sigma, T - t)$ is the Black–Scholes price of an index put option with a strike price of $X = \$96.51$, an index level of $S = \$100$, a risk-free rate of 6.3 percent (the sample average), a market volatility of 0.15 (standard deviation of monthly market returns multiplied by the square root of 12), and a time until expiration of one month. Using these parameter estimates implies that the put option is worth \$0.40 and the cost of the replicating portfolio is \$100.33, \$0.33 more expensive than the risk arbitrage portfolio. Thus, risk arbitrage generates excess returns of 33 basis points per month (4.0 percent annually) after controlling for transaction costs and other practical limitations.

Our analysis of risk and return in risk arbitrage uses a monthly time horizon. However, there is no reason, a priori, to base the analysis on monthly returns. In fact, annual returns shown in Table II suggest that risk arbitrage almost always generates positive returns when the horizon is one year. To determine whether the nonlinearity in returns exists when an annual horizon is used, the piecewise linear regression analysis was performed using annual returns. Results from this regression indicate that the market beta is 0.17 in both appreciating and depreciating markets. The implied excess return from this regression is 3.6

percent per year, very close to the estimate obtained using monthly returns. These results are consistent with the notion that the excess return in risk arbitrage reflects compensation for providing liquidity in merger stocks, especially during market downturns.

E. Contingent Claims Analysis Using Actual Put Prices

Jackwerth (2000) argues that a change in investors' risk aversion level after the October 1987 crash created the opportunity to profit from a trading strategy consisting of selling index put options. Because this increase in risk aversion does not enter the Black–Scholes formula, we also modeled the replicating portfolio over the 1987 to 1996 time period using actual S&P 500 index put option prices. In each month, we built a portfolio consisting of a risk-free bond with a face value of $\$100(1 + r_f)$ and a short position in index put options. To get option prices, we first calculated implied volatilities using prices from options that had one month until expiration and were approximately 4 percent out of the money. Option prices were adjusted to correct for this approximation by using implied volatilities from the actual option prices, together with Black–Scholes and the correct strike price (Strike = $\$100 - \$100(0.04 - r_f)$). At the end of each month, we calculated the payoff from our option position. This payoff, combined with the payoff from the risk-free bond is used to calculate portfolio monthly returns. The number of options sold was adjusted to mimic the risk arbitrage payoff profile.

Returns from this procedure are compared to returns from risk arbitrage over the same sample period (1987 to 1996). Results from this comparison indicate that risk arbitrage produces excess returns of approximately 29 basis points per month (3.5 percent annually). This estimate is lower than the estimate obtained using the Black–Scholes formula; the difference stems from the gap between actual market volatility and volatilities implied by index put option prices. Nevertheless, even when these higher volatilities are taken into consideration, risk arbitrage generates significant excess returns.

F. Contingent Claims Analysis versus CAPM

Because of the nonlinear relationship between risk arbitrage returns and market returns, linear asset pricing models are not appropriate for estimating excess returns associated with risk arbitrage. However, it would be interesting to know the magnitude of the error that one would make by incorrectly using CAPM. To estimate this error, we calculate the excess return using the contingent claims approach and CAPM for various subsamples of our data.

Results from our analysis suggest that, in general, CAPM provides an accurate assessment of excess returns. The largest differences between the CAPM-estimated excess return and the contingent-claims estimate occurs for subsamples with severe nonlinearities and large "up-market" intercepts. For example, when the sample is limited to cash deals in the 1980s, CAPM underestimates the excess return by 8 basis points per month (1.0 percent annually) relative to the contingent claims approach. Conversely, for subsamples where the relationship between risk arbitrage returns and market returns is closer to being linear, the difference in excess return estimates is small. When the sample is limited to stock transactions in the 1990s, CAPM overestimates the excess return by only 3 basis points per month (0.35 percent annually). This finding has implications for evaluating hedge-fund managers. Alphas estimated using linear asset pricing models will generate greater errors for fund

managers that accept greater risk in depreciating markets and generate large monthly "put premiums" in flat and appreciating markets.

To determine whether the large excess returns reported in previous studies result from inaccurate measures of risk, we performed our contingent claims analysis using VWRA returns. These returns assume that there are no transaction costs or other practical limitations. Using results from Panel A of Table IV, the VWRA portfolio generates excess returns of 82 basis points per month (10.3 percent annually). Although this is far smaller than excess returns estimated in most other studies, it is greater than—not less than—the 74 basis-point-per-month (9.25 percent annually) excess return obtained using CAPM (Table III). Thus, transaction costs, not inaccurate measures of risk, explain most of the large excess returns found in other studies.

VI. Hedge Fund Returns

A. Characteristics of Risk and Return

In addition to examining the profile of risk arbitrage returns generated by our index portfolio, we also examine the merger arbitrage return series published by Hedge Fund Research (HFR), a research and consulting firm that tracks the hedge fund industry. Their merger arbitrage monthly return series is compiled by averaging the net-of-fees returns from a sample of active merger arbitrage hedge funds over the 1990 to 1998 time period. Panel A of Figure 5 shows a scatter plot of HFR merger arbitrage returns versus market returns. For comparison purposes, RAIM returns versus market returns over the same period are shown in Panel B of Figure 5. Figure 5 shows that the payoff profile generated using our index approach is similar to that generated by HFR's sample of active hedge fund managers.

Table VIII presents piecewise linear regressions using HFR returns. To facilitate comparisons between these results and those presented in Table IV for the index portfolio, we use a threshold (kink point) excess market return of –4 percent. As is the case with the RAIM returns, –4 percent is the threshold that minimizes the sum of squared residuals. To gauge the sensitivity of the results to the choice of threshold, results are also presented using thresholds of –3 percent and –5 percent.

Results obtained using HFR returns are similar to those obtained using our index risk arbitrage portfolio returns—the market beta increases dramatically during market downturns. In depreciating markets, the HFR market beta is 0.60, slightly greater than the 0.47 market beta obtained using our RAIM portfolio returns. The intercepts are also similar—61 basis points per month using HFR index returns compared to 54 basis points per month using RAIM portfolio returns. In flat and appreciating markets, the HFR hedge fund index generates a positive market beta equal to 0.10. This compares to a beta of –0.03 using returns generated from our RAIM portfolio. Thus, in addition to being short a fraction of a put option on the market index, active managers are also long 0.10 call options on the market index.

B. Correlation Between Hedge Fund Returns and RAIM Returns

The similarity between our RAIM portfolio returns and hedge fund returns suggest that the RAIM returns may be a useful benchmark for evaluating the value added by active risk

Figure 5. This figure compares RAIM returns and hedge fund returns during 1990– 1998. Panel A presents hedge fund returns obtained from Hedge Fund Research's merger arbitrage index and Panel B presents RAIM returns. Data labels correspond to months (i.e., 9808 is August, 1998). Fitted lines from a piecewise linear regression are also shown.

arbitrage hedge fund managers. To examine the differences and similarities between our RAIM returns and those generated by active managers, we examine the correlation structure between RAIM returns, HFR returns, and individual hedge fund returns.[12] Individual fund returns are self-reported and were obtained from a large investor in merger arbitrage hedge funds.[13] Funds are included in our analysis if they have data for at least seven of the nine years between 1990 and 1998.

Table VIII Piecewise Linear Regressions: Hedge Fund Returns versus Market Returns

This table presents results from the following piecewise linear regression relating risk arbitrage hedge fund returns to market returns:

$$R_{Hedge\ Fund} - R_f = (1-\delta)[\alpha_{Mkt\ Low} + \beta_{MktLow}(R_{Mkt} - R_f)] + \delta[\alpha_{Mkt\ High} + \beta_{Mkt\ High}(R_{Mkt} - R_f)],$$

where $R_{Hedge\ Fund}$ is the mean monthly return of actively managed merger arbitrage funds tracked by Hedge Fund Research, R_f is the risk-free rate, R_{Mkt} is the monthly return on the value-weighted CRSP index, and δ is a dummy variable equal to one if the market return is greater than a threshold and zero otherwise. Results for three thresholds (–3 percent, –4 percent, –5 percent) are presented. To ensure continuity, the following restriction is imposed:

$$\alpha_{Mkt\ Low} - \beta_{Mkt\ Low}(Threshold) = \alpha_{Mkt\ High} - \beta_{Mkt\ High}(Threshold).$$

The sample consists of monthly returns over the 1990–1998 time period. Standard errors are in parentheses.

Dependent Variable	$\alpha_{Mkt\ High}$	$\beta_{Mkt\ Low}$	$\beta_{Mkt\ High}$	Adj. R^2	Sample Size
Panel A: Threshold = –3%					
Hedge fund returns	0.0067 (0.0012)***	0.5464 (0.0696)***	0.0862 (0.0346)*	0.457	108
Panel B: Threshold = –4%					
Hedge fund returns	0.0061 (0.0011)***	0.5985 (0.0787)***	0.1042 (0.0324)**	0.458	108
Panel C: Threshold = –5%					
Hedge fund returns	0.0055 (0.0011)***	0.6296 (0.0902)***	0.1223 (0.0314)***	0.443	108

*, **, *** indicate significance at the 0.05, 0.01, and 0.001 levels, respectively.

Table IX shows correlations between RAIM returns, HFR returns, and individual risk arbitrage fund returns. Panel A of Table IX shows that RAIM returns are positively correlated with both HFR returns and individual fund returns. However, the correlation between RAIM returns and a given fund's returns is generally lower than the correlation between two arbitrary funds' returns. To investigate this further, we examine the correlations after segmenting the data into two subgroups according to whether the market return minus the risk-free rate is greater than or less than –4 percent. Results, shown in Panel B for depreciating markets and Panel C for flat and appreciating markets, indicate that the correlation between RAIM returns and the HFR returns is high (0.66) in depreciating markets and close to zero (–0.02) in flat and appreciating markets. A similar effect is evident when comparing RAIM returns to individual fund returns. This pattern, however, does not hold when using quarterly, rather than monthly, returns. Panels D and E of Table IX show the results for depreciating markets and appreciating markets, respectively, using quarterly returns. Unlike the correlations calculated using monthly returns, the correlations using quarterly returns are much stronger, even in appreciating markets.[14] The correlation between RAIM and HFR is 0.38 and is statistically different from zero at the 5 percent level.

Table IX Correlation Between RAIM Returns and Hedge Fund Returns, 1990–1998

This table presents correlations coefficients between RAIM returns, HFR returns, and individual hedge fund returns. RAIM returns are generated from our sample of cash and stock swap merger transactions; HFR returns represent an average of hedge fund returns assembled by Hedge Fund Research; individual fund returns are self-reported returns obtained from a large hedge fund investor. Panels A, B, and C present correlations using monthly returns. Panels D and E present correlations using quarterly returns. The monthly return threshold used to distinguish between depreciating and appreciating markets is –4 percent, whereas the quarterly return threshold is 0 percent. Using a 0 percent threshold for quarterly returns ensures that there is an adequate sample of returns in both appreciating and depreciating markets.

	RAIM	HFR	Fund A	Fund B	Fund C	Fund D	Fund E	Fund F	Fund G	Fund H	Fund I	Fund J
Panel A: Correlations using Monthly Returns, Complete Sample												
RAIM	1.00											
HFR	0.36	1.00										
Fund A	0.17	0.61	1.00									
Fund B	0.09	0.59	0.43	1.00								
Fund C	0.22	0.68	0.62	0.58	1.00							
Fund D	0.14	0.52	0.20	0.31	0.30	1.00						
Fund E	0.15	0.66	0.55	0.51	0.59	0.32	1.00					
Fund F	0.41	0.84	0.46	0.38	0.50	0.28	0.53	1.00				
Fund G	0.34	0.84	0.59	0.53	0.67	0.39	0.56	0.67	1.00			
Fund H	0.32	0.62	0.45	0.26	0.42	0.14	0.42	0.92	0.58	1.00		
Fund I	0.23	0.60	0.45	0.45	0.54	0.31	0.47	0.54	0.43	0.23	1.00	
Fund J	0.40	0.69	0.46	0.34	0.47	0.30	0.48	0.52	0.69	0.49	0.42	1.00
Panel B: Correlations using Monthly Returns, Depreciating Markets												
RAIM	1.00											
HFR	0.66	1.00										
Fund A	0.65	0.82	1.00									
Fund B	0.40	0.66	0.53	1.00								
Fund C	0.70	0.82	0.64	0.75	1.00							
Fund D	0.39	0.82	0.46	0.53	0.69	1.00						
Fund E	0.61	0.81	0.68	0.33	0.78	0.77	1.00					
Fund F	0.74	0.90	0.80	0.42	0.62	0.64	0.69	1.00				
Fund G	0.58	0.91	0.87	0.55	0.78	0.71	0.85	0.74	1.00			
Fund H	0.43	0.49	0.74	–0.02	0.12	0.20	0.51	0.95	0.66	1.00		
Fund I	0.62	0.78	0.50	0.49	0.71	0.60	0.49	0.91	0.55	0.09	1.00	
Fund J	0.76	0.90	0.83	0.41	0.69	0.62	0.84	0.91	0.88	0.72	0.70	1.00

(*Continued*)

Table IX (*Continued*)

	RAIM	HFR	Fund A	Fund B	Fund C	Fund D	Fund E	Fund F	Fund G	Fund H	Fund I	Fund J
Panel C: Correlations using Monthly Returns, Flat and Appreciating Markets												
RAIM	1.00											
HFR	–0.02	1.00										
Fund A	–0.14	0.43	1.00									
Fund B	–0.20	0.45	0.29	1.00								
Fund C	–0.12	0.47	0.53	0.44	1.00							
Fund D	–0.01	0.39	0.06	0.16	0.12	1.00						
Fund E	–0.19	0.48	0.42	0.49	0.40	0.13	1.00					
Fund F	0.09	0.71	0.19	0.32	0.35	0.10	0.30	1.00				
Fund G	0.04	0.62	0.35	0.38	0.46	0.20	0.18	0.44	1.00			
Fund H	0.11	0.68	0.22	0.32	0.46	0.04	0.26	0.86	0.41	1.00		
Fund I	0.00	0.47	0.35	0.36	0.38	0.19	0.35	0.30	0.22	0.16	1.00	
Fund J	0.10	0.40	0.20	0.16	0.18	0.11	0.18	0.09	0.41	0.20	0.21	1.00
Panel D: Correlations using Quarterly Returns, Depreciating Markets												
RAIM	1.00											
HFR	0.69	1.00										
Fund A	0.37	0.87	1.00									
Fund B	0.18	0.42	0.43	1.00								
Fund C	0.64	0.87	0.83	0.73	1.00							
Fund D	0.55	0.55	0.16	0.15	0.31	1.00						
Fund E	0.59	0.88	0.86	0.26	0.75	0.46	1.00					
Fund F	0.67	0.89	0.72	–0.14	0.70	0.20	0.67	1.00				
Fund G	0.61	0.92	0.92	0.23	0.80	0.26	0.84	0.90	1.00			
Fund H	0.21	0.56	0.67	–0.30	0.30	0.02	0.60	0.91	0.75	1.00		
Fund I	0.65	0.73	0.61	0.81	0.91	0.31	0.50	0.64	0.61	–0.04	1.00	
Fund J	0.81	0.87	0.76	0.15	0.75	0.36	0.80	0.85	0.93	0.57	0.64	1.00
Panel E: Correlations using Quarterly Returns, Flat and Appreciating Markets												
RAIM	1.00											
HFR	0.38	1.00										
Fund A	0.09	0.34	1.00									
Fund B	0.36	0.26	0.23	1.00								
Fund C	0.03	0.32	0.53	0.20	1.00							
Fund D	0.20	0.59	0.08	0.19	0.07	1.00						
Fund E	–0.01	0.36	0.49	0.25	0.29	0.09	1.00					
Fund F	0.32	0.61	–0.03	0.24	0.07	0.08	0.38	1.00				
Fund G	0.35	0.61	0.20	0.23	0.39	0.46	0.10	0.44	1.00			
Fund H	0.51	0.60	0.06	0.43	0.18	0.04	0.06	0.89	0.65	1.00		
Fund I	–0.07	0.50	0.19	–0.22	0.19	0.49	0.44	0.28	0.02	–0.07	1.00	
Fund J	0.23	0.38	0.29	–0.04	0.26	0.22	0.15	–0.01	0.35	0.17	0.13	1.00

This correlation is similar in magnitude to the correlations between HFR returns and individual fund returns, which is surprising given that the HFR average is comprised of the individual funds' returns. Overall these results suggest that our RAIM portfolio provides a useful benchmark for evaluating hedge fund returns in depreciating markets, both for monthly and quarterly horizons. It also provides a useful benchmark in appreciating markets when a quarterly horizon is used. However, it does not reflect monthly variations of hedge fund returns in flat and appreciating markets.

There are a number of possible explanations for the lack of correlation between monthly RAIM returns and hedge fund returns in flat and appreciating markets. One possibility is that RAIM returns are generated from investments in simple cash and stock swap mergers whereas actively managed hedge fund returns reflect investments in other types of corporate transactions. In addition to investing in spin-offs and carve-outs, active hedge funds commonly invest in "collar" merger transactions. In a collar transaction, the amount paid to target shareholders depends on the acquirer's stock price during a period of time near the merger closing date. The typical collar results in a lower payment to target shareholders when the acquirer's stock price falls below a prespecified level and a higher payment if the acquirer's stock price rises above a prespecified level. Because of the concavity in the lower part of the collar and the convexity in the upper part of the collar, the return generated by an arbitrage investment in a collar deal decreases as the acquirer's stock price falls and increases as the acquirer's stock price rises. Since the acquirer's stock price is more likely to increase in appreciating markets, arbitrage returns generated by investments in collar transactions are likely to have a greater correlation with the market than simple stock transactions. The fact that hedge fund portfolios typically have positions in collar deals whereas our RAIM portfolio does not may explain why, in appreciating markets, individual fund returns are correlated with each other but not with the RAIM returns. This might also explain why the HFR returns have a beta that is more positive than the RAIM beta in appreciating markets and more negative in depreciating markets. To the degree that active managers use financial leverage, these inherent differences in betas will be amplified.

Results from our analysis suggest that three parameters, estimated with a piecewise linear regression, should be used in evaluating return series generated by risk arbitrage hedge funds. The three parameters are the down-market beta, the up-market beta, and the constant. RAIM regressions presented in Table IV provide parameter estimates that could be achieved using an index (i.e., no active information acquisition) approach. Superior hedge fund managers will have smaller down-market betas, larger up-market betas, and larger constants.

VII. Conclusion

Using a comprehensive sample of cash and stock-for-stock mergers, we examine returns generated from risk arbitrage. Our index portfolio starts with a fixed amount of cash and invests in every merger subject to three constraints. First, an investment in any merger cannot exceed 10 percent of total capital. Second, position sizes are limited by the liquidity of the underlying securities. A maximum price impact of 5 percent is allowed when investing in any position. Finally, the index fund must have an adequate amount of cash reserves to undertake the investment (the fund cannot use leverage). Returns obtained from the index portfolio are net of transaction costs including price impact and brokerage commissions. These costs are substantial. Whereas ignoring them would result in an annualized return to

risk arbitrage of 16.05 percent per year, including them reduces the return to 10.64 percent per year.

In addition to the index portfolio, we calculate value-weighted average risk arbitrage returns. In this approach, we assume transactions are costless and that an unlimited amount of capital can be invested, earning the average risk arbitrage return. Although this approach is clearly unrealistic, it provides a benchmark useful for comparisons to previous studies that use a similar approach.

Our results indicate that in most market environments, risk arbitrage returns are uncorrelated with market returns. However, during market downturns, the correlation between market returns and risk arbitrage returns increases dramatically. This effect is asymmetric—similar increases are not observed in market rallies. We document similar patterns for out-of-sample tests, namely, the actual returns to professional risk arbitrage activity during the 1990s. Because of this similarity, our nonlinear analysis of risk arbitrage index manager returns can be used to generate a benchmark for evaluating risk arbitrage hedge fund managers.

These results suggest that risk arbitrage returns are similar to those obtained from writing uncovered index put options. In most states of the world, a small put premium is collected. However, in rare states, a large payment is made. This payoff profile suggests that risk arbitrage may be better evaluated using a contingent claims analysis rather than a linear asset pricing model such as CAPM. However, our analysis shows that when measuring excess returns, the error associated with using CAPM is significant only when the nonlinearity in returns is severe. This tends to be the case in time periods when cash, rather than stock, is the predominant form of merger consideration. Although linear asset pricing models mask the true risk in risk arbitrage, they do not result in large errors when measuring excess returns.

Results from our analysis indicate that risk arbitrage generates excess returns of roughly four percent annually. For individual investors that typically do not receive interest on their short proceeds, the excess return is only two percent. This compares to estimates from other studies that range between 11 percent and 100+ percent. Most of the difference between our estimates and those obtained in other studies can be attributed to transaction costs. Although our estimate is far less than estimates reported in other studies, it is still substantial. We postulate that this excess return reflects a premium paid to risk arbitrageurs for providing liquidity, especially during severe market downturns.

Appendix

A. Indirect Trading Costs

Breen et al. (1999) estimate the price impact of a trade of specified size based on liquidity characteristics of the underlying security. The price impact equation is given as

$$\frac{\Delta P}{P} = \beta(NTO) \tag{A1}$$

where $\Delta P/P$ is the price impact, equal to the percentage change in price resulting from a trade with net turnover equal to NTO. Net turnover is defined as one-tenth of the buyer-initiated volume minus seller-initiated volume divided by shares outstanding. Using the

above equation, Breen et al. estimate β from price changes and net turnover over 5-minute and 30-minute intervals. The βs are then used in a cross-sectional regression to obtain the following price impact model (Breen et al. (1999, Table 5)).

$$\beta = 8.77 + 2.52X_1 - 1.84X_2 - 1.39X_3 - 1.92X_4 - 27.5X_5 - 8.29X_6 - 0.02X_7 - 0.38X_8 - 0.63X_9 - 0.08X_{10} - 0.39X_{11}, \quad \text{(A2)}$$

where

X_1 = log of market capitalization,
X_2 = log of previous quarter's trading volume,
X_3 = price at the end of the previous month divided by price 6 months prior,
X_4 = dummy variable equal to one if the equity is included in the S&P 500,
X_5 = dividend yield,
X_6 = R^2 of returns versus NYSE obtained from regressing monthly returns over the prior 36 months,
X_7 = NYSE inclusion dummy,
X_8 = NASDAQ inclusion dummy,
X_9 = dummy variable equal to one if last earnings release was more than 2 months ago,
X_{10} = percentage institutional ownership,
X_{11} = dummy variable equal to one if there are options traded on the security.

Table AI Pre-1975 Direct Trading Costs

Brokerage Commissions	
Size of Trade	**Commission**
\$100–\$2,499	1.3% + \$12.00
\$2,500–\$19,999	0.9% + \$22.00
\$20,000–\$29,999	0.6% + \$82.00
\$30,000–\$300,000	0.4% + \$142
Over \$300,000	Negotiable
Round-lot Surcharge	
Number of Round Lots	**Charge per Round-lot**
0–10	\$6.00
>10	\$6.00 for first 10, \$4 for each additional round-lot
Transfer Tax	
Stock Price	**Transfer Tax (per share)**
Less than \$5	\$0.0125
Between \$5 and \$10	\$0.025
Between \$10 and \$15	\$0.0375
More than \$15[a]	\$0.05

[a] According to Francis (1980), the transfer tax increases to \$0.05 when the stock price exceeds \$20. He does not indicate the magnitude of the tax for stock prices between \$15 and \$20. Therefore, we assume the \$0.05 tax applies to all stocks with a price above \$15.

Table AII Post-1975 Direct Trading Costs

Date	Per Share Trading Cost
1975–1979	$0.10
1980–1989	$0.05
1990–1998	$0.04

In this paper, we use equation (A2) to estimate b for both the acquirer and the target.[15] For any arbitrary price impact level (e.g., $\Delta P/P = 5$ percent), we then use the estimate of β to calculate the maximum allowable number of shares that can be traded, assuming this maximum does not result in a position that exceeds 10 percent of the portfolio's total value. Equations (A1) and (A2) are also used to compute the indirect cost of trading. For every transaction in our index portfolio, we subtract transaction costs equal to the price impact implied by equations (A1) and (A2), divided by 10. The factor of 10 is used to account for the fact that traders attempt to limit the price impact of their trades by placing many small orders to accumulate a large position.

B. Direct Trading Costs

To calculate realistic returns using the risk arbitrage index portfolio, direct trading costs must be estimated. For the pre-1975 sample, this is a straightforward task. During that time period, per-share trading costs were regulated by the NYSE. The regulated direct trading costs consisted of three main components: (1) brokerage commission, (2) round-lot surcharge for orders of 200 shares or more, and (3) transfer taxes based on the price of the stock being bought or sold. Table AI, based on Francis (1980), outlines each of these costs.

After 1975, brokerage houses were free to compete on price. Because there is no set transaction cost after 1975, we assume the costs per share outlined in Table AII.

Endnotes

[1] The sample includes stock swap mergers, cash mergers, and cash tender offers. Constructing returns from individual mergers allows us to avoid the sample selection issues inherent in recent studies that use hedge fund returns to assess the risk/reward profile of risk arbitrage. For example, Fung and Hsieh (1997), Ackermann, McEnally, and Ravenscraft (1999) and Agarwal and Naik (1999) provide analyses of hedge fund returns. Fung and Hsieh (2000) present a discussion of the sample selection biases inherent in using these returns.

[2] Shleifer and Vishny (1997), p. 37.

[3] It is possible that the complicated transactions that we drop are systematically different from the simple transactions that we retain. To get an idea of whether there are systematic differences between these two groups, we examined the average takeover premium, acquirer stock price reaction, the percentage return obtained from taking a long position in the target (without the corresponding hedge) from announcement through consummation, failure probability, target size, acquirer size, and percentage of friendly transactions for the two groups. For all of these variables, differences between the two groups are not significantly different from zero at or below the five percent level.

[4] Although large funds receive interest on short proceeds, individual investors typically do not (unless they have a substantial amount of capital invested with their broker). Results presented in this paper assume that the risk-free rate is paid on short proceeds. Results from unreported analyses indicate that annual returns are reduced by approximately two percent if interest is not paid on short proceeds.

[5] Whether or not this is a reasonable assumption is debatable. The alternative is to calculate the weighted average of returns for transaction days across all days. This approach generates higher returns, but makes the implicit assumption that capital is never idle.

[6] For successful deals, there are no transaction costs associated with closing a position. In the case of a cash deal, the target's stock is traded for the cash consideration. In a stock deal, the number of shares of the acquirer's stock that is exchanged for the target's stock is exactly equal to the number of acquirer's shares initially shorted. Thus, for both successful cash deals and successful stock deals, no securities are sold and no transactions costs are incurred. Transaction costs are incurred when closing out positions in failed deals.

[7] The choice of initial capital is not inconsequential. Particularly in the early 1960s, there was a dearth of mergers. If restrictions are placed on the amount that can be invested in any one deal (due to illiquidity or diversification requirements), a significant amount of the initial capital must be invested in cash, thereby distorting the returns from risk arbitrage.

[8] To make sure that the large returns associated with the value weighting procedure are not driven by the choice of weights, risk arbitrage returns are also calculated using an equal weighting procedure. Whereas the VWRA average return is 16.05 percent, the equal weighted average is 18.08 percent.

[9] A formal test of the null hypothesis that the risk arbitrage market beta is the same in appreciating and depreciating markets rejects at the 0.001 level (see Jagannathan and Korajczyk (1986) for a description of the procedures used to test for nonlinearity in return series).

[10] Schwert (2000) uses the same definition for one of the four hostility variables in his examination of the economic distinction (based on accounting and stock price data) between hostile and friendly deals. Schwert also notes that hostile deals have a lower likelihood of deal completion.

[11] In addition to the probit model described in Table V, we also examined the effect of a market decline on the ratio of failed deals in the month to active deals in the month. Results from this analysis are consistent with those obtained from the probit model. A 5 percent decline in the market in the previous month increases the fail/active ratio from 0.050 to 0.059, an increase of 18 percent. This effect is significant at the 0.1 percent level.

[12] We are grateful to an anonymous referee for suggesting this analysis.

[13] In addition to merger arbitrage, many large hedge funds pursue other relative value strategies (e.g., convertible bond arbitrage). To check whether alternative investments affect our results, we performed our analyses using average returns from a select group of hedge funds that, based on interviews, we are reasonably sure focus primarily on event arbitrage (mergers, spin-offs, carve-outs, self tender offers). Results obtained from this subgroup of managers are both quantitatively and qualitatively similar to those presented.

[14] To distinguish between depreciating and appreciating markets when using quarterly returns, we use a market return threshold of zero percent per quarter. This ensures that we have an adequate sample size in both depreciating and appreciating markets.

[15] Particularly for older transactions, we do not have all of the independent variables required for obtaining predicted values from equation (A2). Specifically, we are lacking X_4, X_9, X_{10}, and X_{11}. For these variables, the means from the Breen et al. (1999) sample are used.

References

Ackermann, Carl, Richard McEnally, and David Ravenscraft, 1999, The performance of hedge funds: Risk, returns, and incentives, *The Journal of Finance* 54, 833–874.

Agarwal, Vikas, and Narayan Naik, 1999, On taking the "alternative" route: Risks, rewards, style and performance persistence of hedge funds, Working paper, London Business School.

Baker, Malcolm, and Serkan Savasoglu, 2002, Limited arbitrage in mergers and acquisitions, *Journal of Financial Economics*, forthcoming.

Bhagat, Sanjai, James Brickley, and Uri Loewenstein, 1987, The pricing effects of interfirm cash tender offers, *The Journal of Finance* 42, 965–986.

Breen, William, Laurie Simon Hodrick, and Robert Korajczyk, 1999, The determinants of equity illiquidity, Working paper, Northwestern University.

Brown, Keith, and Michael Raymond, 1986, Risk arbitrage and the prediction of successful corporate takeovers, *Financial Management*, 15:3, 54.

Canina, Linda, Roni Michaely, Richard Thaler, and Kent Womack, 1998, Caveat compounder: A warning about using the daily CRSP equal weighted index to compute long-run excess returns, *The Journal of Finance* 53, 403–416.

Dukes, William, Cheryl Frohlich, and Christopher Ma, 1992, Risk arbitrage in tender offers: Handsome rewards—and not for insiders only, *Journal of Portfolio Management*, 18:4, 47–55.

Fama, Eugene, and Ken French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Francis, Jack, 1980, *Investments Analysis and Management*, 3rd ed. (McGraw-Hill, New York).

Fung, William, and David Hsieh, 1997, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *Review of Financial Studies*, 10, 275–302.

Fung, William, and David Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies,* 14, 313–341.

Fung, William, and David Hsieh, 2000b, Performance characteristics of hedge funds and CTA funds: Natural versus spurious biases, *Journal of Quantitative and Financial Analysis* 35, 291–307.

Glosten, Larry, and Ravi Jagannathan, 1994, A contingent claim approach to performance evaluation, *Journal of Empirical Finance* 1, 133–160.

Jackwerth, Jens Carsten, 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.

Jagannathan, Ravi, and Robert Korajczyk, 1986, Assessing the market timing performance of managed portfolios, *Journal of Business* 59, 217–235.

Jarrell, Gregg, 1984, Change at the exchange: The causes and effects of deregulation, *Journal of Law and Economics* 27, 273–314.

Jindra, Jan, and Ralph Walkling, 1999, Arbitrage spreads and the market pricing of proposed acquisitions, Working paper, Ohio State University.

Karolyi, G. Andrew, and John Shannon, 1998, Where's the risk in risk arbitrage? Working paper, Richard Ivey School of Business, The University of Western Ontario.

Larker, David, and Thomas Lys, 1987, An empirical analysis of the incentives to engage in costly information acquisition: The case of risk arbitrage, *Journal of Financial Economics* 18, 111–126.

Merton, Robert, 1981, On market timing and investment performance I: An equilibrium theory of value for market forecasts, *Journal of Business* 54, 363–406.

Samuelson, William, and Leonard Rosenthal, 1986, Price movements as indicators of tender offer success, *The Journal of Finance* 41, 481–499.

Schwert, G. William, 2000, Hostility in takeovers: In the eyes of the beholder? *The Journal of Finance* 55, 2599–2640.

Shleifer, Andrei, and Robert Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

An Alternative Future: Part I

An exploration of the role of hedge funds

Clifford Asness

One of the main financial market stories of the last three to five years has been the explosive growth of hedge funds. Depending on whom you ask, hedge funds are either the wave of the future, or they are a dangerous fad that has been grossly over-capitalized and all will end in ruin. Both sentiments contain elements of truth. The good news for hedge funds is that a portfolio structure that divides capital between traditional index funds to obtain beta (or market exposure) and hedge funds to earn alpha is very appealing.

Traditional active management attempts to add alpha by adjusting index holdings in an arbitrary, confusing, and constraining manner.[1] It can be viewed as a tie-in sale between an index fund and a very constrained hedge fund. Hedge funds allow for a much clearer separation of the unrelated activities of obtaining index exposure and generating alpha, thus leading to clean portfolio construction, performance attribution, and fee breakdowns. Furthermore, sources of alpha and techniques unavailable in traditional mandates are available in this format.

To fulfill their promise, hedge funds need to recognize and improve on shortcomings. A companion piece to this article explores hedge fund fees in more depth, examines various dark sides to hedge fund investing, and recommends future evolutionary changes needed to help hedge funds achieve their potential (see Asness [2004]).[2]

Given my generally pro-hedge fund slant here, some fair disclosure is in order. I am a hedge fund manager.

Simple Hedged Investment Strategy

A simple definition of a hedge fund investment strategy might go as follows: a strategy that trades relatively liquid assets (versus some other alternatives like private equity and real estate), seeks to make positive average returns over time, and provides diversification

Reprinted from *The Journal of Portfolio Management's* 30th Anniversary Issue, 2004, with permission.

versus traditional stock and bond markets. A more quantitative definition might say we want positive expected return with low correlation.

Of course, the trillion dollar question is how to create such an asset. We can start with a simple and familiar example, a traditional, actively managed equity portfolio. This is a collection of stocks, usually a relatively small subset of those in a stock index, that an active manager believes will outperform the others in the index. Define these portfolio weights for the active manager as A and the weights in the index as I. The weights that make up A and I separately sum to 100%.

We can represent the active manager's holdings A as:

$$A = I + [A - I] \tag{1}$$

Now what does this say beyond that it is one of the world's simplest equations? Well, it says you can define what the active manager owns in two parts. First, he owns the index, and second he owns a zero investment long-short portfolio that represents his views. Since A – I is the difference in weights across all stocks between the active portfolio and the index, and since the weights in A and I both sum separately to 100%, A – I is a *zero investment portfolio*. In traditional long-only investment management the expected return to A – I is usually referred to as the manager's alpha and the volatility of the return to A – I as the manager's tracking error.[3]

So how do we make a hedge fund? Well, we can make one very simple hedge fund by shorting the index (say, with a futures contract). When the active manager shorts one unit of a futures contract, she subtracts the excess return (over cash) of the index from her portfolio. Let's add the short futures position to Equation (1) and define a new set of weights H. Note that we can interpret these equations in terms of portfolio weights or in terms of the returns on these portfolios:[4]

$$H = I + [A - I] - [I - CASH] \tag{2}$$

$$H = CASH + [A - I]$$

What is this new beast H? Well, it is a very simple hedge fund. It will have low to zero correlation with the index I. It will have a positive expected return if the manager has skill or alpha. Basically, assuming that the original manager is trying to beat the index through stock-picking skill, this new asset H delivers the return on that skill alone, separate from the index return.[5]

Why might this new investment H be interesting? What does it give us beyond what was achievable with simple index funds (I) and active management (A)? Well, first and foremost, it allows us to choose our exposure to the index and the active manager's skill separately.[6]

Say there is an active manager who is quite skillful, but takes relatively limited tracking error from the benchmark. Say, for instance, the volatility of I is 15% and that of A – I is 3%. A believer in this manager's skill might want more than a 1:5 ratio versus market risk. If the manager is skillful enough (implausible perhaps to fans of efficient markets, but they are probably not the ones buying active management to begin with), investors might not get enough of this skill with traditional active management.

In fact, one can view traditional active management as a tie-in sale. In this example, you get five units of an index fund for every one unit of active management in the only choice available to you under the traditional structure. Once H exists, though, investors can choose the ratio of H and I.

There is another and more subtle benefit to creating asset H. It allows for a much easier and more natural analysis of the manager's skill and the fees paid for that skill. When investing in active Portfolio A, it is quite easy to be imprecise about such things. When investing in some combination of I and H, it is hard not to be more precise, as the sole purpose of I is index exposure, and the sole purpose of H is exposure to the manager's skill. This decomposition and separation is itself quite clarifying.

Indeed, I argue even a die-hard fan of efficient markets who would not be caught dead owning H will like this way of thinking. The clarity behind this approach is less likely to allow underperforming stock pickers to be saved by a bull market. In the hedge fund-plus-index fund construct, this underperformance does not show up as making less than the stock market, but rather as H losing to cash. This tends to be overlooked less often.

Next let's consider leverage L, and modify Equation (2):

$$H = CASH + L\,[A - I] \tag{3}$$

Once a hedge fund manager is shorting the index, he, assuming for simplicity that the cost of borrowing is zero, can rather easily apply leverage to the remaining lower-volatility, pure active bet. Depending on L, a hedge fund manager's volatility can be equal to, higher than, or lower than the traditional manager's (A's) tracking error. Does $L > 1$ make the hedge fund manager more risky than the active manager?

Well, say A's tracking error is 300 basis points (bp), and the hedge fund manager is levered 3:1 as above. In this case, the hedge fund manager's volatility is 3×300 bp = 9%. The index fund's volatility is 15%, and the actively managed portfolio is very slightly riskier (the square root of 15% squared plus 3% squared). Thus, the acts of selling futures and levering clearly do not automatically lead to higher risk.

An interesting question now regards the fair fee for H. Say the traditional active manager charges a 65 bp fixed fee for taking 300 bp of tracking error. What is the fair fee for a hedge fund manager levered 3:1?

Well, the active manager A is delivering a hedge fund plus an index fund. Assume the index fund would cost 5 bp alone. Obviously the cost of A's skill is an extra 60 bp. The hedge fund manager H is delivering 3 times the skill, with no index exposure, so the fee equivalent to A's fee for the hedge fund is 180 bp a year.[7]

While this seems gigantic by conventional standards, it's simply a consistent consequence of the math. Traditional active management is a tie-in sale of a cheap good (an index fund) and an expensive one (manager skill). When you break the tie-in sale and lever up the expensive part, it costs more.

It's very possible that (when viewed this way) the common perception that hedge funds charge far more than traditional active managers is exaggerated. It may remain an open question as to whether certain necessary assumptions hold (such as low-beta hedge funds delivering more active risk per dollar than traditional managers), but the comparison of hedge fund and traditional active management fees is clearly more complex than it seems at first glance when sticker shock can be blinding.

To sum this up, a simple application of financial engineering allows the isolation of skill as a stand-alone investable asset from index exposure. This separation provides three main benefits: 1) It allows investors to choose the allocation between market exposure and skill; 2) it potentially improves the monitoring of risk and performance attribution; and 3) it sheds light on fees.

More Complex Hedge Fund Strategies

A casual review of the hedge fund world reveals very few that resemble my "long active management, short futures" strategy. This simple strategy captures the essence of hedge fund investing in perhaps as simple a form as possible, but that is its only real virtue.

We need to generalize the ideas behind hedge fund investing, and also understand some of the major categories of hedge funds. They generally all share the goal of making money on average without high correlation with traditional markets, and usually share the characteristics of less constrained investing with the ability to short, lever, or use derivatives to try to meet this goal.

First, staying within the world of individual stocks, and still trying to benefit from stock-picking skills, we can generalize the H construct to include not just shorting of stock-index futures, but of specific individual stocks. While there are all kinds of variations on the theme, two broad categories of hedge funds that do this are called *market-neutral* and *long-short equity*. Both combine actively chosen long and short positions, with the general idea that if alpha exists in choosing stocks to purchase, it probably also exists in choosing stocks to short, and a superior portfolio can be put together by doing both.[8]

Market-neutral investing is often a quantitative style that seeks to be balanced long and short (equal dollar or more rigorously equal beta), and usually consists of very diverse long and short portfolios. Long-short equity investing generally refers to a less quantitative strategy of stock-picking. Long-short managers often run to a long bias, and will change their net and gross exposure over time. Both strategies, of course, seek to achieve a positive average return with lower or much lower correlations with market indexes than those of traditional long-only stock portfolios.

Other long-short strategies seek to achieve the same goals through different means from individual stock-picking. One is *merger arbitrage*.

When a new merger is announced, typically the target price rises, but typically by less than implied by the full terms of the merger; as the market recognizes some probability the merger will fail. The merger arbitrage hedge fund manager generally goes long the target and short the acquirer, betting that the deal will go through, and the rest of the spread will close for a profit. The risk is that the deal falls through, and the spread widens precipitously. The merger arbitrage manager will obviously profit over the long run if the small gains from many deals going through outweigh the generally larger losses from a much smaller number of deals failing.

Another example is a strategy called *statistical arbitrage*.[9] Some use this term to refer to any quantitative investment strategy. I use the term here to mean a narrower set of short-term contrarian strategies. This strategy, on a very short-term basis (less than a month but down to days or even hours), basically goes long very recent losing stocks and short recent winners. Its practitioners typically try to hedge the long positions with the short positions

along a bevy of risk characteristics (e.g., beta, size, or industry). In general they make money when recent winners and losers reverse course.

One interpretation is that these managers are betting that the winners and losers got that way through temporary and uninformed price pressure—that is, somebody big was buying or selling for reasons without particular insight, and they lose when they get on the wrong side of an informed trader.

What is interesting about merger and statistical arbitrage is that both trade individual stocks only, but are very difficult to make relevant within a typical, traditional long-only structure. For instance, simply buying an announced takeover target firm naked without shorting the acquirer is a far less precise bet on the merger succeeding (assuming the merger is stock-for-stock, not a cash deal).[10]

Imagine one computer company is buying another. The merger arbitrageur's bet is just that the deal will occur, although idiosyncratic news on one of the companies can still matter. Just buying the target also bets on the merger occurring, but adds a large bet that the market will reward computer companies over the same time horizon. Essentially, the hedge fund structure allows strategies to become relevant that would otherwise not matter much, as they would be too imprecise and risky.

A generalized form of *fixed-income arbitrage* involves attempting to find a set of bonds to long and short, where the longs and shorts hedge each other somewhat, and the longs are expected to outperform the shorts. Why a hedge fund for this? Well, for many of these strategies, the instruments are not very volatile, and the hedges are very good (i.e., the longs and the shorts move together well).

Leverage can often be required to make this a strategy worth pursuing. Short-selling is obviously required to bet on the relative returns between the longs and the shorts at any material risk level (beyond just underweighting versus an index), and not on the absolute direction of bond markets. For example, a fixed-income arbitrageur might have the opinion that the five-year U.S. Treasury Bond is attractive versus the two-year and ten-year, but no opinion on the direction of bonds in general.

Another strategy that would not make much sense outside the hedge fund structure (if shorting and leverage are not permitted) is called *convertible arbitrage.* Essentially, this strategy buys convertible bonds that can be traded in at some prescribed price for shares of common stock.

An outright purchase of these bonds entails many risks, including exposure to fluctuations in the price of the common stock. Hedge fund managers can remove this risk (or at least try to, as hedging is not always a precise science) by shorting the common stock. While there are many variations, convertible arbitrage managers generally try to profit from finding convertible bonds that are trading too cheaply, and locking in this value with their hedges.[11]

Leverage is often applied, as a diversified portfolio of hedged convertibles is fairly low-risk. Without shorting and leverage, the only way to bet that a convertible is too cheap would be an outright purchase of the bond, clearly a far less precise and more risky way to express such a view. This is another case where the hedge fund construct takes potential sources of return that may be irrelevant with traditional tools and makes them relevant.

Two other major classes of hedge fund strategies are *macro* and *commodity trading advisors* (CTAs). Macro is a very general class that can do anything, but is thought of as looking across borders, making bets on the absolute and relative attractiveness of asset classes, countries (in any asset class), and currencies. Again the hedge fund structure

makes something potentially important out of strategies and trades with little allowable impact on traditional portfolios.

Either by constraint, or by specialization of mandate, many cross-country and cross-asset class positions are not common to traditional managers, and some must be pursued through shorting or leverage to matter (e.g., a view that two non-U.S. currencies will move relative to each other with no view on the U.S. dollar usually has little place in traditional portfolios). Like macro managers, CTAs also trade a wide variety of global assets such as equity indexes, government bonds, currencies, and physical commodities through the futures markets. While macro managers are thought of as making medium-to-long-term valuation calls, CTAs are generally thought of as trading on short-to medium-term trend-following or momentum.

The hedge fund structure overall allows skill to be packaged as a pure investment in a variety of asset classes. It also allows meaningful implementation of certain strategies that may be skill-based, but also might represent fair compensation for undertaking certain risks or providing liquidity (e.g., merger or statistical arbitrage). Without short-selling and leverage, which are typically part of the hedge fund structure, and a generally less constrained environment, much of this would not be possible.

There is a positive effect here not just for hedge fund managers and investors but for financial markets in general. Without merger arbitrageurs, spreads after announced deals would likely remain much wider, putting stockholders of the target company at risk if they want to reap the merger's benefit. Without statistical arbitrage, markets would be less liquid, as a big seller would find fewer willing buyers. Without fixed income arbitrage, strange discontinuities in yield curves, abnormally high prices for risky bonds, and other anomalies would all be larger and last longer, potentially warping real investment decisions. Thus many hedge fund strategies make markets more efficient.

Why Now?

So why are hedge funds gaining such popularity now when these strategies have been around for quite a while?

Learning

One possibility is that hedge funds are a superior investment structure but this is not a simple thing to ascertain, and we just were not ready to understand it until recently.

Many of these strategies involve much more financial engineering than more typical investments. The development of options exchanges in the 1970s, the birth and development of modern portfolio theory, and numerous technological advancements all were necessary for the acceptance of such strategies to grow.

At one time in history, we would have said: Why mutual funds now? Why money market funds now? Why index funds now?

A Low Equity Risk Premium

There is a growing belief or acceptance that traditional markets will not provide the risk premiums (returns over riskless assets) going forward that they have in the past. In particular,

many investors and researchers believe the equity risk premium in the U.S. going forward will not be what it was historically.

Look at the S&P 500's P/E in the Exhibit. Many authors have shown that when prices are high versus fundamentals, as they are now, expected future real stock market returns are low. One might note that the P/E in the Exhibit is currently far lower than the peak of the bubble in early 2000, but that is damning with faint praise. Current P/Es are very high versus 120 years of history, auguring a lower equity risk-premium going forward.

The generally poorer prospects for traditional equity markets make the pursuit of a higher Sharpe ratio (expected excess return per unit of volatility) through diversifying strategies more important than it once was. In the past, a very attractive risk premium was available just for going long an equity index fund. This is no longer the case. The bear stock market of 2000–2002 may have reinforced this lesson and stepped up the rush to hedge funds as a needed diversifier, and some institutional investors may also have realized that certain other assets like venture capital and private equity were not the diversifiers they once thought.

It May Be a Fad

Part of the popularity of hedge funds might be faddish. They are generally perceived to be the investment of choice of the rich and the informed, and they are more interesting and fun to discuss than your Vanguard index fund. And, unlike the stock market of 2000–2002, they have not failed for a few years now (1998 is much longer ago than 2002). Fads are difficult to predict or explain, but even the most pro-hedge fund observer would have to admit that a faddish rush accounts for at least some of the recent hedge fund explosion.

The Hurdle for Portfolio Improvement Is Low

Perhaps the most important reason for the newfound popularity of hedge funds, particularly among institutions, is the realization of exactly how low a Sharpe ratio is needed from hedge

Exhibit S&P 500 P/E (price divided by 10-year real earnings)

Sources: Calculated from Shiller website and DataStream.

funds to improve, or even radically improve, the risk-adjusted return of a portfolio. If you can find a hedge fund with zero correlation with your current portfolio, and if it has any positive expected return over cash, adding some of it will improve your overall expected results.

While many in the hedge fund world seek or claim to have annual Sharpe ratios of 1.0, 2.0, 3.0, or more, in reality numbers like 0.5, 0.2, and even 0.1 would warrant inclusion in a traditional portfolio. Now, it's a different question as to whether a hedge fund with a 0.1 Sharpe ratio is a viable business. Such a manager will have to deal with being down in very close to half the years, and the time horizon needed to be reasonably certain they provide value is gigantic.

Specifics aside, the general realization among institutions of the low hurdle for including diversifying assets in their portfolio has clearly radically raised demand for hedge funds. While traditional hedge fund investors have for years demanded their hedge funds make money all the time, more modern hedge fund investors are closer to demanding their hedge funds make money a bit more often than they lose, as long as the losing periods are relatively unrelated to the losing periods in traditional markets.

Hedge Fund Alpha versus Beta

The concepts of alpha and beta are familiar from the world of stock-picking. In this traditional investing world, beta is a portfolio's exposure to the stock market; a beta of 1.0 indicates that if the market goes up 10%, the portfolio would be expected to go up 10%. Alpha is a manager's expected return (or often ex post their realized return) above or below the return attributable to their beta and that attributable to the risk-free rate.

The typical interpretation of alpha is that of *skill.* If over the long term a portfolio manager has had returns in excess of what you would expect simply observing her beta, we say the manager has demonstrated positive alpha. It is the alpha that the construct H in Equation (2) is trying to deliver in a pure form.

Even in traditional markets, the world of one single market beta seems dated. For stock-picking, researchers have generalized the concept of beta to include exposures to other risk factors or styles.

The most famous generalization is the Fama and French [1993, 1996] three-factor model. In this model there are three betas representing 1) systematic exposure to the market (regular old beta), 2) exposure to the excess performance of value stocks over growth stocks, and 3) exposure to the excess performance of small stocks over large stocks. The idea behind the three-factor model is not to credit (or blame) managers for the returns to known strategies available to all. Such exposures are not normally considered skill.

Generalizing outside the stock-picking world, new potential betas arise. Exposure to interest-rate risk and credit risk can both be thought of as systematic exposure. We can easily imagine measuring the alpha of a manager net of her average exposures to all the various betas. If a manager averages, and is always expected to average, long high-yield bonds, we want to know not just if she posts returns above cash, but if she posts returns above what one would expect, given this passive exposure to credit.

The concept of alpha and beta can also be applied to the hedge fund world. Asness [2003], Jensen and Rotenberg [2003], Dunn [2004], and Siegel [2004] make a similar observation. In the long-only world, market beta is something available to all, a known

strategy with an explanation of why it should over time deliver positive returns above the risk-free rate. Hedge fund betas are a very similar concept.

Notice above when I discuss strategies such as merger arbitrage, convertible arbitrage, and statistical arbitrage, I am describing a systematic strategy (e.g., "the merger arbitrage manager generally goes long the target and short the acquirer"). I am not stressing manager skill.

Imagine a merger arbitrage strategy that participates in every announced merger. This is not enough to specify a strategy completely. How would we weight each announced merger? When after an announcement does a strategy enter into a position?

Should such ambiguity bother us when we try to define this concept of a hedge fund beta? I argue no. It is inherent in any traditional long-only index fund attempting to deliver the theoretically easy yet practically hard to define "market exposure." In other words, for a traditional index fund do we include 500, 1,000, or 5,000 stocks? Do we have a hard rule like capitalization, or a committee to choose the stocks? When do we add new stocks and replace old ones? Do we use straight market capitalization weighting or some alternative scheme?

Thus, while no two practical implementations of a merger arbitrage beta portfolio will be precisely alike, the only real question is whether such a construct helps us understand hedge fund returns, and represents a real, viable, and potentially important investment choice. The answer to these questions is yes.

Of course, merger arbitrage is only one example. Many other hedge fund strategies are amenable to similar thinking. For instance, it is easy to imagine systematic and easily explainable strategies for convertible arbitrage, statistical arbitrage, and even fixed-income arbitrage.

Precise models will of course vary. I am not claiming the breakdown or classification of any given strategy into beta and alpha represents a clear bright-line test. It is a useful concept, though, to think of any hedge fund manager as a set of exposures to 1) traditional betas (market exposure, value versus growth, small versus large, interest rates, credit) and 2) hedge fund betas (basic merger arbitrage, convertible arbitrage, fixed-income arbitrage). Any skill hedge fund managers have comes on top of these exposures. What both these betas have in common is that they represent a known implementable strategy, and thus a source of potentially common systematic risk.[12]

An interesting and potentially contentious question is whether the expected return to each hedge fund beta is positive. While that is not necessary to make this concept useful (a systematic risk factor that explains why a group of managers go up or down can be useful even if that risk factor is not rewarded on average over time), we generally think of these betas as offering long-term positive results.

For every hedge fund strategy, there is some argument for why it should have an expected positive Sharpe ratio. In general, like those in the long-only literature, these arguments can fall into the efficient market camp (where investors are rational to insist on being paid to bear some undiversifiable risk) or the inefficient market camp (where a strategy works on average because it capitalizes on known investor biases or structural reasons why some investors will accept lower returns so the hedge fund can reap higher ones).[13]

Again take the example of merger arbitrage. Once a deal is announced, going long the target and short the acquirer has a strange pattern of returns (see Mitchell and Pulvino [2000]). While the average returns have been positive over the long term, individual positions are highly negatively skewed, meaning that when deals fail you lose a lot more than

you can gain when they succeed. Furthermore, an entire portfolio of mergers has a positive market exposure, and interestingly a negative coskewness with the market (when the market suffers large declines, the merger arbitrage strategy suffers more than one would guess if all returns were distributed in a normal fashion).

Has merger arbitrage had a positive long-term return because managers rationally demand to be paid for bearing this risk? Or has it been successful because investors irrationally will pay to avoid the chance that an individual deal will blow up? Or, finally, have managers just been lucky for 20 or 30 years?

All these questions are not unique to merger arbitrage or hedge fund betas in general, but rather mirror the arguments over more traditional risk factors such as value versus growth and small versus large.

Now, given a model for traditional and hedge fund betas, we have not eliminated the role of skill. Skill or alpha is simply the return net of these betas (in either an ex ante expected sense or in the noisier realized sense). The betas are exposures to strategies (traditional and hedge fund) that are known, common to many managers (so they explain a lot of common variation) and generally support why managers get paid (have positive expected returns over time). Alpha is what is left over.

To continue using merger arbitrage as an example, imagine again a systematic merger strategy. Now imagine a particular merger arbitrage manager who generally practices this strategy, but invests in only a subset of the available deals, at weights of his own choosing, and at a degree of leverage that varies through time. This manager's returns versus the basic merger arbitrage strategy over time would be a measure of his alpha.

Interestingly, it is reasonable that conclusions may change over time. Thirty years ago merger arbitrage was pursued by many fewer managers than today. Thus the strategy in all likelihood delivered a much higher risk-adjusted return then, as spreads were much wider with less capital invested.[14] So, back then, was basic merger arbitrage an alpha or a beta?

Well, there is a strong argument it was an alpha (and the first person to notice it would be offended if I called it beta). It was not a widely known strategy, and it delivered very attractive risk-adjusted returns.

In fact, one way to think of the alpha versus beta question is what we want to credit managers for as skill. Thirty years ago, simply noticing and then implementing merger arbitrage would seem to deserve tremendous credit, and recognizing its attractiveness deserved to be called skill. Nowadays, it seems far more likely that, in its simple form, it is a beta. More generally, over time alpha can turn into beta.

This last example brings us to another potential insight. Do these alphas and betas last forever? Well, it is hard to argue that, in a relentlessly competitive market, true alpha can last forever. In that sense, beta, while certainly less glamorous, has something going for it. If beta is rewarded with a positive but modest expected return because in a rational market it represents a risk some investors will pay others to bear for them, or liquidity some investors are willing to pay others to provide, it can last forever.

So, stepping back, why is this split into alpha and beta useful? Well, first, it demands that we answer the question: "Why does this strategy make money?" Second, it has direct implications for portfolio construction and correlation. That is, if you do not realize that your collection of hedge fund managers is largely about a similar beta exposure, it is easy to underestimate how the funds will move together in a crisis. Third, it helps in

understanding hedge fund fees. What you should pay for real alpha is different from what you should pay for hedge fund beta, which in turn is different from what you should pay for traditional beta. Asness [2004] explores the issue of fees in more depth.

Conclusion

Allocating to hedge funds will not solve the problem that there is a low equity-risk premium, or that bonds offer low real interest rates, and certainly not that cash rates are infinitesimal. The hedge fund structure does not create investing skill out of thin air, where other traditional structures have failed to do so, and the tools that hedge funds use (leverage, short-selling and derivatives) certainly come with risk.

That said, the potential benefits of hedge fund investing are real. To the extent investing skill exists, the hedge fund structure allows that skill to be offered alone without the tie-in sale of a traditional index fund. In addition, skill applied in an unconstrained hedge fund structure can produce a better outcome than more constrained traditional formats. The hedge fund structure can serve to clarify risk, performance attribution, and what you're paying for versus traditional active management. Hedge funds can also work to make the market more efficient by aligning prices better with reality.

Hedge funds also, through a set of reasonably well-defined systematic strategies (hedge fund betas), allow liquidity to be provided to those who need it from those who have it, and risk to be transferred from those who do not want it to those who do. Many of these strategies would not be possible without the tools of leverage, short-selling, and derivatives common to hedge funds.

However, the story is not over as all is not wine and roses in the hedge fund world. In Asness [2004] I will explore some dark sides to hedge fund investing that must be addressed, and some less dramatic but important evolutionary changes that are needed for hedge funds to fully fulfill the potential I have outlined here.

Endnotes

The author thanks Gregor Andrade, Kent Clark, Thomas Dunn, Ira Handler, David Kabiller, Robert Krail, John Liew, Mark Mitchell, and Todd Pulvino for very helpful comments on this manuscript, as well as Jerome Baesel, Trey Beck, Anne Casscells, George Main, Laurence Siegel, and Todd Tibbets for many conversations over the years that have greatly aided his understanding of this topic.

The views and opinions expressed here are those of the author and do not necessarily reflect the views of his employer, its affiliates, or its employees. The information has been obtained or derived from sources believed by the author to be reliable, but the author does not make any representation or warranty, express or implied, as to the information's accuracy or completeness, nor does the author recommend that the information serve as the basis for any investment decision.

[1]For instance, see Clarke, de Silva, and Thorley [2002].

[2]Space constraints did not permit addressing all these issues in one article. Readers who wish to avoid becoming too elated or too depressed about hedge funds should probably read them at one time and in sequence.

[3]For simplicity assume the risk exposures are the same for the manager and the index.

[4]See Asness [1998] for more details on this approach.

[5]This analysis assumes that market risk is the only priced risk factor (i.e., a CAPM type world). I am not arguing this is the case, but it suffices for this example. When recombined with an index (often a different one from the original I) this type of pure skill is often called portable alpha.

[6]More generally, Roll [1992] shows that the traditional method, first deciding on a set of index exposures, and then having a set of active managers maximize risk-adjusted return versus the benchmark, does not generally lead to a mean-variance optimal outcome.

[7]Hedge funds traditionally charge a combination of fixed and performance fees, typically 1% fixed and 20% of some definition of the hedge fund's profits.

[8]Sometimes *active* is taken not as the opposite of *passive*, but also to mean a judgment style as contrasted with a quantitative one. What I mean by *active* is an attempt to add value. I have known some quants who are quite frenetic.

[9]Quite the oxymoron: random riskless profit. The other name for merger arbitrage is risk arbitrage, which translated literally means "risky riskless profit," not much better.

[10]The traditional active manager can participate in a stock-for-stock merger only by not owning the acquirer (so she is under weight the acquirer versus her benchmark) and buying the target. This is a very limited form of betting on the merger that is likely meaningful only when the acquirer is a large company, so the underweight from not owning it is significant.

[11]A convertible bond can be cheap for a variety of reasons. Credit and the implied volatility of the option are the two most likely candidates.

[12]Note one potentially confusing point. Calling systematic hedge fund strategies "beta," and distinguishing this from true alpha, does not mean these hedge fund betas are not value-added propositions. Adding hedge fund beta to traditional portfolios in all likelihood increases their Sharpe ratios, and hedge fund beta requires a far greater variety of skills to implement than do traditional index funds. This itself might be a working definition of alpha in a different sense.

[13]Structural reasons might include common institutional investor guidelines that starve one area of the market for capital, like ratings or maturity constraints on short-selling, the presence of non-economic actors in some markets (e.g., central banks), and many others. Biases include things like the home bias (the fact that many investors overweight assets that are familiar to them), herding behavior (wanting to do what the crowd does), overextrapolation (assuming what's been happening will keep happening), and general myopia (too much of a focus on the short term).

[14]Large-scale mergers were also a relatively new phenomenon starting in the 1960s, raising the possibility that ex ante managers at the time thought this strategy was much more volitile and its Sharpe ratio lower (meaning they did not think it was a great alpha strategy, as it was later shown to be).

References

Asness, Clifford. "An Alternative Future: Part II." *The Journal of Portfolio Management*, forthcoming, Fall 2004.

_____. "Hedge Fund Alphas and Betas." Presentation to NMS Management's Annual Endowments & Foundations Conference, January 26, 2003.

_____. "Market-Neutral Investing: Putting the 'Hedge' in Hedge Funds." Working paper, AQR Capital Management, 1998.

Clarke, Roger, Harindra de Silva, and Steven Thorley. "Portfolio Constraints and the Fundamental Law of Active Management." *Financial Analysts Journal*, September/October 2002.

Dunn, Thomas. "ABP Hedge Fund Investments." Presentation at the University of Chicago Graduate School of Business, May 14, 2004.

Fama, Eugene F., and Kenneth R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993).

_____. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance,* March 1996.

Jensen, Greg, and Jason Rotenberg. "The Flood Into Hedge Funds." *Bridgewater Daily Observations,* May 9, 2003.

Mitchell, Mark, and Todd Pulvino. "Characteristics of Risk and Return in Risk Arbitrage." *Journal of Finance,* 56 (6) (2000), pp. 2135–2175.

Roll, Richard. "A Mean/Variance Analysis of Tracking Error." *The Journal of Portfolio Management,* Summer 1992.

Siegel, Laurence B. "The Risk of Paying Alpha Fees for Beta Performance." Presentation to PRMIA, August 26, 2004.

This page intentionally left blank

AQR

TWENTY YEARS

EST. 1998

1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

REFLECTIONS

Our twentieth anniversary is an opportunity to recognize our partnership with our clients, how far we have come as a firm, and all that we have ahead of us.

THE BEGINNING

Applying academia

1989-1997

AQR's story begins at the University of Chicago, where the foundation for the firm's philosophy was established.

Cliff Asness, in his second year of the PhD program, earned the coveted appointment of teaching and research assistant to Gene Fama, known for market efficiency research. For his dissertation, however, Cliff chose to focus on momentum investing, a challenge to market efficiency. It was in Fama's class that Cliff met his future co-founders, John Liew and Robert Krail, both PhD students at the time.

While still working on his dissertation, Cliff joined Goldman Sachs for what was to be a yearlong assignment, completing his PhD on nights and weekends. When the year came to a close, he was asked to lead a new quantitative research team for Goldman Sachs Asset Management.

John Liew and Robert Krail joined him, and the new team applied what they learned in academia to their investment philosophy. Based on the strength of their research, Goldman seeded a hedge fund. Long-only strategies and mutual funds utilizing the same investment process followed shortly thereafter.

Around the same time, David Kabiller, who worked closely with the team in his role overseeing relationships with some of Goldman Sachs' largest pension and investment funds, co-authored "Hedge Funds Demystified: Their Potential Role in Institutional Portfolios." The paper made the connection between hedge funds and non-correlated sources of returns — and the diversification benefits for institutional investors.

David believed — and eventually convinced his partners — that starting a new firm was the path forward.

"You would be surprised how much effort it takes to turn the academic into the practical. Managing money for clients involves controlling costs, managing risk and building real portfolios. Academic work is just the tip of the iceberg."

— John Liew

THE LAUNCH

A promising start

1998

With a strong partnership and a successful track record, the founders established AQR Capital Management in 1998. The name was chosen for a very simple reason: Applied Quantitative Research was exactly what the team did, and both the applied side and the research side were of equal importance.

The firm received nearly $2 billion in subscriptions, though only took investments of $1 billion, which was one of the largest hedge fund launches at the time.

AQR's Absolute Return strategy launched on August 3, 1998 — the start of a particularly bad month for markets. The Russian government defaulted, Long Term Capital Management blew up and the MSCI World Index fell -13.4%.

Cliff Asness writing during the early days of the firm.

The firm was unaffected, and at the outset, it looked like the strong track record the team had experienced for years before would continue.

AQR's Founders: Robert Krail, David Kabiller, Cliff Asness, John Liew

From the very beginning, AQR differentiated itself from others in the industry. The firm was one of the first hedge fund managers to voluntarily register with the SEC.

It was a choice that was made because the founders believed in the importance of holding the firm to the highest ethical standards.

THE EARLY YEARS

Trusting the process, staying the course

1999-2000

Within just a few months of AQR's launch, the tech bubble began.

While it can be said that all value investors were in a lot of pain during this time, for AQR's fund, a levered, market-neutral value-oriented strategy, the losses were far worse.

The fund's performance fell by nearly 40% and AUM was halved. The firm held fast to its investment philosophy and was open and transparent with clients throughout this difficult period. AQR emerged from the tech bubble with a "depression-era" perspective, with many lessons learned that ultimately shaped the future investment process.

The firm enhanced its risk management, improved drawdown controls and continued to evolve its model's signals.

The firm's belief in diversification was only strengthened. In 2000, AQR launched its first long-only product, an international equity fund, to complement the higher volatility hedge fund.

"The worst times in your life can sometimes end up as the times you look back on with the most pride. The tech bubble began in earnest nearly to the day we started AQR.

I'm proud of sticking with our philosophy. I'm proud of how hard, and how successfully, we fought to keep our clients."

— Cliff Asness

AQR's Founders and earliest employees at the firm's first office in New York City.

TOP ROW:
Cliff Asness, Jacques Friedman, Robert Krail, John Liew

SECOND ROW:
David Kabiller, Ronen Israel, Lars Nielsen

THIRD ROW:
Oktay Kurbanov, Brian Hurst

BOTTOM ROW:
John Liew, Brad Asness

THE CASE FOR DIVERSIFICATION

Building a lasting business

2004-2010

By 2004, the firm was managing over $12 billion in assets — approximately half in hedge funds and half in long-only strategies — at a time when managing both was somewhat unusual.

That same year, the firm moved from New York City to its current headquarters in Greenwich, Connecticut, and soon started expanding globally with the opening of its first international office in Australia in 2005. Even as the firm continued to grow and expand, there were some headwinds. In 2007, the quant crisis weakened or shuttered many quantitative funds. While AQR emerged from this short-lived episode, the global financial crisis of 2008 followed, which shook the entire industry.

In the wake of the crisis, the firm became even more resolute in its belief in the importance of diversification. AQR recognized that many individual investors were not properly diversified going into the crisis and, as a result, suffered significant losses.

AQR's investment team hosting educational events for clients.

In 2009, the firm became one of the first alternative investment managers to offer mutual funds, giving individual investors access to the same set of diversifying strategies previously offered only to institutional investors.

Throughout this period, AQR broadened its product offerings, introducing some of the firm's signature strategies. The firm's first risk parity fund launched in 2006. DELTA, which combines a mix of classic hedge fund styles, launched in 2008, while AQR's first Managed Futures fund launched in 2009.

Together these strategies represent the firm's innovative approach to product development — backed by rigorous research and designed to solve clients' investment challenges.

As AQR expanded its solutions, the firm invested time and resources in helping clients better understand its strategies and the broader investment landscape.

By 2010, the firm was managing over $33 billion in assets for a growing client base around the world.

ACADEMIC ENGAGEMENT

Exchanging ideas and advancing financial knowledge

2011-2015

Whether through publishing research or teaching at top universities, applied academia is a core part of AQR's DNA. In support of new ideas and the next generation of researchers, the firm established a series of prestigious academic awards and partnerships.

The firm introduced the AQR Insight Award in 2011, which honors exceptional academic working papers with an annual $100,000 prize, embodying the firm's culture of intellectual curiosity and application of innovative research ideas.

A year later, the AQR Top Finance Graduate Award at Copenhagen Business School was founded to recognize PhD graduates whose dissertation and research carry the greatest potential impact in both practice and academia. And in 2015, the AQR Asset Management Institute at the London Business School was established to promote excellence in asset management with awards that recognize young research talent, and events where practitioners, policymakers and academics debate, create and share ideas.

With 74 PhDs and 19 current or former university professors, the firm resembles a leading academic finance department.

AQR Principals and professors: Toby Moskowitz (right) presenting at a client event and Andrea Frazzini (below) introducing the AQR Insight Award presentations.

Our research has been recognized with over 50 awards and is among the most frequently cited in top journals, alongside academic institutions such as the University of Chicago, Yale University and University of Pennsylvania.

As of June 30, 2018. Source: AQR and ssrn.com

FOCUS ON LIFELONG LEARNING

Educating and inspiring employees and clients

2015-2018

In 2015, AQR launched QUANTA Academy, the firm's learning and development program. The curriculum takes a differentiated, holistic approach, offering employees over 350 classes and events across three core pillars — technical skills and knowledge, leadership and management, and personal enrichment.

Employees gain expertise and perspective through experiences such as the Insights Book Club, where discussions of great books are led by renowned authors and professors, or Tech Talks, where leaders, innovators and disrupters discuss trends and developments in technology.

Through QUANTA, every individual has the opportunity to become the best version of themselves because the firm believes that when employees reach their full potential, clients are best served. Recognizing the benefits QUANTA offers both professionally and personally, in 2018, AQR began extending the program to clients.

"Learning is a journey we embark upon together. It helps us grow as individuals and as a firm, and creates a rich culture of innovation and excellence."

— David Kabiller

An AQR professional participating in a QU∀NTA session.

What is QU∀NTA?

QU∀NTA, the plural of quantum, is the smallest unit of many forms of energy. The name represents continuous learning through countless small interactions and formal programs. The inverted A — a formal logic symbol meaning "for all" — represents learning as a shared value across the firm.

REFLECTING ON 20 YEARS

Our culture of excellence and unwavering commitment to clients

Today, AQR is at the nexus of data, technology, economics and statistics. We offer a broad, diversified set of strategies based on a unified set of underlying principles.

We continue to challenge the status quo and apply rigorous research to our investment process. We have built a strong culture of excellence with a relentless commitment to our clients. We pursue better solutions to meet our clients' objectives and make their challenges our most exciting opportunities.

We are proud of the lasting relationships we have built through this approach: 20 years later, many of our first investors remain clients of AQR.

We are grateful for the past 20 years and look forward to what the next two decades will bring.

— Cliff, David and John

- OVER 900 EMPLOYEES
- 8 OFFICES AROUND THE GLOBE
- $226 BILLION IN ASSETS UNDER MANAGEMENT
- CLIENTS REPRESENTING SOME OF THE LARGEST AND MOST SOPHISTICATED INVESTORS AROUND THE WORLD
- SOLUTIONS ACROSS ASSET CLASSES, SPANNING LONG-ONLY AND ALTERNATIVE INVESTMENTS

TOP:
Co-founder David Kabiller in discussion with Principal Ronen Israel.

OTHER:
AQR employees collaborating across the firm.

As of June 30, 2018. Includes assets managed by AQR and its advisory affiliates.

AQR

TWENTY
YEARS

EST. 1998

PART V

Pioneering Factor Investing

Time Series Momentum

Value and Momentum Everywhere

Betting Against Beta

Common Factors in Corporate Bond Returns

Size Matters If You Control Your Junk

This page intentionally left blank

Time Series Momentum*

Tobias J. Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen

We document significant "time series momentum" in equity index, currency, commodity, and bond futures for *each* of the 58 liquid instruments we consider. We find persistence in returns for one to 12 months that partially reverses over longer horizons, consistent with sentiment theories of initial under-reaction and delayed over-reaction. A diversified portfolio of time series momentum strategies across all asset classes delivers substantial abnormal returns with little exposure to standard asset pricing factors and performs best during extreme markets. Examining the trading activities of speculators and hedgers, we find that speculators profit from time series momentum at the expense of hedgers.

1. Introduction: A Trending Walk Down Wall Street

We document an asset pricing anomaly we term "time series momentum," which is remarkably consistent across very different asset classes and markets. Specifically, we find strong positive predictability from a security's own past returns for almost five dozen diverse futures and forward contracts that include country equity indexes, currencies, commodities, and sovereign bonds over more than 25 years of data. We find that the past 12-month excess return of each instrument is a positive predictor of its future return. This

*We thank Cliff Asness, Nick Barberis, Gene Fama, John Heaton, Ludger Hentschel, Brian Hurst, Andrew Karolyi, John Liew, Matt Richardson, Richard Thaler, Adrien Verdelhan, Robert Vishny, Robert Whitelaw, Jeff Wurgler, and seminar participants at NYU and the 2011 AFA meetings in Denver, CO for useful suggestions and discussions, and Ari Levine and Haibo Lu for excellent research assistance. Moskowitz thanks the Initiative on Global Markets at the University of Chicago Booth School of Business and CRSP for financial support.

Reprinted from *The Journal of Financial Economics*, May 2012, Vol. 104 (2), with permission.

Recipient of the 2013 Whitebox Prize.

time series momentum or "trend" effect persists for about a year and then partially reverses over longer horizons. These findings are robust across a number of subsamples, look-back periods, and holding periods. We find that 12-month time series momentum profits are positive, not just on average across these assets, but for *every* asset contract we examine (58 in total).

Time series momentum is related to, but different from, the phenomenon known as "momentum" in the finance literature, which is primarily cross-sectional in nature. The momentum literature focuses on the *relative* performance of securities in the *cross-section,* finding that securities that recently outperformed their peers over the past three to 12 months continue to outperform their peers on average over the next month.[1] Rather than focus on the relative returns of securities in the cross-section, time series momentum focuses purely on a security's *own* past return.

We argue that time series momentum directly matches the predictions of many prominent behavioral and rational asset pricing theories. Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999) all focus on a single risky asset, therefore having direct implications for time series, rather than cross-sectional, predictability. Likewise, rational theories of momentum (Berk, Green, and Naik, 1999; Johnson, 2002; Ahn, Conrad, and Dittmar, 2003; Liu and Zhang, 2008; Sagi and Seasholes, 2007) also pertain to a single risky asset.

Our finding of positive time series momentum that partially reverses over the long-term may be consistent with initial under-reaction and delayed over-reaction, which theories of sentiment suggest can produce these return patterns.[2] However, our results also pose several challenges to these theories. First, we find that the correlations of time series momentum strategies across asset classes are larger than the correlations of the asset classes themselves. This suggests a stronger common component to time series momentum across different assets than is present among the assets themselves. Such a correlation structure is not addressed by existing behavioral models. Second, very different types of investors in different asset markets are producing the same patterns at the same time. Third, we fail to find a link between time series momentum and measures of investor sentiment used in the literature (Baker and Wurgler, 2006; Qiu and Welch, 2006).

To understand the relationship between time series and cross-sectional momentum, their underlying drivers, and relation to theory, we decompose the returns to a time series and cross-sectional momentum strategy following the framework of Lo and Mackinlay (1990) and Lewellen (2002). This decomposition allows us to identify the properties of returns that contribute to these patterns, and what features are common and unique to the two strategies. We find that positive auto-covariance in futures contracts' returns drives most of the time series and cross-sectional momentum effects we find in the data. The contribution of the other two return components—serial cross-correlations and variation in mean returns—is small. In fact, negative serial cross-correlations (i.e., lead-lag effects across securities), which affect cross-sectional momentum, are negligible and of the "wrong" sign among our instruments to explain time series momentum. Our finding that time series and cross-sectional momentum profits arise due to auto-covariances is consistent with the theories mentioned above.[3] In addition, we find that time series momentum captures the returns associated with individual stock (cross-sectional) momentum, most notably Fama and French's UMD factor, despite time series momentum being constructed from a completely different set of securities. This finding indicates strong correlation structure between time

series momentum and cross-sectional momentum even when applied to different assets and suggests that our time series momentum portfolio captures individual stock momentum.

To better understand what might be driving time series momentum, we examine the trading activity of speculators and hedgers around these return patterns using weekly position data from the Commodity Futures Trading Commission (CFTC). We find that speculators trade with time series momentum, being positioned, on average, to take advantage of the positive trend in returns for the first 12 months and reducing their positions when the trend begins to reverse. Consequently, speculators appear to be profiting from time series momentum at the expense of hedgers. Using a vector auto-regression (VAR), we confirm that speculators trade in the same direction as a return shock and reduce their positions as the shock dissipates, whereas hedgers take the opposite side of these trades.

Finally, we decompose time series momentum into the component coming from spot price predictability versus the "roll yield" stemming from the shape of the futures curve. While spot price changes are mostly driven by information shocks, the roll yield can be driven by liquidity and price pressure effects in futures markets that affect the return to holding futures without necessarily changing the spot price. Hence, this decomposition may be a way to distinguish the effects of information dissemination from hedging pressure. We find that both of these effects contribute to time series momentum, but only spot price changes are associated with long-term reversals, consistent with the idea that investors may be over-reacting to information in the spot market but that hedging pressure is more long-lived and not affected by over-reaction.

Our finding of time series momentum in virtually every instrument we examine seems to challenge the "random walk" hypothesis, which in its most basic form implies that knowing whether a price went up or down in the past should not be informative about whether it will go up or down in the future. While rejection of the random walk hypothesis does not necessarily imply a rejection of a more sophisticated notion of market efficiency with time-varying risk premiums, we further show that a diversified portfolio of time series momentum across all assets is remarkably stable and robust, yielding a Sharpe ratio greater than one on an annual basis, or roughly 2.5 times the Sharpe ratio for the equity market portfolio, with little correlation to passive benchmarks in each asset class or a host of standard asset pricing factors. The abnormal returns to time series momentum also do not appear to be compensation for crash risk or tail events. Rather, the return to time series momentum tends to be largest when the stock market's returns are most extreme—performing best when the market experiences large up and down moves. Hence, time series momentum may be a hedge for extreme events, making its large return premium even more puzzling from a risk-based perspective. The robustness of time series momentum for very different asset classes and markets suggest that our results are not likely spurious, and the relatively short duration of the predictability (less than a year) and the magnitude of the return premium associated with time series momentum present significant challenges to the random walk hypothesis and perhaps also to the efficient market hypothesis, though we cannot rule out the existence of a rational theory that can explain these findings.

Our study relates to the literature on return autocorrelation and variance ratios that also finds deviations from the random walk hypothesis (Fama and French, 1988; Lo and Mackinlay, 1988; Poterba and Summers, 1988). While this literature is largely focused on US and global equities, Cutler, Poterba, and Summers (1991) study a variety of assets including housing and collectibles. The literature finds positive return autocorrelations at

daily, weekly, and monthly horizons and negative autocorrelations at annual and multi-year frequencies. We complement this literature in several ways. The studies of autocorrelation examine, by definition, return predictability where the length of the "look-back period" is the same as the "holding period" over which returns are predicted. This restriction masks significant predictability that is uncovered once look-back periods are allowed to differ from predicted or holding periods. In particular, our result that the past 12 *months* of returns strongly predicts returns over the next one *month* is missed by looking at one-year autocorrelations. While return continuation can also be detected implicitly from variance ratios, we complement the literature by explicitly documenting the extent of return continuation and by constructing a time series momentum factor that can help explain existing asset pricing phenomena, such as cross-sectional momentum premiums and hedge fund macro and managed futures returns. Also, a significant component of the higher frequency findings in equities is contaminated by market microstructure effects such as stale prices (Richardson, 1993; Ahn, Boudoukh, Richardson, and Whitelaw, 2002). Focusing on liquid futures instead of individual stocks and looking at lower frequency data mitigates many of these issues. Finally, unique to this literature, we link time series predictability to the dynamics of hedger and speculator positions and decompose returns into price changes and roll yields.

Our paper is also related to the literature on hedging pressure in commodity futures (Keynes, 1923; Fama and French, 1987; Bessembinder, 1992; de Roon, Nijman, and Veld, 2000). We complement this literature by showing how hedger and speculator positions relate to past futures returns (and not just in commodities), finding that speculators' positions load positively on time series momentum, while hedger positions load negatively on it. Also, we consider the relative return predictability of positions, past price changes, and past roll yields. Gorton, Hayashi, and Rouwenhorst (2008) also link commodity momentum and speculator positions to the commodities' inventories.

The rest of the paper is organized as follows. Section 2 describes our data on futures returns and the positioning of hedgers and speculators. Section 3 documents time series momentum at horizons less than a year and reversals beyond that. Section 4 defines a time series momentum factor, studying its relation to other known return factors, its performance during extreme markets, and correlations within and across asset classes. Section 5 examines the relation between time series and cross-sectional momentum, showing how time series momentum is a central driver of cross-sectional momentum as well as macro and managed futures hedge fund returns. Section 6 studies the evolution of time series momentum and its relation to investor speculative and hedging positions. Section 7 concludes.

2. Data and Preliminaries

We describe briefly the various data sources we use in our analysis.

2.1. Futures Returns

Our data consist of futures prices for 24 commodities, 12 cross-currency pairs (from nine underlying currencies), nine developed equity indexes, and 13 developed government bond

futures, from January 1965 through December 2009. These instruments are among the most liquid futures contracts in the world.[4] We focus on the most liquid instruments to avoid returns being contaminated by illiquidity or stale price issues and to match more closely an implementable strategy at a significant trade size. Appendix A provides details on each instrument and their data sources, which are mainly Datastream, Bloomberg, and various exchanges.

We construct a return series for each instrument as follows. Each day, we compute the daily excess return of the most liquid futures contract (typically the nearest or next nearest-to-delivery contract), and then compound the daily returns to a cumulative return index from which we can compute returns at any horizon. For the equity indexes, our return series are almost perfectly correlated with the corresponding returns of the underlying cash indexes in excess of the Treasury bill rate.[5]

As a robustness test, we also use the "far" futures contract (the next maturity after the most liquid one). For the commodity futures, time series momentum profits are in fact slightly stronger for the far contract, and, for the financial futures, time series momentum returns hardly change if we use far futures.

Table 1 presents summary statistics of the excess returns on our futures contracts. The first column reports when the time series of returns for each asset starts, and the next two columns report the time series mean (arithmetic) and standard deviation (annualized) of each contract by asset class: commodities, equity indexes, bonds, and currencies. As Table 1 highlights, there is significant variation in sample mean returns across the different contracts. Equity index, bonds, and currencies yield predominantly positive excess returns, while various commodity contracts yield positive, zero, and even negative excess average returns over the sample period. Only the equity and bond futures exhibit statistically significant and consistent positive excess average returns.

More striking are the differences in volatilities across the contracts. Not surprisingly, commodities and equities have much larger volatilities than bond futures or currency forward contracts. But, even among commodities, there is substantial cross-sectional variation in volatilities. Making comparisons across instruments with vastly different volatilities or combining various instruments into a diversified portfolio when they have wide-ranging volatilities is challenging. For example, the volatility of natural gas futures is about 50 times larger than that of 2-year US bond futures. We discuss below how we deal with this issue in our analysis.

2.2. Positions of Traders

We also use data on the positions of speculators and hedgers from the Commodity Futures Trading Commission (CFTC) as detailed in Appendix A. The CFTC requires all large traders to identify themselves as commercial or non-commercial which we, and the previous literature (e.g., Bessembinder, 1992; de Roon, Nijman, and Veld, 2000), refer to as hedgers and speculators, respectively. For each futures contract, the long and short open interest held by these traders on Tuesday are reported on a weekly basis.[6]

Using the positions of speculators and hedgers as defined by the CFTC, we define the Net speculator position for each asset as follows:

$$\text{Net speculator position} = \frac{\text{Speculator long positions} - \text{Speculator short positions}}{\text{Open interest}}$$

Table 1 Summary statistics on futures contracts.

Reported are the annualized mean return and volatility (standard deviation) of the futures contracts in our sample from January 1965 to December 2009 as well as the mean and standard deviation of the Net speculator long positions in each contract as a percentage of open interest, covered and defined by the CFTC data, which are available over the period January 1986 to December 2009. For a detailed description of our sample of futures contracts, see Appendix A.

	Data start date	Annualized mean	Annualized volatility	Average net speculator long positions	Std. dev. net speculator long positions
Commodity futures					
ALUMINUM	Jan-79	0.97%	23.50%		
BRENTOIL	Apr-89	13.87%	32.51%		
CATTLE	Jan-65	4.52%	17.14%	8.1%	9.6%
COCOA	Jan-65	5.61%	32.38%	4.9%	14.0%
COFFEE	Mar-74	5.72%	38.62%	7.5%	13.6%
COPPER	Jan-77	8.90%	27.39%		
CORN	Jan-65	–3.19%	24.37%	7.1%	11.0%
COTTON	Aug-67	1.41%	24.35%	–0.1%	19.4%
CRUDE	Mar-83	11.61%	34.72%	1.0%	5.9%
GASOIL	Oct-84	11.95%	33.18%		
GOLD	Dec-69	5.36%	21.37%	6.7%	23.0%
HEATOIL	Dec-78	9.79%	33.78%	2.4%	6.4%
HOGS	Feb-66	3.39%	26.01%	5.1%	14.5%
NATGAS	Apr-90	–9.74%	53.30%	–1.6%	8.9%
NICKEL	Jan-93	12.69%	35.76%		
PLATINUM	Jan-92	13.15%	20.95%		
SILVER	Jan-65	3.17%	31.11%	20.6%	14.3%
SOYBEANS	Jan-65	5.57%	27.26%	8.2%	12.8%
SOYMEAL	Sep-83	6.14%	24.59%	6.7%	11.2%
SOYOIL	Oct-90	1.07%	25.39%	5.7%	12.8%
SUGAR	Jan-65	4.44%	42.87%	10.0%	14.2%
UNLEADED	Dec-84	15.92%	37.36%	7.8%	9.6%
WHEAT	Jan-65	–1.84%	25.11%	4.3%	12.1%
ZINC	Jan-91	1.98%	24.76%		
Equity index futures					
ASX SPI 200 (AUS)	Jan-77	7.25%	18.33%		
DAX (GER)	Jan-75	6.33%	20.41%		
IBEX 35 (ESP)	Jan-80	9.37%	21.84%		
CAC 40 10 (FR)	Jan-75	6.73%	20.87%		
FTSE/MIB (IT)	Jun-78	6.13%	24.59%		
TOPIX (JP)	Jul-76	2.29%	18.66%		
TOPIX (JP)	Jul-76	2.29%	18.66%		

(*Continued*)

Table 1 (*Continued*)

	Data start date	Annualized mean	Annualized volatility	Average net speculator long positions	Std. dev. net speculator long positions
AEX (NL)	Jan-75	7.72%	19.18%		
FTSE 100 (UK)	Jan-75	6.97%	17.77%		
S&P 500 (US)	Jan-65	3.47%	15.45%	–4.6%	5.4%
Bond futures					
3-year AUS	Jan-92	1.34%	2.57%		
10-year AUS	Dec-85	3.83%	8.53%		
2-year EURO	Mar-97	1.02%	1.53%		
5-year EURO	Jan-93	2.56%	3.22%		
10-year EURO	Dec-79	2.40%	5.74%		
30-year EURO	Dec-98	4.71%	11.70%		
10-year CAN	Dec-84	4.04%	7.36%		
10-year JP	Dec-81	3.66%	5.40%		
10-year UK	Dec-79	3.00%	9.12%		
2-year US	Apr-96	1.65%	1.86%	1.9%	11.3%
5-year US	Jan-90	3.17%	4.25%	3.0%	9.2%
10-year US	Dec-79	3.80%	9.30%	0.4%	8.0%
30-year US	Jan-90	9.50%	18.56%	–1.4%	6.2%
Currency forwards					
AUD/USD	Mar-72	1.85%	10.86%	12.4%	28.8%
EUR/USD	Sep-71	1.57%	11.21%	12.1%	18.7%
CAD/USD	Mar-72	0.60%	6.29%	4.7%	24.1%
JPY/USD	Sep-71	1.35%	11.66%	–6.0%	23.8%
NOK/USD	Feb-78	1.37%	10.56%		
NZD/USD	Feb-78	2.31%	12.01%	38.8%	33.8%
SEK/USD	Feb-78	–0.05%	11.06%		
CHF/USD	Sep-71	1.34%	12.33%	–5.2%	26.8%
GBP/USD	Sep-71	1.39%	10.32%	2.7%	25.4%

This signed measure shows whether speculators are net long or short in aggregate, and scales their net position by the open interest or total number of contracts outstanding in that futures market. Since speculators and hedgers approximately add up to zero (except for a small difference denoted "non-reported" due to measurement issues of very small traders), we focus our attention on speculators. Of course, this means that net hedger positions constitute the opposite side (i.e., the negative of Net speculator position).

The CFTC positions data do not cover all of the futures contracts we have returns for and consider in our analysis. Most commodity and foreign exchange contracts are covered, but only the US instruments among the stock and bond futures contracts are covered.

The third and fourth columns of Table 1 report summary statistics on the sample of futures contracts with Net speculator positions in each contract over time. Speculators are net long, on average, and hence hedgers are net short, for most of the contracts, a result consistent with Bessembinder (1992) and de Roon, Nijman, and Veld (2000) for a smaller set of contracts over a shorter time period. All but two of the commodities (natural gas and cotton) have net long speculator positions over the sample period, with silver exhibiting the largest average net long speculator position. This is consistent with Keynes' (1923) conjecture that producers of commodities are the primary hedgers in markets and are on the short side of these contracts as a result. For the other asset classes, other than the S&P 500, the 30-year US Treasury bond, and the $US/Japanese and $US/Swiss exchange rates, speculators exhibit net long positions, on average. Table 1 also highlights that there is substantial variation over time in Net speculator positions per contract and across contracts. Not surprisingly, the standard deviation of Net speculator positions is positively related to the volatility of the futures contract itself.

2.3. Asset Pricing Benchmarks

We evaluate the returns of our strategies relative to standard asset pricing benchmarks, namely the MSCI World equity index, Barclay's Aggregate Bond Index, S&P GSCI Index, all of which we obtain from Datastream, the long-short factors *SMB, HML,* and *UMD* from Ken French's Web site, and the long-short value and cross-sectional momentum factors across asset classes from Asness, Moskowitz, and Pedersen (2010).

2.4. Ex Ante Volatility Estimate

Since volatility varies dramatically across our assets (illustrated in Table 1), we scale the returns by their volatilities in order to make meaningful comparisons across assets. We estimate each instrument's ex ante volatility σt at each point in time using an extremely simple model: the exponentially weighted lagged squared daily returns (i.e., similar to a simple univariate GARCH model). Specifically, the ex ante annualized variance σ_t^2 for each instrument is calculated as follows:

$$\sigma_t^2 = 261\sum_{i=0}^{\infty}(1-\delta)\delta^i(r_{t-1-i}-\overline{r}_t)^2, \tag{1}$$

where the scalar 261 scales the variance to be annual, the weights $(1-\delta)\delta i$ add up to one, and $\overline{r}_t$ is the exponentially weighted average return computed similarly. The parameter δ is chosen so that the center of mass of the weights is $\sum_{i=0}^{\infty}(1-\delta)\delta^i i = \delta/(1-\delta) = 60$ days. The volatility model is the same for all assets at all times. While all of the results in the paper are robust to more sophisticated volatility models, we chose this model due to its simplicity and lack of look-ahead bias in the volatility estimate. To ensure no look-ahead bias contaminates our results, we use the volatility estimates at time t–1 applied to time-t returns throughout the analysis.

3. Time Series Momentum: Regression Analysis and Trading Strategies

We start by examining the time series predictability of futures returns across different time horizons.

3.1. Regression Analysis: Predicting Price Continuation and Reversal

We regress the excess return r_t^s for instrument s in month t on its return lagged h months, where both returns are scaled by their ex ante volatilities σ_{t-1}^s (defined above in Section 2.4):

$$r_t^s / \sigma_{t-1}^s = \alpha + \beta_h r_{t-h}^s / \sigma_{t-h-1}^s + \varepsilon_t^s. \tag{2}$$

Given the vast differences in volatilities (as shown in Table 1), we divide all returns by their volatility to put them on the same scale. This is similar to using Generalized Least Squares instead of Ordinary Least Squares (OLS).[7] Stacking all futures contracts and dates, we run a pooled panel regression and compute t-statistics that account for group-wise clustering by time (at the monthly level). The regressions are run using lags of $h = 1, 2, \ldots, 60$ months.

Panel A of Fig. 1 plots the t-statistics from the pooled regressions by month lag h. The positive t-statistics for the first 12 months indicate significant return continuation or trends. The negative signs for the longer horizons indicate reversals, the most significant of which occur in the year immediately following the positive trend.

Another way to look at time series predictability is to simply focus only on the *sign* of the past excess return. This even simpler way of looking at time series momentum underlies the trading strategies we consider in the next section. In a regression setting, this strategy can be captured using the following specification:

$$r_t^s / \sigma_{t-1}^s = \alpha + \beta_h \, sign\left(r_{t-h}^s\right) + \varepsilon_t^s. \tag{3}$$

We again make the left-hand side of the regression independent of volatility (the right-hand side is too since *sign* is either +1 or –1), so that the parameter estimates are comparable across instruments. We report the t-statistics from a pooled regression with standard errors clustered by time (i.e., month) in Panel B of Fig. 1.

The results are similar across the two regression specifications: strong return continuation for the first year and weaker reversals for the next 4 years. In both cases, the data exhibit a clear pattern, with all of the most recent 12-month lag returns positive (and nine statistically significant) and the majority of the remaining lags negative. Repeating the panel regressions for each asset class separately, we obtain the same patterns: one to 12-month positive time series momentum followed by smaller reversals over the next 4 years as seen in Panel C of Fig. 1.

Figure 1. Time series predictability across all asset classes. We regress the monthly excess return of each contract on its own lagged excess return over various horizons. Panel A uses the size of the lagged excess return as a predictor, where returns are scaled by their ex ante volatility to make them comparable across assets, Panel B uses the sign of the lagged excess return as a predictor, where the dependent variable is scaled by its ex ante volatility to make the regression coefficients comparable across different assets, and Panel C reports the results of the sign regression by asset class. Reported are the pooled regression estimates across all instruments with *t*-statistics computed using standard errors that are clustered by time (month). Sample period is January 1985 to December 2009. (A) *Panel A*: $r_t^s/\sigma_{t-1}^s = \alpha + \beta_h r_{t-h}^s/\sigma_{t-h-1}^s + \varepsilon_t^s$; (B) *Panel B*: $r_t^s/\sigma_{t-1}^s = \alpha + \beta_h sign\left(r_{t-h}^s\right) + \varepsilon_t^s$; (C) *Panel C*: Results by asset class.

3.2. Time Series Momentum Trading Strategies

We next investigate the profitability of a number of trading strategies based on time series momentum. We vary both the number of months we lag returns to define the signal used to form the portfolio (the "look-back period") and the number of months we hold each portfolio after it has been formed (the "holding period").

For each instrument *s* and month *t,* we consider whether the excess return over the past *k* months is positive or negative and go long the contract if positive and short if negative, holding the position for *h* months. We set the position size to be inversely proportional to the instrument's ex ante volatility, $1/\sigma^{s}_{t-1}$, each month. Sizing each position in each strategy to have constant ex ante volatility is helpful for two reasons. First, it makes it easier to aggregate strategies across instruments with very different volatility levels. Second, it is helpful econometrically to have a time series with relatively stable volatility so that the strategy is not dominated by a few volatile periods.

For each trading strategy *(k,h),* we derive a *single* time series of monthly returns even if the holding period *h* is more than one month. Hence, we do not have overlapping observations. We derive this single time series of returns following the methodology used by Jegadeesh and Titman (1993): The return at time *t* represents the average return across *all* portfolios at that time, namely the return on the portfolio that was constructed last month, the month before that (and still held if the holding period *h* is greater than two), and so on for all currently "active" portfolios.

Specifically, for each instrument, we compute the time-*t* return based on the sign of the past return from time t – *k* – 1 to t – 1. We then compute the time-*t* return based on the sign of the past return from $t-k-2$ to $t-2$, and so on until we compute the time-*t* return based on the final past return that is still being used from $t-k-h$ to $t-h$. For each (*k*,*h*), we get a single time series of monthly returns by computing the average return of all of these *h* currently "active" portfolios (i.e., the portfolio that was just bought and those that were bought in the past and are still held). We average the returns across all instruments (or all instruments within an asset class), to obtain our time series momentum strategy returns, $r_t^{TSMOM(k,h)}$

To evaluate the abnormal performance of these strategies, we compute their alphas from the following regression:

$$\begin{aligned} r_t^{TSMOM(k,h)} &= \alpha + \beta_1 MKT_t + \beta_2 BOND_t + \beta_3 GSCI_t + sSMB_t \\ &\quad + hHML_t + mUMD_t + \varepsilon_t, \end{aligned} \tag{4}$$

where we control for passive exposures to the three major asset classes—the stock market *MKT,* proxied by the excess return on the MSCI World Index, the bond market *BOND,* proxied by the Barclays Aggregate Bond Index, the commodity market *GSCI,* proxied by the S&P GSCI Index—as well as the standard Fama-French stock market factors *SMB, HML,* and *UMD* for the size, value, and (cross-sectional) momentum premiums. For the evaluation of time series momentum strategies, we rely on the sample starting in 1985 to ensure that a comprehensive set of instruments have data (see Table 1) and that the markets had significant liquidity. We obtain similar (and generally more significant) results if older data are included going back to 1965, but given the more limited breadth and liquidity of the instruments during this time, we report results post-1985.

Table 2 shows the *t*-statistics of the estimated alphas for each asset class and across all assets. The existence and significance of time series momentum is robust across horizons and asset classes, particularly when the look-back and holding periods are 12 months or less. In addition, we confirm that the time series momentum results are almost identical if we use the cash indexes for the stock index futures. The other asset classes do not have cash indexes.

4. Time Series Momentum Factor

For a more in-depth analysis of time series momentum, we focus our attention on a single time series momentum strategy. Following the convention used in the cross-sectional momentum literature (and based on the results from Fig. 1 and Table 2), we focus on the properties of the 12-month time series momentum strategy with a 1-month holding period (e.g., $k = 12$ and $h = 1$), which we refer to simply as TSMOM.

4.1. TSMOM by Security and the Diversified TSMOM Factor

We start by looking at each instrument and asset separately and then pool all the assets together in a diversified TSMOM portfolio. We size each position (long or short) so that it has an ex ante annualized volatility of 40%. That is, the position size is chosen to be 40%/ $\sigma_{t\text{-}1}$ where $\sigma_{t\text{-}1}$ is the estimate of the ex ante volatility of the contract as described above. The choice of 40% is inconsequential, but it makes it easier to intuitively compare our portfolios to others in the literature. The 40% annual volatility is chosen because it is similar to the risk of an average individual stock, and when we average the return across all securities (equal-weighted) to form the portfolio of securities which represent our TSMOM factor, it has an annualized volatility of 12% per year over the sample period 1985–2009, which is roughly the level of volatility exhibited by other factors such as those of Fama and French (1993) and Asness, Moskowitz, and Pedersen (2010).[8] The TSMOM return for any instrument *s* at time *t* is therefore:

$$r_{t,t+1}^{TSMOM,s} = sign\left(r_{t-12,t}^{s}\right)\frac{40\%}{\sigma_t^s} r_{t,t+1}^{s}. \tag{5}$$

We compute this return for each instrument and each available month from January 1985 to December 2009. The top of Fig. 2 plots the annualized Sharpe ratios of these strategies for each futures contract. As the figure shows, *every* single futures contract exhibits positive predictability from past one-year returns. All 58 futures contracts exhibit positive time series momentum returns and 52 are statistically different from zero at the 5% significance level.

If we regress the TSMOM strategy for each security on the strategy of always being long (i.e., replacing "sign" with a 1 in Eq. (5)), then we get a positive alpha in 90% of the cases (of which 26% are statistically significant; none of the few negative ones are significant). Thus, a time series momentum strategy provides additional returns over and above a passive long position for most instruments.

Table 2 *t*-statistics of the alphas of time series momentum strategies with different look-back and holding periods.
Reported are the *t*-statistics of the alphas (intercepts) from time series regressions of the returns of time series momentum strategies over various look-back and holding periods on the following factor portfolios: MSCI World Index, Lehman Brothers/Barclays Bond Index, S&P GSCI Index, and HML, SMB, and UMD Fama and French factors from Ken French's Web site. Panel A reports results for all asset classes, Panel B for commodity futures, Panel C for equity index futures, Panel D for bond futures, and Panel E for currency forwards.

		Holding period (months)							
		1	**3**	**6**	**9**	**12**	**24**	**36**	**48**
Panel A: All assets									
Lookback period (months)	1	4.34	4.68	3.83	4.29	5.12	3.02	2.74	1.90
	3	5.35	4.42	3.54	4.73	4.50	2.60	1.97	1.52
	6	5.03	4.54	4.93	5.32	4.43	2.79	1.89	1.42
	9	6.06	6.13	5.78	5.07	4.10	2.57	1.45	1.19
	12	6.61	5.60	4.44	3.69	2.85	1.68	0.66	0.46
	24	3.95	3.19	2.44	1.95	1.50	0.20	–0.09	–0.33
	36	2.70	2.20	1.44	0.96	0.62	0.28	0.07	0.20
	48	1.84	1.55	1.16	1.00	0.86	0.38	0.46	0.74
Panel B: Commodity futures									
Lookback period (months)	1	2.44	2.89	2.81	2.16	3.26	1.81	1.56	1.94
	3	4.54	3.79	3.20	3.12	3.29	1.51	1.28	1.62
	6	3.86	3.53	3.34	3.43	2.74	1.59	1.25	1.48
	9	3.77	4.05	3.89	3.06	2.31	1.27	0.71	1.04
	12	4.66	4.08	2.64	1.85	1.46	0.58	0.14	0.57
	24	2.83	2.15	1.24	0.58	0.18	–0.60	–0.33	–0.14
	36	1.28	0.74	0.07	–0.25	–0.34	–0.03	0.34	0.65
	48	1.19	1.17	1.04	1.01	0.92	0.75	1.16	1.29

(*Continued*)

Table 2 (*Continued*)

		Holding period (months)							
		1	**3**	**6**	**9**	**12**	**24**	**36**	**48**
Panel C: Equity index futures									
Lookback period (months)	1	1.05	2.36	2.89	3.08	3.24	2.28	1.93	1.28
	3	1.48	2.23	2.21	2.81	2.78	2.00	1.57	1.14
	6	3.50	3.18	3.49	3.52	3.03	2.08	1.36	0.88
	9	4.21	3.94	3.79	3.30	2.64	1.96	1.21	0.75
	12	3.77	3.55	3.03	2.58	2.02	1.57	0.78	0.33
	24	2.04	2.22	1.96	1.70	1.49	0.87	0.43	0.13
	36	1.86	1.66	1.26	0.90	0.66	0.34	0.02	0.08
	48	0.81	0.84	0.58	0.44	0.36	0.12	0.01	0.23
Panel D: Bond futures									
Lookback period (months)	1	3.31	2.66	1.84	2.65	2.88	1.76	1.60	1.40
	3	2.45	1.52	1.10	1.99	1.80	1.27	1.05	1.00
	6	2.16	2.04	2.18	2.53	2.24	1.71	1.36	1.37
	9	2.93	2.61	2.68	2.55	2.43	1.83	1.17	1.40
	12	3.53	2.82	2.57	2.42	2.18	1.47	1.12	0.96
	24	1.87	1.55	1.62	1.66	1.58	1.01	0.90	0.64
	36	1.97	1.83	1.70	1.62	1.73	1.13	0.75	0.91
	48	2.21	1.80	1.53	1.43	1.26	0.72	0.73	1.22
Panel E: Currency forwards									
Lookback period (months)	1	3.16	3.20	1.46	2.43	2.77	1.22	0.83	–0.42
	3	3.90	2.75	1.54	3.05	2.55	1.02	0.10	–0.84
	6	2.59	1.86	2.32	2.82	2.08	0.62	–0.16	–1.14
	9	3.40	3.16	2.65	2.35	1.72	0.20	–0.38	–1.17
	12	3.41	2.40	1.65	1.25	0.71	–0.29	–1.01	–1.67
	24	1.78	0.99	0.53	0.27	–0.05	–1.15	–1.88	–2.27
	36	0.73	0.42	–0.04	–0.42	–0.96	–1.67	–2.04	–2.42
	48	–0.55	–1.05	–1.41	–1.62	–1.79	–2.02	–2.34	–2.32

Figure 2. Sharpe ratio of 12-month time series momentum by instrument. Reported are the annualized gross Sharpe ratio of the 12-month time series momentum or trend strategy for each futures contract/instrument. For each instrument in every month, the trend strategy goes long (short) the contract if the excess return over the past 12 months of being long the instrument is positive (negative), and scales the size of the bet to be inversely proportional to the ex ante volatility of the instrument to maintain constant volatility over the entire sample period from January 1985 to December 2009. The second figure plots a normalized value of the illiquidity of each futures contract measured by ranking contracts within each asset class by their daily trading volume (from highest to lowest) and reporting the standard normalized rank for each contract within each asset class. Positive (negative) values imply the contract is more (less) illiquid than the median contract for that asset class.

The overall return of the strategy that diversifies across all the S_t securities that are available at time t is

$$r_{t,t+1}^{TSMOM} = \frac{1}{S_t}\sum_{s=1}^{S_t} sign\left(r_{t-12,t}^{s}\right)\frac{40\%}{\sigma_t^s} r_{t,t+1}^{s}.$$

We analyze the risk and return of this factor in detail next. We also consider TSMOM strategies by asset class constructed analogously.

4.2. Alpha and Loadings on Risk Factors

Table 3 examines the risk-adjusted performance of a diversified TSMOM strategy and its factor exposures. Panel A of Table 3 regresses the excess return of the TSMOM strategy on the returns of the MSCI World stock market index and the standard Fama-French factors *SMB, HML,* and *UMD,* representing the size, value, and cross-sectional momentum premium among individual stocks. The first row reports monthly time series regression results and the second row uses quarterly non-overlapping returns (to account for any non-synchronous trading effects across markets). In both cases, TSMOM delivers a large and significant alpha or intercept with respect to these factors of about 1.58% per month or 4.75% per quarter. The TSMOM strategy does not exhibit significant betas on the market, *SMB,* or *HML* but loads significantly positively on *UMD,* the cross-sectional momentum factor. We explore the connection between cross-sectional and time series momentum more fully in the next section, but given the large and significant alpha, it appears that time series momentum is not fully explained by cross-sectional momentum in individual stocks.

Panel B of Table 3 repeats the regressions using the Asness, Moskowitz, and Pedersen (2010) value and momentum ''everywhere'' factors (i.e., factors diversified across asset classes) in place of the Fama and French factors. Asness, Moskowitz, and Pedersen (2010) form long-short portfolios of value and momentum across individual equities from four international markets, stock index futures, bond futures, currencies, and commodities. Similar to the Fama and French factors, these are cross-sectional factors. Once again, we find no significant loading on the market index or the value everywhere factor, but significant loading on the cross-sectional momentum everywhere factor. However, the returns to TSMOM are not fully captured by the cross-sectional everywhere factor—the alpha is still an impressive 1.09% per month with a t-stat of 5.40 or 2.93% per quarter with a t-stat of 4.12.

4.3. Performance over Time and in Extreme Markets

Fig. 3 plots the cumulative excess return to the diversified time series momentum strategy over time (on a log scale). For comparison, we also plot the cumulative excess returns of a diversified passive long position in all instruments, with an equal amount of risk in each instrument. (Since each instrument is scaled by the same constant volatility, both portfolios have the same ex ante volatility except for differences in correlations among time series

Table 3 Performance of the diversified time series momentum strategy.
Panel A reports results from time series regressions of monthly and non-overlapping quarterly returns on the diversified time series momentum strategy that takes an equal-weighted average of the time series momentum strategies across all futures contracts in all asset classes, on the returns of the MSCI World Index and the Fama and French factors SMB, HML, and UMD, representing the size, value, and cross-sectional momentum premiums in US stocks. Panel B reports results using the Asness, Moskowitz, and Pedersen (2010) value and momentum "everywhere" factors instead of the Fama and French factors, which capture the premiums to value and cross-sectional momentum globally across asset classes. Panel C reports results from regressions of the time series momentum returns on the market (MSCI World Index), volatility (VIX), funding liquidity (TED spread), and sentiment variables from Baker and Wurgler (2006, 2007), as well as their extremes.

***Panel A*: Fama and French factors**

		MSCI World	SMB	HML	UMD	Intercept	R^2
Monthly	Coefficient	0.09	–0.05	–0.01	0.28	1.58%	14%
	(*t*-Stat)	(1.89)	(–0.84)	(–0.21)	(6.78)	(7.99)	
Quarterly	Coefficient	0.07	–0.18	0.01	0.32	4.75%	23%
	(*t*-Stat)	(1.00)	(–1.44)	(0.11)	(4.44)	(7.73)	

***Panel B*: Asness, Moskowitz, and Pedersen (2010) factors**

		MSCI World	VAL Everywhere	MOM Everywhere	Intercept	R^2
Monthly	Coefficient	0.11	0.14	0.66	1.09%	30%
	(*t*-Stat)	(2.67)	(2.02)	(9.74)	(5.40)	
Quarterly	Coefficient	0.12	0.26	0.71	2.93%	34%
	(*t*-Stat)	(1.81)	(2.45)	(6.47)	(4.12)	

(*Continued*)

Table 3 (*Continued*)

***Panel C*: Market, volatility, liquidity, and sentiment extremes**

		MSCI World	MSCI World squared	TED spread	TED spread top 20%	VIX	VIX top 20%
Quarterly	Coefficient	–0.01	1.99				
	(*t*-Stat)	(–0.17)	(3.88)				
Quarterly	Coefficient			–0.001	–0.008		
	(*t*-Stat)			(– .06)	(–0.29)		
Quarterly	Coefficient					0.001	–0.003
	(*t*-Stat)					(0.92)	(–0.10)

		Sentiment	Sentiment top 20%	Sentiment bottom 20%	Change in sentiment	Change in sentiment top 20%	Change in sentiment bottom 20%
Quarterly	Coefficient	0.03	–0.01	–0.01			
	(*t*-Stat)	(0.73)	(–0.27)	(–0.12)			
Quarterly	Coefficient				–0.01	0.02	0.01
	(*t*-Stat)				(–1.08)	(1.25)	(0.66)

Figure 3 Cumulative excess return of time series momentum and diversified passive long strategy, January 1985 to December 2009. Plotted are the cumulative excess returns of the diversified TSMOM portfolio and a diversified portfolio of the possible long position in every futures contract we study. The TSMOM portfolio is defined in Eq. (5) and across all futures contracts summed.

momentum strategies and passive long strategies.) As Fig. 3 shows, the performance over time of the diversified time series momentum strategy provides a relatively steady stream of positive returns that outperforms a diversified portfolio of passive long positions in all futures contracts (at the same ex ante volatility).

We can also compute the return of the time series momentum factor from 1966 to 1985, despite the limited number of instruments with available data. Over this earlier sample, time series momentum has a statistically significant return and an annualized Sharpe ratio of 1.1, providing strong out-of-sample evidence of time series momentum.[9]

Fig. 3 highlights that time series momentum profits are large in October, November, and December of 2008, which was at the height of the Global Financial Crisis when commodity and equity prices dropped sharply, bond prices rose, and currency rates moved dramatically. Leading into this period, time series momentum suffers losses in the third quarter of 2008, where the associated price moves caused the TSMOM strategy to be short in many contracts, setting up large profits that were earned in the fourth quarter of 2008 as markets in all these asset classes fell further. Fig. 3 also shows that TSMOM suffers sharp losses when the crisis ends in March, April, and May of 2009. The ending of a crisis constitutes a sharp trend reversal that generates losses on a trend following strategy such as TSMOM.

More generally, Fig. 4 plots the TSMOM returns against the S&P 500 returns. The returns to TSMOM are largest during the biggest up and down market movements. To test the statistical significance of this finding, the first row of Panel C of Table 3 reports coefficients from a regression of TSMOM returns on the market index return and squared market index return. While the beta on the market itself is insignificant, the coefficient on the market return squared is significantly positive, indicating that TSMOM delivers its highest profits during the most extreme market episodes. TSMOM, therefore, has payoffs similar to an option straddle on the market. Fung and Hsieh (2001) discuss why trend following has straddlelike payoffs and apply this insight to describe the performance of hedge funds. Our TSMOM strategy generates this payoff structure because it tends to go long when the market has a major upswing and short when the market crashes.

Figure 4 The time series momentum smile. The non-overlapping quarterly returns on the diversified (equally weighted across all contracts) 12-month time series momentum or trend strategy are plotted against the contemporaneous returns on the S&P 500.

These results suggest that the positive average TSMOM returns are not likely to be compensation for crash risk. Historically, TSMOM does well during "crashes" because crises often happen when the economy goes from normal to bad (making TSMOM go short risky assets), and then from bad to worse (leading to TSMOM profits), with the recent financial crisis of 2008 being a prime example.

4.4. Liquidity and Sentiment

We test whether TSMOM returns might be driven or exaggerated by illiquidity. We first test whether TSMOM performs better for more illiquid assets in the cross-section, and then we test whether the performance of the diversified TSMOM factor depends on liquidity indicators in the time series. For the former, we measure the illiquidity of each futures contract using the daily dollar trading volume obtained from Reuters and broker feeds. We do not have historical time series of daily volume on these contracts, but use a snapshot of their daily volume in June 2010 to examine cross-sectional differences in liquidity across the assets. Since assets are vastly different across many dimensions, we first rank each contract within an asset class by their daily trading volume (from highest to lowest) and compute the standard normalized rank of each contract by demeaning each rank and dividing by its standard deviation, i.e., *(rank-mean(rank))/std(rank)*. Positive (negative) values imply a contract is more (less) illiquid than the median contract for that asset class. As shown in the bottom of Figure 2, we find little relation between the magnitude of the Sharpe ratio of TSMOM for a particular contract and its illiquidity, as proxied by daily dollar trading volume. The correlation between illiquidity and Sharpe ratio of a time series momentum strategy by contract is –0.16 averaged across all contracts, suggesting that, if anything, more *liquid* contracts exhibit greater time series momentum profits.

We next consider how TSMOM returns co-vary in aggregate with the time series of liquidity. The second row of Panel C of Table 3 reports results using the Treasury Eurodollar (TED) spread, a proxy for funding liquidity as suggested by Brunnermeier and Pedersen (2009), Asness, Moskowitz, and Pedersen (2010), and Garleanu and Pedersen (2011), and the top 20% most extreme realizations of the TED spread to capture the most illiquid

funding environments. As the table shows, there is no significant relation between the TED spread and TSMOM returns, suggesting little relationship with funding liquidity. The third row of Panel C of Table 3 repeats the analysis using the VIX index to capture the level of market volatility and the most extreme market volatility environments, which also seem to correspond with illiquid episodes. There is no significant relationship between TSMOM profitability and market volatility either.

At the bottom of Panel C of Table 3, we also examine the relationship between TSMOM returns and the sentiment index measures used by Baker and Wurgler (2006, 2007). We examine both the level of sentiment and its monthly changes (first differences) and examine the top and bottom extremes (20%) of these variables. As the regressions indicate, we find no significant relationship between TSMOM profitability and sentiment measures, even at the extremes.

4.5. Correlation Structure

Table 4 examines the correlation structure of the time series momentum strategies and compares them to the correlation structure of passive long positions in the contracts. The first row of Panel A of Table 4 reports the average pair-wise correlation of time series momentum returns among contracts within the same asset class. The correlations are positive within each asset class, ranging from 0.37 to 0.38 for equities and fixed income futures to 0.10 and 0.07 for commodities and currencies. Part of this correlation structure reflects the comovement of the returns to simply being passive long (or short) in each instrument at the same time. The second row of Panel A of Table 4 reports the average pair-wise correlation of passive long positions within each asset class and, except for currencies, passive long strategies exhibit higher correlations than time series momentum strategies within an asset class.

Panel B of Table 4 shows the average correlation of time series momentum strategies *across* asset classes. Here, we first compute the return of a diversified portfolio of time series momentum strategies within each asset class and then estimate the correlation of returns for TSMOM portfolios across asset classes. All of the correlations are positive, ranging from 0.05 to 0.21. For comparison, the table also shows the correlations across asset classes of diversified passive long positions. For every asset class comparison, the correlation of time series momentum strategies across asset classes is larger than the corresponding correlation of passive long strategies, many of which are negative.

Summing up the results from both panels, time series momentum strategies are positively correlated within an asset class, but less so than passive long strategies. However, across asset classes, time series momentum strategies exhibit positive correlation with each other, while passive long strategies exhibit zero or negative correlation across asset classes. This last result suggests that there is a common component affecting time series momentum strategies across asset classes simultaneously that is not present in the underlying assets themselves, similar to the findings of Asness, Moskowitz, and Pedersen (2010) who find common structure among cross-sectional momentum strategies across different asset classes.

Table 4 Correlations of time series momentum strategy returns within and across asset classes.

Panel A reports within each asset class the average pair-wise correlation of each instrument's 12-month time series momentum strategy returns, as well as a passive long position in each instrument. Panel B reports the correlation of time series momentum strategies and passive long positions *across* asset classes, where an equal-weighted average of the instruments within each asset class is first formed and then the correlations between the equal-weighted strategies across asset classes are calculated. Correlations are calculated from monthly returns over the period January 1985–December 2009.

***Panel A*: Average pair-wise correlation within asset class**				
	Commodities	**Equities**	**Fixed income**	**Currencies**
TSMOM strategies	0.07	0.37	0.38	0.10
Passive long positions	0.19	0.60	0.63	–0.04
***Panel B*: Average correlation across asset classes**				
	Correlations of TSMOM strategies			
Commodities	1			
Equities	0.20	1		
Fixed income	0.07	0.21	1	
Currencies	0.13	0.20	0.05	1
	Correlations of passive long positions			
Commodities	1			
Equities	0.17	1		
Fixed income	–0.12	–0.03	1	
Currencies	–0.12	–0.20	0.02	1

5. Time Series vs. Cross-Sectional Momentum

Our previous results show a significant relationship between time series momentum and cross-sectional momentum. In this section, we explore that relationship further and determine how much overlap and difference exist between our time series momentum strategies and the cross-sectional momentum strategies commonly used in the literature.

5.1. Time Series Momentum Regressed on Cross-Sectional Momentum

Panel A of Table 5 provides further evidence on the relationship between time series momentum (TSMOM) and cross-sectional momentum (XSMOM) by regressing the returns to our time series momentum strategies— diversified across all instruments and within each asset class—on the returns of cross-sectional momentum strategies applied to the same assets. Specifically, following Asness, Moskowitz, and Pedersen (2010), we apply a cross-sectional momentum strategy based on the relative ranking of each asset's past 12-month returns and form portfolios that go long or short the assets in proportion to their ranks relative to the median rank.[10]

The first row of Panel A of Table 5 reports results from the TSMOM strategy diversified across all assets regressed on the XSMOM strategy that is diversified across those same assets. As before, time series momentum and cross-sectional momentum are significantly related, with the beta of TSMOM on XSMOM equal to 0.66 with a t-statistic of 15.17 and R-square of 44%. However, as the intercept indicates, TSMOM is not fully captured by XSMOM, exhibiting a positive and significant alpha of 76 basis points per month with a *t*-statistic of 5.90. So, TSMOM and XSMOM are related, but are not the same.

The second row of Panel A of Table 5 repeats the regression using XSMOM strategies for each asset class, including individual stocks. TSMOM is related to XSMOM across all of the different asset classes, including individual equity momentum, which is not even included in the TSMOM strategy, and this is *after* controlling for exposure to XSMOM from the other four asset classes. TSMOM still exhibits a significant alpha, however, and is therefore not fully captured by cross-sectional momentum strategies in these asset classes.

Repeating these regressions using the TSMOM returns for each asset class separately, we find a consistent pattern. TSMOM is related to XSMOM within each asset class, with *R*-squares ranging from 56% in currencies (FX) to 14% in fixed income, but TSMOM is not captured by XSMOM. The alphas of TSMOM remain significantly positive for every asset class. We also see some interesting cross-asset relationships among TSMOM and XSMOM. For instance, not only is TSMOM for commodities correlated with XSMOM for commodities, but also with XSMOM for currencies. Likewise, TSMOM among equity index futures is not only correlated with XSMOM among those equity indexes but also with XSMOM among individual stocks. And, TSMOM for fixed income is correlated with XSMOM for fixed income and XSMOM for equity indexes. These results indicate significant correlation structure in time series and cross-sectional momentum across different asset classes, consistent with our earlier results and those of Asness, Moskowitz, and Pedersen (2010).

5.2. A Simple, Formal Decomposition

We can more formally write down the relationship between time series (TSMOM) and cross-sectional (XSMOM) momentum. Following Lo and Mackinlay (1990) and Lewellen (2002), we can describe a simple cross-sectional and time series momentum strategy on the same assets as follows. For cross-sectional momentum, we let the portfolio weight of instrument *i* be $w_t^{XS,i} = (1/N)\left(r_{t-12,t}^i - r_{t-12,t}^{EW}\right)$, that is, the past 12-month excess return over the equal-weighted average return, $r_{t-12,t}^{EW} = (1/N)\sum_{i=1}^{N} r_{t-12,t}^i$. The return to the portfolio is therefore

$$r_{t,t+1}^{XS} = \sum_{i=1}^{N} W_t^{XS,i} r_{t,t+1}^i .$$

Next, assuming that the monthly expected return is[11] $\mu^i = E\left(r^i_{t,t+1}\right) = E\left(r^i_{t-12,t}\right)/12$ and letting $\mu = \left[\mu^1,....,\mu^N\right]'$, $R_{t,s} = \left[r^1_{t,s},....,r^N_{t,s}\right]'$, and $\Omega = E\left[\left(R_{t-12,t} - 12\mu\right)\left(R_{t,t+1} - \mu\right)'\right]$, the expected return to cross-sectional momentum (XSMOM) can be decomposed as

$$\begin{aligned} E\left[r^{XS}_{t,t+1}\right] &= \frac{tr\left(\Omega\right)}{N} - \frac{1'\,\Omega 1}{N^2} + 12\sigma^2_\mu \\ &= \frac{N-1}{N^2} tr\left(\Omega\right) - \frac{1}{N^2}\left[1'\,\Omega 1 - tr\left(\Omega\right)\right] + 12\sigma^2_\mu, \end{aligned} \tag{6}$$

where *tr* is the trace of a matrix, 1 is an (N × 1) vector of ones, and σ^2_μ is the cross-sectional variance of the mean monthly returns μ^i

Eq. (6) shows that cross-sectional momentum profits can be decomposed into an auto-covariance component between lagged 1-year returns and future 1-month returns (the diagonal elements of Ω captured by the first term), a cross-covariance component capturing the temporal leads and lags across stocks (the off-diagonal elements of Ω captured by the second term), and the cross-sectional variation in unconditional mean returns (the third term). As emphasized by Lewellen (2002), cross-sectional momentum profits need not be generated by positive autocorrelation in returns (i.e., time series predictability). If cross-serial covariances across stocks are negative, implying that high past returns of an asset predict lower future returns of other assets, this, too, can lead to momentum profits. Likewise, large cross-sectional variation in mean returns can also lead to momentum profits since, on average, assets with the highest mean returns will have the highest realized returns.

The returns to time series momentum can be decomposed similarly if we let the portfolio weights be $W^{TS,i}_t = \left(1/N\right) r^i_{t-12,t}$. Then the expected return is

$$E\left(r^{TS}_{t,t+1}\right) = E\left(\sum_{i=1}^{N} w^{TS,i}_t r^i_{t,t+1}\right) = \frac{tr\left(\Omega\right)}{N} + 12\frac{\mu'\mu}{N} \tag{7}$$

As these equations highlight, time series momentum profits are primarily driven by time series predictability (i.e., positive auto-covariance in returns) if the average squared mean returns of the assets is small. Comparing Eqs. (6) and (7), we see that time series momentum profits can be decomposed into the autocovariance term that also underlies cross-sectional momentum (plus the average squared mean excess return). The equations thus provide a link between time series and cross-sectional momentum profitability, which we can measure in the data to determine how related these two phenomena are.

Panel B of Table 5 computes each of the components of the diversified 12-month cross-sectional and time series momentum strategies across all assets and within each asset class according to the equations above. We report the three components of the cross-sectional momentum strategy: "Auto" refers to the auto-covariance or time series momentum component, "Cross" refers to the cross-serial covariance or lead-lag component, and "Mean" refers to the contribution from unconditional mean returns, as well as their sum ("Total").

Table 5 Time series momentum vs. cross-sectional momentum.
Panel A reports results from regressions of the 12-month time series momentum strategies by asset class (TSMOM) on 12-month cross-sectional momentum strategies (XSMOM) of Asness, Moskowitz, and Pedersen (2010). Panel B reports results from the decomposition of cross-sectional momentum and time series momentum strategies according to Section 4.2, where Auto is the component of profits coming from the auto-covariance of returns, Cross is the component coming from cross-serial correlations or lead-lag effects across the asset returns, Mean is the component coming from cross-sectional variation in unconditional mean returns, and Mean squared is the component coming from squared mean returns. Panel C reports results from regressions of several XSMOM strategies in different asset classes, the Fama-French momentum, value, and size factors, and two hedge fund indexes obtained from Dow Jones/Credit Suisse on our benchmark TSMOM factor.

***Panel A*: Regression of TSMOM on XSMOM**

		Independent variables							
		XSMOM ALL	XSMOM COM	XSMOM EQ	XSMOM FI	XSMOM FX	XSMOM US stocks	Intercept	R^2
Dependent variable	TSMOM ALL	0.66						0.76%	44%
		(15.17)						(5.90)	
	TSMOM ALL		0.31	0.20	0.17	0.37	0.12	0.73%	46%
			(7.09)	(4.25)	(3.84)	(8.11)	(2.66)	(5.74)	
	TSMOM COM		0.65					0.57%	42%
			(14.61)					(4.43)	
	TSMOM COM		0.62	0.05	0.02	0.14	0.05	0.51%	45%
			(13.84)	(1.01)	(0.50)	(3.08)	(1.06)	(3.96)	
	TSMOM EQ			0.39				0.47%	15%
				(7.32)				(3.00)	
	TSMOM EQ		0.07	0.28	0.04	0.06	0.24	0.43%	22%
			(1.29)	(5.07)	(0.67)	(1.11)	(4.26)	(2.79)	
	TSMOM FI				0.37			0.59%	14%
					(6.83)			(3.77)	
	TSMOM FI		–0.03	0.18	0.34	0.01	0.03	0.50%	17%
			(–0.62)	(3.05)	(6.19)	(0.20)	(0.48)	(3.15)	
	TSMOM FX					0.75		0.42%	56%
						(19.52)		(3.75)	
	TSMOM FX		0.04	0.00	–0.01	0.75	–0.01	0.40%	56%
			(1.07)	(–0.04)	(–0.17)	(18.89)	(–0.24)	(3.49)	

(*Continued*)

Table 5 (*Continued*)

***Panel B*: Decomposition of TSMOM and XSMOM**

	XSMOM decomposition				TSMOM decomposition		
	Auto	**Cross**	**Mean**	**Total**	**Auto**	**Mean squared**	**Total**
ALL	0.54%	–0.03%	0.12%	0.61%	0.54%	0.29%	0.83%
COM	0.41%	–0.13%	0.11%	0.39%	0.43%	0.17%	0.59%
EQ	0.74%	–0.62%	0.02%	0.14%	0.83%	0.17%	1.00%
FI	0.32%	–0.10%	0.05%	0.27%	0.35%	0.70%	1.05%
FX	0.80%	–0.55%	0.02%	0.18%	0.80%	0.17%	0.96%

***Panel C*: What factors does TSMOM explain?**

		Independent variable		
		TSMOM ALL	**Intercept**	R^2
Dependent variable	XSMOM ALL	0.66	–0.16%	44%
		(15.17)	(–1.17)	
	XSMOM COM	0.65	–0.09%	42%
		(14.61)	(–0.66)	
	XSMOM EQ	0.39	0.29%	15%
		(7.32)	(1.86)	
	XSMOM FI	0.37	–0.14%	14%
		(6.83)	(–0.87)	
	XSMOM FX	0.75	–0.19%	56%
		(19.52)	(–1.71)	
	UMD	0.49	–0.28%	13%
		(6.56)	(–0.93)	
	HML	–0.07	0.43%	1%
		(–1.46)	(2.08)	
	SMB	–0.01	0.10%	0%
		(–0.26)	(0.49)	
	DJCS MF	0.55	–0.30%	33%
		(9.60)	(–1.37)	
	DJCS MACRO	0.32	0.52%	14%
		(5.64)	(2.38)	

We also report the two components to TSMOM: the Auto and mean-squared return components, as well as their sum.

As Panel B of Table 5 shows, time series and cross-sectional momentum are related but different. The auto-covariance component contributes just about all of the cross-sectional momentum strategy profits across all assets. The cross-serial or lead-lag component contributes *negatively* to XSMOM and the cross-sectional variation in means has a small positive contribution to cross-sectional momentum profits. The contribution of these components to cross-sectional momentum strategies is also fairly stable across asset classes, with the dominant component consistently being the auto-covariance or time series piece.

The decomposition of time series momentum shows that the main component is the auto-covariance of returns. Squared mean excess returns are a much smaller component of TSMOM profits, except for fixed income. Since the cross-sectional correlation of lead-lag effects among assets contributes negatively to XSMOM, it is not surprising that TSMOM, which does not depend on the cross-serial correlations across assets, produces higher profits than XSMOM.

We also regress the returns of time series momentum as defined in Eq. (7) (which is linear as opposed to using the sign of the past return as before) on the returns to cross-sectional momentum as defined in Eq. (6). We find that time series momentum has a significant alpha to cross-sectional momentum, consistent with our earlier results in Panel A of Table 5 that use a slightly different specification. We investigate next whether the reverse is true. Does TSMOM explain XSMOM?

5.3. Does TSMOM Explain Cross-sectional Momentum and Other Factors?

Panel C of Table 5 uses TSMOM as a right-hand-side variable, testing its ability to explain other factors. We first examine XSMOM to see if TSMOM can capture the returns to cross-sectional momentum across all asset classes as well as within each asset class. As the first five rows of Panel C of Table 5 show, TSMOM is able to fully explain cross-sectional momentum across all assets as well as within each asset class for commodities, equity indexes, bonds, and currencies. The intercepts or alphas of XSMOM are statistically no different from zero, suggesting TSMOM captures the return premiums of XSMOM in these markets. The only positive alpha is for XSMOM in equity indexes, which has a marginal 1.86 t-statistic. We also regress the Fama-French cross-sectional momentum factor for individual US equities, UMD, on our TSMOM portfolio. The UMD factor is created from individual US equities and hence has no overlap with any of the assets used to comprise our TSMOM factor. Nevertheless, TSMOM is able to capture the return premium to UMD, which has a positive 0.49 loading on TSMOM and an insignificant –0.28 alpha (t-stat = –0.93). We also examine the other Fama-French factors HML, the value factor, and SMB, the size factor. HML loads negatively on TSMOM, so TSMOM naturally cannot explain the value effect, and SMB has a loading close to zero.

Finally, we also examine the returns to two popular hedge fund indexes that trade globally across many assets: the "Managed Futures" hedge fund index and "Global Macro" hedge fund index obtained from Dow Jones/Credit Suisse from 1994 to 2009. As the last

two rows of Panel C of Table 5 show, both hedge fund indexes load significantly on our TSMOM factor and in the case of the Managed Futures index, the TSMOM factor captures its average return entirely. Hence, TSMOM is a simple implementable factor that captures the performance metric of Fung and Hsieh (2001), which they show explains hedge fund returns as well.[12]

The strong performance of TSMOM and its ability to explain some of the prominent factors in asset pricing, namely, cross-sectional momentum as well as some hedge fund strategy returns, suggests that TSMOM is a significant feature of asset price behavior. Future research may well consider what other asset pricing phenomena might be related to time series momentum.

6. Who Trades on Trends: Speculators or Hedgers?

To consider who trades on time series momentum, Fig. 5 shows the Net speculator position broken down by the sign of the past 12-month return for each instrument with available CFTC data. Specifically, for each futures contract, Fig. 5 plots the average Net speculator position in, respectively, the subsample where the past 12-month return on the contract is positive ("Positive TSMOM") and negative ("Negative TSMOM") de-meaned using the average Net speculator position for each instrument. The figure illustrates that speculators are, on average, positioned to benefit from trends, whereas hedgers, by definition, have the opposite positions. Speculators have longer-than-average positions following positive past 12-month returns, and smaller-than-average positions following negative returns, on average. Said differently, speculators have larger positions in an instrument following positive returns than following negative returns. This trend-following pattern of speculative positions is found for every contract except the S&P 500 futures, where Net speculator positions have opposite signs, though are close to zero. Since we also know that time series momentum is associated with positive abnormal returns, these results indicate that speculators profit, on average, from these position changes at the expense of hedgers.

6.1. The Evolution of TSMOM

We next consider the dynamics of these trading positions over time. Our previous results suggest that time series momentum lasts about a year and then starts to reverse. We investigate these return patterns in more depth and attempt to link them to the evolution of trading positions.

Examining the evolution of TSMOM and trading patterns may help distinguish theories for momentum. For example, if initial under-reaction to news is the cause of TSMOM, then this part of the trend should *not* reverse, whereas the part of a trend that is driven by over-reaction *should* reverse as prices eventually gravitate back toward fundamentals.

To consider the evolution of TSMOM, we perform an event study as follows. For each month and instrument, we first identify whether the previous 12-month excess returns are positive or negative. For all the time-instrument pairs with positive 12-month past returns, we compute the average return from 12 months prior to the "event date" (portfolio formation date) to 36 months after. We do the same for the time-instrument pairs with negative past 12-month returns. We then standardize the returns to have a zero mean across time

Figure 5 Net speculator positions. For each futures contract, the figure plots the average de-meaned Net speculator position in, respectively, the subsample where the past 12-month returns on the contract are positive (''Positive TSMOM'') and negative (''Negative TSMOM''). The figure illustrates that speculators are on average positioned to benefit from trends, whereas hedgers, by definition, have the opposite positions.

and across the two groups (for ease of comparison), and compute the cumulative returns of the subsequent months following positive past-year returns (''Positive TSMOM'') and negative past-year returns (''Negative TSMOM''), respectively, where we normalize the cumulative returns to be one at the event date.

Panel A of Fig. 6 shows the cumulative returns conditional on positive and negative time series momentum. The returns to the left of the event date are, of course, positive and negative by construction. To the right of the event date, we see that the positive preformation returns continue upward after the portfolio formation for about a year, consistent with a time series momentum effect, and then partially reverse thereafter. This is consistent with both initial under-reaction and delayed over-reaction as predicted by sentiment theories such as Daniel, Hirshleifer, and Subrahmanyam (1998) and Barberis, Shleifer, and Vishny (1998). While the reversal after a year suggests over-reaction, the fact that only part of the post-formation upward trend is reversed suggests that under-reaction appears to be part of the story as well. Similarly, the negative preformation trend is continued for a year until it partially reverses as well.

Panel B of Fig. 6 shows the evolution of Net speculator positions that coincide with the positive and negative time series momentum returns. Specifically, for each instrument and month, we compute the average Net speculator position for each month from 12 months prior to the event (portfolio formation date) to 36 months after portfolio formation for both positive and negative trends. We see that for positive TSMOM, speculators increase their positions steadily from months –12 to 0, building up to the formation date.

Figure 6 Event study of time series momentum. For each month and instrument, we identify whether the previous 12-month returns are positive or negative and compute the average return from 12 months prior to the "event date" (portfolio formation date) to 36 months after following positive past-year returns ("Positive TSMOM") and negative past-year returns ("Negative TSMOM"). We standardize the returns to have a zero mean across time and across the two groups (for ease of comparison), and compute the cumulative returns of the subsequent months, where we normalize the cumulative returns to be 1 at the event date. Panel B plots the Net speculator position as defined by the CFTC conditional on positive and negative past returns. (A) *Panel A*: Cumulative returns in event time; (B) *Panel B*: Net speculator positions in event time.

Likewise, speculators decrease their positions steadily from the negative TSMOM event date. These patterns are not by construction since we split the sample based on returns and not based on Net speculator positions. After the event date, speculators' positions begin to mean-revert towards their (positive) average levels, and plateau at about a year (and maybe slightly longer for negative TSMOM), which is when the trend in returns starts to reverse.

The patterns in Fig. 6 indicate that while speculator positions are consistent with trading on TSMOM, they do not appear to keep piling into the trade with a lag. In fact, speculators appear to reduce their trend chasing up to the point where positive returns from following TSMOM disappear. Conversely, hedgers, who are on the other side of these trades, appear to be increasing their positions steadily in the direction of the trend. This suggests that if over-reaction is caused by such trading, it would have to come from hedgers, not speculators. While the direction of causality between returns and trading positions is indeterminate, these results also suggest that trading positions

of speculators and hedgers are closely linked to the profitability of time series momentum, where speculators appear to be profiting from trends and reversals at the expense of hedgers.

6.2. Joint Dynamics of Returns and Trading Positions

For a more formal analysis of trading patterns and returns, we study the joint dynamics of time series momentum returns and the change in Net speculator positions using a vector autoregressive (VAR) model. We estimate a monthly bivariate VAR with 24 months of lags of returns and changes in Net speculator positions and plot the impulse response of returns and Net speculator positions from a return shock. We need to include more than 12 months of lags to capture delayed reversal, but our results are robust to choosing other lag lengths between 12 and 24.

We perform a Cholesky decomposition of the variance-covariance matrix of residuals with the return first, and consider a one-standard-deviation shock to the returns of the contract. (The return response to an initial return shock is qualitatively the same regardless of the ordering of the Cholesky decomposition because of a limited feedback with positions.) The response to this impulse is plotted in Fig. 7, both in terms of the effect on the cumulative return to the contract and the cumulative changes in Net speculator positions. As the figure shows, returns continue to rise for about a year and then partially reverse thereafter following the return shock. Net speculator positions increase contemporaneously with the return shock and then mean-revert to zero at about a year. These results are consistent with our previous findings and confirm that speculative positions match the return patterns of time series momentum. Speculators seem to profit from TSMOM for about a year and then revert to their average positions at the same time the TSMOM effect ends; all at the expense of hedgers.

The patterns indicate that speculators profit from time series momentum, while hedgers pay for it. One explanation might be that speculators earn a premium through time series momentum for providing liquidity to hedgers. We explore this possibility further by examining the predictability of returns using trading positions as well as different components of futures returns. Specifically, we investigate whether changes in the underlying spot price or the shape of the futures curve (e.g., "roll yield") are driving the time series predictability and how each of these lines up with trading positions.

6.3. Predictability of Positions, Price Changes, and Roll Yield

We decompose the past return of each futures contract into the change in the price of the underlying spot asset and the return that is related to the shape of the futures curve, called the "roll return" or "roll yield."

We define the underlying spot price changes in excess of the risk-free rate as

$$Price\ change_{t-12,t} = \frac{Price - Price_{t-12}}{Price_{t-12}} - r^{f}_{t-12,t}$$

Figure 7 Impulse response from a shock to returns. Plotted are the cumulative returns and speculators' net positions in response to a one standard deviation shock to total returns on the futures contract (Panel A), returns on the spot asset (Panel B) and returns to rolling the contract (Panel C). The impulse response is based on an estimated vector autoregressive model using monthly returns with 24 lags of returns and Net speculator positions that assumes coefficients are the same across all contracts, with a Cholesky decomposition of the shock. (A) *Panel A*: Futures returns; (B) *Panel B*: Spot returns; (C) *Panel C*: Roll returns.

where prices are measured as the nearest-to-expiration futures price and $r^f_{t-12,t}$ is the risk-free interest rate over the 12-month period. We then define the roll return by the following decomposition:

$$Futures\ return_{t-12,t} = Price\ change_{t-12,t} + Roll\ return_{t-12,t}$$

In financial futures with little storage costs or convenience yield, the roll return is close to zero, but, in commodity markets, the roll return can be substantial. The futures return is calculated from the nearest-to-expiration or next-to-nearest expiration contract, whose maturity date may not be in the same month as the spot return calculation, which is based only on the nearest-to-expiration contract.

Our conjecture is that hedgers' price pressure affects mostly the roll returns, whereas information diffusion affects mostly spot price changes. To see why, recall first Keynes' (1923) basic idea that hedging pressure must affect required returns to give speculators an incentive to provide liquidity by taking the other side of the trade. Since hedging takes place in futures markets, hedging pressure would affect futures prices and thus lead to a roll yield as each futures contract expires at the spot price. When hedgers, such as commodity producers, are shorting the futures, this leads to positive roll return and what Keynes called "normal backwardation." On the other hand, information diffusion (which is the driver of several of the behavioral theories) would simply affect price changes.

Panel B of Fig. 7 plots the impulse response of spot price changes and Net speculator positions by repeating the VAR we ran above, replacing the total futures returns with the spot price changes only. The impulse response of spot returns and Net speculator positions matches those for total returns: trends exist for about a year and then reverse and Net speculative positions mirror that pattern. This is consistent with initial under-reaction and delayed over-reaction being due to information diffusion rather than hedging pressure.

Panel C of Fig. 7 plots the impulse response from replacing total returns with the roll return in the VAR. Here, the picture looks quite different. A shock to roll returns is associated with a continued upward trend to roll returns and a small effect on Net speculator positions. This is consistent with hedgers having stable positions in the same direction for extended periods of time and being willing to give up roll returns to enjoy hedging benefits. Speculators who take the other side profit from momentum as a premium for providing liquidity to hedgers.

Finally, Table 6 revisits the return predictability regressions we started with, focusing on 12-month return predictability, but examines the predictive power of the spot versus roll return, as well as their interaction with speculative trading positions. We regress the return of each futures contract on the past 12-month return of each contract, the spot price change of each contract, the roll return of each contract, and the change and level of Net speculator positions. The first five rows of Table 6 report the univariate regression results for each of these variables, which are all significant positive predictors of futures returns.

In multivariate regressions, however, the change in Net speculator positions drops slightly and becomes insignificant, indicating that controlling for past returns reduces some of the predictive power of speculative positions. This is consistent with the idea that roll return and speculator positions both capture hedging pressure, though measured differently and neither being a perfect measure for hedging pressure. Spot price changes and roll returns have almost the same predictive regression coefficient in the multivariate

Table 6 Time series predictors of returns: Spot prices, roll returns, and positions.
Reported are results from regressions of the monthly futures return on the previous 12 months' futures return ("Full TSMOM"), previous 12 months' change in spot price ("Spot price MOM"), past 12-month roll return ("Roll MOM"), and the 12-month change and average level in speculators' aggregate net (i.e., long minus short) positions as a percent of open interest ("Net speculator position"). Also reported are interactions between the change in Net speculator positions and the spot and roll returns over the previous 12 months.

	Full TSMOM	Spot price MOM	Roll MOM	Chg net speculator position	Net speculator position	Spot MOM × Chg net spec pos	Roll MOM × Chg net spec pos	Intercept	R^2
Coefficient	0.019							0.09%	0.6%
t-Stat	(3.57)							(1.31)	
Coefficient		0.014						0.12%	0.3%
t-Stat		(2.29)						(1.71)	
Coefficient			0.024					0.08%	0.3%
t-Stat			(3.22)					(1.08)	
Coefficient				0.007				0.12%	0.2%
t-Stat				(2.67)				(1.63)	
Coefficient					0.007			0.08%	0.2%
t-Stat					(2.33)			(1.12)	
Coefficient	0.017			0.004				0.09%	0.7%
t-Stat	(3.13)			(1.65)				(1.33)	
Coefficient	0.018				0.002			0.08%	0.6%
t-Stat	(3.31)				(0.76)			(1.14)	
Coefficient		0.017	0.030					0.08%	0.6%
t-Stat		(2.74)	(3.90)					(1.03)	
Coefficient		0.014	0.030	0.005				0.07%	0.8%
t-Stat		(2.12)	(3.94)	(1.89)				(0.99)	
Coefficient		0.014	0.030	0.005		0.023	0.015	0.06%	0.8%
t-Stat		(2.17)	(3.94)	(1.78)		(1.38)	(0.49)	(0.77)	

regression, hence, their joint predictive power (as measured by the R-square) is the same as the univariate predictability of their sum, which is the total futures return. Finally, the last row of Table 6 includes interaction terms between the spot and roll returns and Net speculator positions. While all the coefficients are positive, indicating that when changes in Net speculator positions move in the same direction as returns, there is stronger positive predictability of future returns, the results are not statistically significant.

The results in Table 6 indicate that time series momentum is not purely driven by one component of futures returns. Both the spot return change and roll yield provide predictive power for futures returns. In addition, as the VAR results show, there is an interesting dynamic between time series momentum and Net speculator and hedging positions. Speculators seem to ride the trend for about a year, eventually reducing their positions and taking the opposite side before it reverses. In the process, they earn positive excess returns at the expense of hedgers, who may be willing to compensate speculators for liquidity provision in order to maintain their hedge.

7. Conclusion

We find a significant time series momentum effect that is remarkably consistent across the nearly five dozen futures contracts and several major asset classes we study over the last 25 years. The time series momentum effect is distinct from cross-sectional momentum, though the two are related. Decomposing both time series and cross-sectional momentum profits, we find that the dominant force to both strategies is significant positive autocovariance between a security's excess return next month and its lagged 1-year return. This evidence is consistent with initial under-reaction stories, but may also be consistent with delayed over-reaction theories of sentiment as the time series momentum effect partially reverses after one year.

Time series momentum exhibits strong and consistent performance across many diverse asset classes, has small loadings on standard risk factors, and performs well in extreme periods, all of which present a challenge to the random walk hypothesis and to standard rational pricing models. The evidence also presents a challenge to current behavioral theories since the markets we study vary widely in terms of the type of investors, yet the pattern of returns remains remarkably consistent across these markets and is highly correlated across very different asset classes. Indeed, correlation among time series momentum returns is stronger than the correlation of passive long positions across the same asset classes, implying the existence of a common component to time series momentum that is not present in the underlying assets themselves.

Finally, the link between time series momentum returns and the positions of speculators and hedgers indicates that speculators profit from time series momentum at the expense of hedgers. This evidence is consistent with speculators earning a premium via time series momentum for providing liquidity to hedgers. Decomposing futures returns into the effect of price changes, which captures information diffusion, and the roll return, which captures how hedging pressure affects the shape of the futures curve, we find that shocks to both price changes and roll returns are associated with time series momentum profits. However, only shocks to price changes partially reverse, consistent with behavioral theories of delayed over-reaction to information, and not hedging pressure.

Time series momentum represents one of the most direct tests of the random walk hypothesis and a number of prominent behavioral and rational asset pricing theories. Our findings present new evidence and challenges for those theories and for future research.

Appendix A. Data sources

A.1. Equity indexes

The universe of equity index futures consists of the following nine developed equity markets: SPI 200 (Australia), CAC 40 (France), DAX (Germany), FTSE/MIB (Italy), TOPIX (Japan), AEX (Netherlands), IBEX 35 (Spain), FTSE 100 (UK), and S&P 500 (U.S). Futures returns are obtained from Datastream. We use MSCI country-level index returns prior to the availability of futures returns.

A.2. Bond indexes

The universe of bond index futures consists of the following 13 developed bond markets: Australia 3-year Bond, Australia 10-year Bond, Euro Schatz, Euro Bobl, Euro Bund, Euro Buxl, Canada 10-year Bond, Japan 10-year Bond (TSE), Long Gilt, US 2-year Note, US 5-year Note, US 10-year Note, and US Long Bond. Futures returns are obtained from Datastream. We use JP Morgan country-level bond index returns prior to the availability of futures returns. We scale daily returns to a constant duration of 2 years for 2- and 3-year bond futures, 4 years for 5-year bond futures, 7 years for 10-year bond futures, and 20 years for 30-year bond futures.

A.3. Currencies

The universe of currency forwards covers the following ten exchange rates: Australia, Canada, Germany spliced with the Euro, Japan, New Zealand, Norway, Sweden, Switzerland, UK, and US. We use spot and forward interest rates from Citigroup to calculate currency returns going back to 1989 for all the currencies except for CAD and NZD, which go back to 1992 and 1996, respectively. Prior to that, we use spot exchange rates from Datastream and Interbank Offered Rate (IBOR) short rates from Bloomberg to calculate returns.

A.4. Commodities

We cover 24 different commodity futures. Our data on Aluminum, Copper, Nickel, Zinc are from London Metal Exchange (LME), Brent Crude, Gas Oil, Cotton, Coffee, Cocoa, Sugar are from Intercontinental Exchange (ICE), Live Cattle, Lean Hogs are from Chicago Mercantile Exchange (CME), Corn, Soybeans, Soy Meal, Soy Oil, Wheat are from Chicago Board of Trade (CBOT), WTI Crude, RBOB Gasoline spliced with Unleaded Gasoline, Heating Oil, Natural Gas are from New York Mercantile Exchange (NYMEX), Gold, Silver are from New York Commodities Exchange (COMEX), and Platinum from Tokyo Commodity Exchange (TOCOM).

A.5. Position of Traders Data

We obtain speculator net length and open interest data from the CFTC Commitments of Traders Report Web site for the following futures: Corn, Soybeans, Soy Meal, Soy Oil, Wheat traded on Chicago Board of Trade (CBOT), Live Cattle, Lean Hogs, Australian Dollar, Canadian Dollar, Swiss Franc, British Pound, Japanese Yen, Euro FX, New Zealand Dollar, S&P 500 traded on Chicago Mercantile Exchange (CME), Cotton, Coffee, Cocoa, Sugar traded on Intercontinental Exchange (ICE), WTI Crude, RBOB Gasoline spliced with Unleaded Gasoline, Heating Oil, Natural Gas traded on New York Mercantile Exchange (NYMEX), and Gold, Silver traded on New York Commodities Exchange (COMEX). The data cover the period January 1986 to December 2009.

Endnotes

[1]Cross-sectional momentum has been documented in US equities (Jegadeesh and Titman, 1993; Asness, 1994), other equity markets (Rouwenhorst, 1998), industries (Moskowitz and Grinblatt, 1999), equity indexes (Asness, Liew, and Stevens, 1997; Bhojraj and Swaminathan, 2006), currencies (Shleifer and Summers, 1990), commodities (Erb and Harvey, 2006; Gorton, Hayashi, and Rouwenhorst, 2008), and global bond futures (Asness, Moskowitz, and Pedersen, 2010). Garleanu and Pedersen (2009) show how to trade optimally on momentum and reversal in light of transaction costs, and DeMiguel, Nogales, and Uppal (2010) show how to construct an optimal portfolio based on stocks' serial dependence and find outperformance out-of-sample. Our study is related to but different from Asness, Moskowitz, and Pedersen (2010), who study cross-sectional momentum and value strategies across several asset classes including individual stocks. We complement their study by examining time series momentum and its relation to cross-sectional momentum and hedging pressure in some of the same asset classes.

[2]Under-reaction can result from the slow diffusion of news (Hong and Stein, 1999), conservativeness and anchoring biases (Barberis, Shleifer, and Vishny, 1998; Edwards, 1968), or the disposition effect to sell winners too early and hold on to losers too long (Shefrin and Statman, 1985; Frazzini, 2006). Over-reaction can be caused by positive feedback trading (De Long, Shleifer, Summers, and Waldmann, 1990; Hong and Stein, 1999), over-confidence and self-attribution confirmation biases (Daniel, Hirshleifer, and Subrahmanyam, 1998), the representativeness heuristic (Barberis, Shleifer, and Vishny, 1998; Tversky and Kahneman, 1974), herding (Bikhchandani, Hirshleifer, and Welch, 1992), or general sentiment (Baker and Wurgler, 2006, 2007).

[3]However, this result differs from Lewellen's (2002) finding for equity portfolio returns that temporal lead-lag effects, rather than auto-covariances, appear to be the most significant contributor to cross-sectional momentum. Chen and Hong (2002) provide a different interpretation and decomposition of the Lewellen (2002) portfolios that is consistent with auto-covariance being the primary driver of stock momentum.

[4]We also confirm the time series momentum returns are robust among more illiquid instruments such as illiquid commodities (feeder cattle, Kansas wheat, lumber, orange juice, rubber, tin), emerging market currencies and equities, and more illiquid fixed income futures (not reported).

[5]Bessembinder (1992) and de Roon, Nijman, and Veld (2000) compute returns on futures contracts similarly and also find that futures returns are highly correlated with spot returns on the same underlying asset.

[6]While commercial traders likely predominantly include hedgers, some may also be speculating, which introduces some noise into the analysis in terms of our classification of speculative and hedging trades. However, the potential attenuation bias associated with such misclassification may only weaken our results.

[7]The regression results are qualitatively similar if we run OLS without adjusting for each security's volatility.

[8]Also, this portfolio construction implies a use of margin capital of about 5–20%, which is well within what is feasible to implement in a real-world portfolio.

[9] We thank the referee for asking for this out-of-sample study of old data.

[10]Asness, Moskowitz, and Pedersen (2010) exclude the most recent month when computing 12-month cross-sectional momentum. For consistency, we follow that convention here, but our results do not depend on whether the most recent month is excluded or not.

[11]This relation between mean returns is exact if annual returns are computed by summing over monthly returns. If returns are compounded, this relation is approximate, but an exact relation is straightforward to derive, e.g., using the separate means of monthly and annual returns.

[12]Lequeux and Acar (1998) also show that a simple timing trade tracks the performance of currency hedge funds.

References

Ahn, D.H., Boudoukh, J., Richardson, M., Whitelaw., R.F., 2002. Partial adjustment or stale prices? Implications from stock index and futures return autocorrelations. The Review of Financial Studies 15, 655–689.

Ahn, D.H., Conrad, J., Dittmar, R.F., 2003. Risk adjustment and trading strategies. The Review of Financial Studies 16, 459–485.

Asness, C., 1994. Variables that Explain Stock Returns. Ph.D. Dissertation. University of Chicago.

Asness, C., Liew, J.M., Stevens, R.L., 1997. Parallels between the cross-sectional predictability of stock and country returns. The Journal of Portfolio Management 23, 79–87.

Asness, C., Moskowitz, T.J., Pedersen, L.H., 2010. Value and momentum everywhere. AQR Capital, University of Chicago, and National Bureau of Economic Research.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645–1680.

Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. Journal of Economic Perspectives 21, 129–157.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Berk, J., Green, R.C., Naik, V., 1999. Optimal investment, growth options and security returns. Journal of Finance 54, 1153–1608.

Bessembinder, H., 1992. Systematic risk, hedging pressure, and risk premiums in futures markets. The Review of Financial Studies 5, 637–667.

Bhojraj, S., Swaminathan, B., 2006. Macromomentum: returns predictability in international equity indices. The Journal of Business 79, 429–451.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, 992–1026.

Brunnermeier, M., Pedersen, L.H., 2009. Market liquidity and funding liquidity. The Review of Financial Studies 22, 2201–2238.

Chen, J., Hong, H., 2002. Discussion of 'Momentum and autocorrelation in stock returns. The Review of Financial Studies 15, 565–573.

Cutler, D.M., Poterba, J.M., Summers, L.H., 1991. Speculative dynamics. Review of Economic Studies 58, 529–546.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. A theory of over- confidence, self-attribution, and security market under- and over- reactions. Journal of Finance 53, 1839–1885.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Positive feedback investment strategies and destabilizing rational speculation. The Journal of Finance 45, 379–395.

DeMiguel, V., Nogales, F.J., Uppal, R., 2010. Stock Return Serial Dependence and Out-of-sample Portfolio Performance. Unpublished working paper. London Business School.

de Roon, F., Nijman, T.E., Veld, C., 2000. Hedging pressure effects in futures markets. Journal of Finance 55, 1437–1456.

Edwards, W., 1968. Conservatism in human information processing. In: Kleinmutz, B. (Ed.), Formal Representation of Human Judgment, John Wiley and Sons, New York, pp. 17–52.

Erb, C., Harvey, C., 2006. The tactical and strategic value of commodity futures. Financial Analysts Journal 62, 69–97.

Fama, E., French, K., 1987. Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. Journal of Business 60, 55–73.

Fama, E., French, K., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.

Fama, E., French, K., 1993. Common risk factors in the returns of stocks and bonds. Journal of Financial Economics 33, 3–56.

Frazzini, A., 2006. The disposition effect and underreaction to news. Journal of Finance 61, 2017–2046.

Fung, W., Hsieh, D., 2001. The risk in hedge fund strategies: theory and evidence from trend followers. Review of Financial Studies 14, 313–341.

Garleanu, N., Pedersen, L.H., 2009. Dynamic Trading with Predictable Returns and Transaction Costs. Unpublished working paper. Berkeley and New York University.

Garleanu, N., Pedersen, L.H., 2011. Margin-based asset pricing and deviations from the law of one price. The Review of Financial Studies 24, 1980–2022.

Gorton, G.B., Hayashi, F., Rouwenhorst, K.G., 2008. The Fundamentals of Commodity Futures Returns. Unpublished working paper. Yale ICF.

Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Johnson, T.C., 2002. Rational momentum effects. Journal of Finance 57, 585–608.

Keynes, J.M., 1923. Some aspects of commodity markets. Manchester Guardian Commercial, European Reconstruction Series Section 13, 784–786.

Lequeux, P., Acar, E., 1998. A dynamic index for managed currencies funds using CME currency contracts. European Journal of Finance 4, 311–330.

Lewellen, J., 2002. Momentum and autocorrelation in stock returns. Review of Financial Studies 15, 533–564.

Liu, L., Zhang, L., 2008. Momentum profits, factor pricing, and macro- economic risk. Review of Financial Studies 21, 2417–2448.

Lo, A., Mackinlay, C., 1988. Stock prices do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Lo, A., Mackinlay, C., 1990. When are contrarian profits due to stock market overreaction? Review of Financial Studies 3, 175–206.

Moskowitz, T.J., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54, 1249–1290.

Poterba, J.M., Summers, L.H., 1988. Mean reversion in stock prices: evidence and implications. Journal of Financial Economics 22, 27–59.

Qiu, L.X., Welch, I., 2006. Investor Sentiment Measures. Unpublished working paper. Brown University.

Richardson, M., 1993. Temporary components of stock prices: a skeptic's view. Journal of Business & Economic Statistics 11, 199–207.

Rouwenhorst, K.G., 1998. International momentum strategies. Journal of Finance 53, 267–284.

Sagi, J.S., Seasholes, M.S., 2007. Firm-specific attributes and the cross-section of momentum. Journal of Financial Economics 84, 389–434.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: theory and evidence. Journal of Finance 40, 777–791.

Shleifer, A., Summers, L.H., 1990. The noise trader approach to finance. Journal of Economic Perspectives 4, 19–33.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.

Value and Momentum Everywhere

*Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen**

We find consistent value and momentum return premia across eight diverse markets and asset classes, and a strong common factor structure among their returns. Value and momentum returns correlate more strongly across asset classes than passive exposures to the asset classes, but value and momentum are negatively correlated with each other, both within and across asset classes. Our results indicate the presence of common global risks that we characterize with a three-factor model. Global funding liquidity risk is a partial source of these patterns, which are identifiable only when examining value and momentum jointly across markets. Our findings present a challenge to existing behavioral, institutional, and rational asset pricing theories that largely focus on U.S. equities.

Two of the most studied capital market phenomena are the relation between an asset's return and the ratio of its "long-run" (or book) value relative to its current market value, termed the "value" effect, and the relation between an asset's return and its recent relative performance history, termed the "momentum" effect. The returns to value and momentum strategies have become central to the market efficiency debate

*Asness is at AQR Capital Management; Moskowitz is at the University of Chicago Booth School of Business and NBER and is a consultant to AQR Capital; and Pedersen is at the New York University Stern School of Business, Copenhagen Business School, AQR, CEPR, FRIC, and NBER. We thank Aaron Brown, John Cochrane, Kent Daniel, Gene Fama, Kenneth French, Cam Harvey (the Editor), Ronen Israel, Robert Krail, John Liew, Harry Mamaysky, Michael Mendelson, Stefan Nagel, Lars Nielsen, Otto Van Hemert, Jeff Wurgler, and an anonymous referee for helpful comments, as well as seminar participants at the University of Chicago, Princeton University, Duke University, the Danish Society of Financial Analysts with Henrik Amilon and Asbjørn Trolle as discussants, and the NBER Summer Institute Asset Pricing Meetings with Kent Daniel as a discussant. We also thank Gunner Arnson, Radhika Gupta, Kelvin Hu, Sarah Jiang, Adam Klein, Ari Levine, Len Lorilla, Wes McKinney, and Karthik Sridharan for research assistance. AQR Capital invests in, among other things, value and momentum strategies. The views expressed here are those of the authors and not necessarily those of any affiliated institution.

Reprinted from *The Journal of Finance*, June 2013, Vol. 68 (3) 929–985, with permission.

and the focal points of asset pricing studies, generating numerous competing theories for their existence. We offer new insights into these two market anomalies by examining their returns jointly across eight diverse markets and asset classes. We find significant return premia to value and momentum in every asset class and strong comovement of their returns across asset classes, both of which challenge existing theories for their existence. We provide a simple three-factor model that captures the global returns across asset classes, the Fama–French U.S. stock portfolios, and a set of hedge fund indices.

The literature on market anomalies predominantly focuses on U.S. individual equities, and often examines value or momentum separately. In the rare case in which value and momentum are studied outside of U.S. equities, they are also typically studied in isolation—separate from each other and separate from other markets. We uncover unique evidence and features of value and momentum by examining them jointly across eight different markets and asset classes (individual stocks in the United States, the United Kingdom, continental Europe, and Japan; country equity index futures; government bonds; currencies; and commodity futures).[1] Although some of these markets have been analyzed in isolation, our joint approach provides unique evidence on several key questions about these pervasive market phenomena. Specifically, how much variation exists in value and momentum premia across markets and asset classes? How correlated are value and momentum returns across these diverse markets and asset classes with different geographies, structures, investor types, and securities? What are the economic drivers of value and momentum premia and their correlation structure? What is a natural benchmark model for portfolios of global securities across different asset classes?

We find consistent and ubiquitous evidence of value and momentum return premia across all the markets we study, including value and momentum in government bonds and value effects in currencies and commodities, which are all novel to the literature. Our broader set of portfolios generates much larger cross-sectional dispersion in average returns than those from U.S. stocks only, providing a richer set of asset returns that any asset pricing model should seek to explain. Most strikingly, we discover significant comovement in value and momentum strategies across diverse asset classes. Value strategies are positively correlated with other value strategies across otherwise unrelated markets, and momentum strategies are positively correlated with other momentum strategies globally. However, value and momentum are negatively correlated with each other within and across asset classes.

The striking comovement pattern across asset classes is one of our central findings and suggests the presence of common global factors related to value and momentum. This common risk structure implies a set of results that we investigate further. For example, using a simple three-factor model consisting of a global market index, a zero-cost value strategy applied across all asset classes, and a zero-cost momentum strategy across all assets, we capture the comovement and the cross section of average returns both globally across asset classes and locally within an asset class. We further show that the global three-factor model does a good job capturing the returns to the Fama and French U.S. stock portfolios as well as a set of hedge fund indices. Our use of a simple three-factor model in pricing a variety of assets globally is motivated by finance research and practice becoming increasingly global and the desire to have a single model that describes returns across asset classes rather than specialized models for each market. We show that separate factors for value and momentum best explain the data, rather than a single factor, since both strategies produce positive returns on average yet are negatively correlated.[2]

We then investigate the source of this common global factor structure. We find only modest links to macroeconomic variables such as the business cycle, consumption, and default risk. However, we find significant evidence that liquidity risk is negatively related to value and positively related to momentum globally across asset classes. Pástor and Stambaugh (2003) and Sadka (2006) find that measures of liquidity risk are positively related to momentum in U.S. individual stocks. We show that this link is also present in other markets and asset classes, and that value returns are significantly negatively related to liquidity risk globally, implying that part of the negative correlation between value and momentum is driven by opposite signed exposure to liquidity risk. Separating market from funding liquidity (see Brunnermeier and Pedersen (2009)), we further find that the primary link between value and momentum returns comes from funding risk, whose importance has increased over time, particularly after the funding crisis of 1998. Importantly, these results cannot be detected by examining a single market in isolation. The statistical power gained by looking across many markets at once—a unique feature of our analysis—allows these factor exposures to be revealed.

In terms of economic magnitudes, however, liquidity risk can only explain a small fraction of value and momentum return premia and comovement. While liquidity risk may partly explain the positive risk premium associated with momentum, because value loads negatively on liquidity risk, the positive premium associated with value becomes an even deeper puzzle. Moreover, a simple equal-weighted combination of value and momentum is immune to liquidity risk and generates substantial abnormal returns. Hence, funding liquidity risk can only provide a partial and incomplete explanation for momentum, but cannot explain the value premium or the value and momentum combination premium.

The evidence we uncover sheds light on explanations for the existence of value and momentum premia. For example, a strong correlation structure among these strategies in otherwise unrelated asset classes may indicate the presence of common global risk factors for which value and momentum premia provide compensation. Conversely, such correlation structure is not a prediction of existing behavioral models (e.g., Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999)).

In addition to assuaging data mining concerns, evidence of consistent value and momentum premia across diverse asset classes may be difficult to reconcile under rational asset pricing theories that rely on firm investment risk or firm growth options as explanations for the value and momentum premia,[3] which are predominantly motivated by firm equity. These theories seem ill equipped to explain the same and correlated effects we find in currencies, government bonds, and commodities.

We also highlight that studying value and momentum jointly is more powerful than examining each in isolation. The negative correlation between value and momentum strategies and their high positive expected returns implies that a simple combination of the two is much closer to the efficient frontier than either strategy alone, and exhibits less variation across markets and over time. The return premium to a combination of value and momentum applied across all asset classes therefore presents an even bigger challenge for asset pricing theories to accommodate (e.g., Hansen and Jagannathan (1997)).

Our work also relates to the recent literature on global asset pricing. Fama and French (2012) examine the returns to size, value, and momentum in individual stocks across global equity markets and find consistent risk premia across markets. Considering both global equities and other global asset classes, Frazzini and Pedersen (2010) find consistent returns to "betting against beta," Koijen et al. (2012) document global "carry" returns, and Moskowitz,

Ooi, and Pedersen (2012) present global evidence of "time series momentum." Time-series momentum is a timing strategy using each asset's own past returns, which is separate from the cross-sectional momentum strategies we study here. Focusing on this different time-series phenomenon, Moskowitz, Ooi, and Pedersen (2012) examine returns to futures contracts on equity indices, bonds, currencies, and commodities—ignoring individual stocks, which comprise half our study here—and address a different set of questions. Our focus is on the interaction between cross-sectional momentum and value strategies and their common factor structure globally, where we find striking comovement across assets and a link to liquidity risk.

The link to funding liquidity risk may also be consistent with global arbitrage activity in the face of funding constraints influencing value and momentum returns (Brunnermeier and Pedersen (2009)). Why does momentum load positively on liquidity risk and value load negatively? A simple and natural story might be that momentum represents the most popular trades, as investors chase returns and flock to the assets whose prices appreciated most recently. Value, on the other hand, represents a contrarian view. When a liquidity shock occurs, investors engaged in liquidating sell-offs (due to cash needs and risk management) will put more price pressure on the most popular and crowded trades, such as high momentum securities, as everyone runs for the exit at the same time (Pedersen (2009)), while the less crowded contrarian/value trades will be less affected.

Vayanos and Wooley (2012) offer a model of value and momentum returns due to delegated management that may be consistent with these results. They argue that flows between investment funds can give rise to momentum effects from inertia due to slow moving capital, and eventually push prices away from fundamentals causing reversals or value effects. Correlation of value and momentum across different asset classes could also be affected by funds flowing simultaneously across asset classes, which could in turn be impacted by funding liquidity. However, matching the magnitude of our empirical findings remains an open question.

The paper proceeds as follows. Section I outlines our data and portfolio construction. Section II examines the performance of value and momentum across asset classes and documents their global comovement. Section III investigates the source of common variation by examining macroeconomic and liquidity risk, and Section IV offers an empirically motivated three-factor model to describe the cross section of returns across asset classes. Section V briefly discusses the robustness of our results to implementation issues. Section VI concludes by discussing the implications of our findings.

I. Data and Portfolio Construction

We describe our data and methodology for constructing value and momentum portfolios across markets and asset classes.

A. Data

A.1. Global Individual Stocks

We examine value and momentum portfolios of individual stocks globally across four equity markets: the United States, the United Kingdom, continental Europe, and Japan. The

U.S. stock universe consists of all common equity in CRSP (sharecodes 10 and 11) with a book value from Compustat in the previous 6 months, and at least 12 months of past return history from January 1972 to July 2011. We exclude ADRs, REITs, financials, closed-end funds, foreign shares, and stocks with share prices less than $1 at the beginning of each month. We limit the remaining universe of stocks in each market to a very liquid set of securities that could be traded for reasonably low cost at reasonable trading volume size. Specifically, we rank stocks based on their beginning-of-month market capitalization in descending order and include in our universe the number of stocks that account cumulatively for 90% of the total market capitalization of the entire stock market.[4] This universe corresponds to an extremely liquid and tradeable set of securities. For instance, over our sample period this universe corresponds to the largest 17% of firms on average in the United States. For the U.S. stock market, at the beginning of the sample period (January 1972) our universe consists of the 354 largest firms and by the end of our sample period (July 2011) the universe comprises the 676 largest names. Hence, our sample of U.S. equities is significantly larger and more liquid than the Russell 1000.

For stocks outside of the United States, we use Datastream data from the United Kingdom, continental Europe (across all European stock markets, excluding the United Kingdom), and Japan. We restrict the universe in each market using the same criteria used for U.S. stocks. On average over the sample period, our universe represents the largest 13%, 20%, and 26% of firms in the United Kingdom, Europe, and Japan, respectively. Data on prices and returns come from Datastream, and data on book values are from Worldscope.

Most studies of individual stocks examine a much broader and less liquid set of securities. We restrict our sample to a much more liquid universe (roughly the largest 20% of stocks in each market) to provide reasonable and conservative estimates of an implementable set of trading strategies and to better compare those strategies with the set of strategies we employ in index futures, currencies, government bonds, and commodity futures, which are typically more liquid instruments. Our results are conservative since value and momentum premia are larger among smaller, less liquid securities over the sample period we study.[5]

All series are monthly and end in July 2011. The U.S. and U.K. stock samples begin in January 1972. The Europe and Japan stock samples begin in January 1974. The average (minimum) number of stocks in each market over their respective sample periods is 724 (354) in the United States, 147 (76) in the United Kingdom, 290 (96) in Europe, and 471 (148) in Japan.

A.2. Global Equity Indices

The universe of country equity index futures consists of the following 18 developed equity markets: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Hong Kong, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States.[6] Returns and price data as well as book values are obtained from MSCI and Bloomberg. The sample covers the period January 1978 to July 2011, with the minimum number of equity indices being 8 and all 18 indices represented after 1980. The returns on the country equity index futures do not include any returns on collateral from transacting in futures contracts, hence these are comparable to returns in excess of the risk-free rate.

A.3. Currencies

We obtain spot exchange rates from Datastream covering the following 10 currencies: Australia, Canada, Germany (spliced with the Euro), Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The data cover the period January 1979 to July 2011, where the minimum number of currencies is 7 at any point in time and all 10 currencies are available after 1980. We compute returns from currency forward contracts or MSCI spot price data and Libor rates, where currency returns are all dollar denominated and implicitly include the local interest rate differential.

A.4. Global Government Bonds

Bond index returns come from Bloomberg and Morgan Markets, short rates and 10-year government bond yields are from Bloomberg, and inflation forecasts are obtained from investment bank analysts' estimates as compiled by Consensus Economics. We obtain government bond data for the following 10 countries: Australia, Canada, Denmark, Germany, Japan, Norway, Sweden, Switzerland, the United Kingdom, and the United States over the period January 1982 to July 2011, where the minimum number of country bond returns is 5 at any point in time and all 10 country bonds are available after 1990.

A.5. Commodity Futures

We cover 27 different commodity futures obtained from several sources. Data on Aluminum, Copper, Nickel, Zinc, Lead, and Tin are from the London Metal Exchange (LME). Brent Crude and Gas Oil are from the Intercontinental Exchange (ICE). Live Cattle, Feeder Cattle, and Lean Hogs are from the Chicago Mercantile Exchange (CME). Corn, Soybeans, Soy Meal, Soy Oil, and Wheat are from the Chicago Board of Trade (CBOT). WTI Crude, RBOB Gasoline, Heating Oil, and Natural Gas are from the New York Mercantile Exchange (NYMEX). Gold and Silver are from the New York Commodities Exchange (COMEX). Cotton, Coffee, Cocoa, and Sugar are from New York Board of Trade (NYBOT), and Platinum data are from the Tokyo Commodity Exchange (TOCOM). The sample covers the period January 1972 to July 2011, with the minimum number of commodities being 10 at any point in time and all 27 commodities available after 1995.

Returns for commodity futures are calculated as follows. Each day we compute the daily excess return of the most liquid futures contract, which is typically the nearest- or next nearest-to-delivery contract, and then compound the daily returns to a total return index from which we compute returns at a monthly horizon. Bessembinder (1992), de Roon, Nijman, and Veld (2000), Moskowitz, Ooi, and Pedersen (2012), and Koijen et al. (2012) compute futures returns similarly. All returns are denominated in U.S. dollars and do not include the return on collateral associated with the futures contract.

B. Value and Momentum Measures

To measure value and momentum, we use the simplest and, to the extent a standard exists, most standard measures. We are not interested in coming up with the best predictors of

returns in each asset class. Rather, our goal is to maintain a simple and fairly uniform approach that is consistent across asset classes and minimizes the pernicious effects of data snooping. As such, if data snooping can be avoided, our results may therefore understate the true gross returns to value and momentum available from more thoughtfully chosen measures.

For individual stocks, we use the common value signal of the ratio of the book value of equity to market value of equity, or book-to-market ratio, *BE/ME* (see Fama and French (1992, 1993) and Lakonishok, Shleifer, and Vishny (1994)), of the stock.[7] Book values are lagged 6 months to ensure data availability to investors at the time, and the most recent market values are used to compute the ratios. For the purposes of this paper, using lagged or contemporary prices rather than market values matched contemporaneously in time as in Fama and French (1992) is not important. When using more recent prices in the value measure, the negative correlation between value and momentum is more negative and the value premium is slightly reduced, but our conclusions are not materially affected. A combination of value and momentum—one of the themes in this paper—obtains nearly identical pricing results regardless of whether we lag price in the value measure. Asness and Frazzini (2012) investigate this issue more thoroughly and argue that using contemporaneous market values can be important and ease interpretation when examining value in the presence of momentum, as we do in this paper. Gerakos and Linnainmaa (2012) decompose value into book and market components and find that the market value of equity drives most of the relevant pricing information.

For momentum, we use the common measure of the past 12-month cumulative raw return on the asset (see Jegadeesh and Titman (1993), Asness (1994), Fama and French (1996), and Grinblatt and Moskowitz (2004)), skipping the most recent month's return, *MOM2–12*. We skip the most recent month, which is standard in the momentum literature, to avoid the 1-month reversal in stock returns, which may be related to liquidity or microstructure issues (Jegadeesh (1990), Lo and MacKinaly (1990), Boudoukh, Richardson, and Whitelaw (1994), Asness (1994), Grinblatt and Moskowitz (2004)).[8]

For all other asset classes, we attempt to define similar value and momentum measures. Momentum is straightforward since we can use the same measure for all asset classes, namely, the return over the past 12 months skipping the most recent month. While excluding the most recent month of returns is not necessary for some of the other asset classes we consider because they suffer less from liquidity issues (e.g., equity index futures and currencies), we do so to maintain uniformity across asset classes. Momentum returns for these asset classes are in fact stronger when we don't skip the most recent month, hence our results are conservative.

For measures of value, attaining uniformity is more difficult because not all asset classes have a measure of book value. For these assets, we try to use simple and consistent measures of value. For country indices, we use the previous month's *BE*/*ME* ratio for the MSCI index of the country. For commodities, we define value as the log of the spot price 5 years ago (actually, the average spot price from 4.5 to 5.5 years ago), divided by the most recent spot price, which is essentially the negative of the spot return over the last 5 years. Similarly, for currencies, our value measure is the negative of the 5-year return on the exchange rate, measured as the log of the average spot exchange rate from 4.5 to 5.5 years ago divided by the spot exchange rate today minus the log difference in the change in CPI in the foreign country relative to the U.S. over the same period. The currency value

measure is therefore the 5-year change in purchasing power parity. For bonds, we use the 5-year change in the yields of 10-year bonds as our value measure, which is similar to the negative of the past 5-year return. These long-term past return measures of value are motivated by DeBondt and Thaler (1985), who use similar measures for individual stocks to identify "cheap" and "expensive" firms. Fama and French (1996) show that the negative of the past 5-year return generates portfolios that are highly correlated with portfolios formed on *BE*/*ME*, and Gerakos and Linnainmaa (2012) document a direct link between past returns and *BE*/*ME* ratios. Theory also suggests a link between long-term returns and book-to-market value measures (e.g., Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999), and Vayanos and Wooley (2012)).

In the Internet Appendix accompanying this paper, we show that individual stock portfolios formed from the negative of past 5-year returns are highly correlated with those formed on *BE*/*ME* ratios in our sample.[9] For example, among U.S. stocks the correlation between returns to a value factor formed from the negative of the past 5-year return and the returns formed from *BE*/*ME* sorts is 0.83. In the United Kingdom, Europe, and Japan the correlation between portfolio returns formed on negative past 5-year returns and *BE*/*ME* ratios is similarly high. Globally, a value factor averaged across all four stock markets estimated from negative past 5-year return sorts has a correlation of 0.86 with a value factor formed from *BE*/*ME* sorts. Hence, using past 5-year returns to measure value seems reasonable.

C. Value and Momentum Portfolios: 48 New Test Assets

Using the measures above, we construct a set of value and momentum portfolios within each market and asset class by ranking securities within each asset class by value or momentum and sorting them into three equal groups. We then form three portfolios—high, middle, and low—from these groups, where for individual stocks we value weight the returns in the portfolios by their beginning-of-month market capitalization, and for the nonstock asset classes we equal weight securities.[10] Given that our sample of stocks focuses exclusively on very large and liquid securities in each market, typically the largest quintile of securities, further value weighting the securities within this universe creates an extremely large and liquid set of portfolios that should yield very conservative results compared to typical portfolios used in the literature. Thus, we generate three portfolios—low, middle, and high—for each of the two characteristics—value and momentum—in each of the eight asset classes, producing $3 \times 2 \times 8 = 48$ test portfolios.

D. Value and Momentum Factors

We also construct value and momentum factors for each asset class, which are zero-cost long-short portfolios that use the entire cross section of securities within an asset class. For any security $i = 1,..., N$ at time t with signal Sit (value or momentum), we weight securities in proportion to their cross-sectional rank based on the signal minus the cross-sectional average rank of that signal. Simply using ranks of the signals as portfolio weights helps mitigate the influence of outliers, but portfolios constructed using the raw signals are similar and generate slightly better performance. Specifically, the weight on security i at time t is

$$w_{it}^{S} = c_t \left(\text{rank}\left(S_{it}\right) - \sum_i \text{rank}\left(S_{it}\right) / N \right), \tag{1}$$

where the weights across all stocks sum to zero, representing a dollar-neutral long-short portfolio. We include a scaling factor c_t such that the overall portfolio is scaled to one dollar long and one dollar short. The return on the portfolio is then

$$r_t^{S} = \sum_i w_{it}^{S} r_{it}, \text{ where } S \in \left(\text{value}, \text{momentum}\right). \tag{2}$$

We also construct a 50/50 equal combination (*COMBO*) factor of value and momentum, whose returns are

$$r_t^{COMBO} = 0.5 r_t^{VALUE} + 0.5 r_t^{MOM}. \tag{3}$$

These zero-cost signal-weighted portfolios are another way to examine the efficacy of value and momentum across markets and are used as factors in our pricing model. Although these factors are not value weighted, the set of securities used to generate them are extremely large and liquid. As we will show, the signal-weighted factor portfolios outperform simple portfolio sort spreads because security weights are a positive (linear) function of the signal, as opposed to the coarseness of only classifying securities into three groups. In addition, the factors are better diversified since more securities in the cross section are given nonzero weight and the weights are less extreme.

II. Value and Momentum Returns and Comovement

Table I shows the consistent performance of value and momentum, and their combination, within each of the major markets and asset classes we study. Other studies examine value and momentum in some of the same asset classes, but not in combination and not simultaneously across asset classes as we do here. In addition, we also discover new evidence for value and momentum premia in asset classes not previously studied—both value and momentum in government bonds and value effects in currencies and commodities. Our emphasis, however, is on the power of applying value and momentum everywhere at once.

A. Return Premia

Table I reports the annualized mean return, *t*-statistic of the mean, standard deviation, and Sharpe ratio of the low (P1), middle (P2), and high (P3) portfolios for value and momentum in each market and asset class as well as the high minus low (P3–P1) spread portfolio and the signal-weighted factor portfolio from equation (2). Also reported are the intercepts or alphas, and their *t*-statistics (in parentheses) from a time-series regression of each return series on the return of the market index for each asset class. The market index for the stock strategies is the MSCI equity index for each country; for country index futures it is the MSCI World Index; and for currencies, fixed income, and commodities, the benchmark is an equal-weighted basket of the securities in each asset class. The last two columns of Table I report the same statistics for the 50/50 combination of value and momentum for the

Table I Performance of Value and Momentum Portfolios across Markets and Asset Classes

Reported are the average raw excess (of the 1-month U.S. T-bill rate) return, *t*-statistic of the average return (in parentheses), standard deviation of returns, and Sharpe ratio of each value, momentum, and equal-weighted 50/50 value and momentum combination strategy in each market and asset class we study: U.S. stocks, U.K. stocks, Europe stocks, Japan stocks, country index futures, currencies, fixed income government bonds, and commodities. Also reported are the intercepts or alphas, and their *t*-statistics (in parentheses) from a time-series regression of each return series on the return of the market index for each asset class. The market index for the stock strategies is the MSCI equity index for each country for all of the individual stock strategies. The MSCI world index is used as the benchmark for strategies of country index futures. For currencies, fixed income, and commodities, the benchmark index is an equal-weighted basket of the securities in each asset class. In each market or asset class the universe of securities is first sorted by either value or momentum and then broken into three equal groups based on those sorts to form three portfolios—low, middle, and high—corresponding to portfolios P1, P2, and P3, respectively. For individual stock strategies (Panel A), stocks within the three portfolios are value weighted by their beginning-of-month capitalization, and for nonstock asset classes (Panel B), securities are equal weighted in the portfolios. Also reported is the high minus low spread in returns (P3–P1) as well as a rank-weighted factor portfolio ("Factor"), which is a zero-investment portfolio that weights each asset in proportion to its rank based on either value or momentum, following equation (1). The 50/50 value/momentum combination strategies are an equal-weighted average of the value and momentum spread strategies (P3–P1 and Factor) for each market/asset class. Results are also reported for an average of all individual stock strategies across all stock markets ("Global stocks"), across all nonstock asset classes ("Global other asset classes"), and across all markets and asset classes ("Global all asset classes"), where average return series are computed using equal volatility weights across the markets and asset classes to account for the large differences in volatility across asset classes (e.g., fixed income vs. commodities). Panel C also reports results for alternative measures of value for fixed income securities. Finally, the last row for each asset class reports the correlation between value and momentum zero cost residual returns from the benchmark in each market or asset class. Statistics are computed from monthly return series but are reported as annualized numbers.

Panel A: Individual Stock Portfolios

		Value Portfolios					Momentum Portfolios					50/50 Combination	
		P1	P2	P3	P3–P1	Factor	P1	P2	P3	P3–P1	Factor	P3–P1	Factor
U.S. stocks 01/1972 to 07/2011	Mean	9.5%	10.6%	13.2%	3.7%	3.9%	8.8%	9.7%	14.2%	5.4%	7.7%	4.6%	5.8%
	(*t*-stat)	(3.31)	(4.33)	(5.19)	(1.83)	(1.66)	(2.96)	(4.14)	(4.82)	(2.08)	(2.84)	(3.98)	(5.40)
	Stdev	17.9%	15.4%	15.9%	12.8%	14.8%	18.6%	14.8%	18.5%	16.4%	17.0%	7.2%	6.8%
	Sharpe	0.53	0.69	0.83	0.29	0.26	0.47	0.66	0.77	0.33	0.45	0.63	0.86
	Alpha	−1.7%	0.8%	3.6%	5.3%	5.8%	−2.3%	0.2%	3.7%	6.0%	8.7%	5.7%	7.2%
	(*t*-stat)	(−1.59)	(1.02)	(3.17)	(2.66)	(2.49)	(−1.68)	(0.29)	(2.34)	(2.30)	(3.22)	(5.05)	(7.06)
									Correlation (Val, Mom) =			−0.53	−0.65
U.K. stocks 01/1972 to 07/2011	Mean	10.8%	12.5%	15.3%	4.5%	5.5%	9.2%	13.8%	15.2%	6.0%	7.2%	6.3%	7.2%
	(*t*-stat)	(3.17)	(3.48)	(4.12)	(1.83)	(2.10)	(2.32)	(3.81)	(4.04)	(2.37)	(3.00)	(4.23)	(5.85)
	Stdev	18.6%	19.7%	20.3%	13.4%	14.4%	24.9%	22.7%	23.7%	15.9%	15.0%	8.1%	6.7%
	Sharpe	0.58	0.64	0.75	0.33	0.38	0.37	0.61	0.64	0.38	0.48	0.77	1.07
	Alpha	−0.2%	0.5%	3.2%	3.5%	4.4%	−3.2%	2.1%	3.5%	6.7%	8.0%	6.0%	7.2%
	(*t*-stat)	(−0.17)	(0.42)	(2.03)	(1.47)	(1.74)	(−2.13)	(2.06)	(2.31)	(2.66)	(3.36)	(4.05)	(5.84)
									Correlation (Val, Mom) =			−0.43	−0.62

(*Continued*)

Table I *(Continued)*

		Value Portfolios					Momentum Portfolios					50/50 Combination	
		P1	P2	P3	P3–P1	Factor	P1	P2	P3	P3–P1	Factor	P3–P1	Factor
Europe stocks	Mean	11.8%	14.6%	16.7%	4.8%	5.2%	9.2%	13.3%	17.3%	8.1%	9.8%	5.9%	6.9%
01/1974 to	(*t*-stat)	(3.53)	(4.43)	(4.61)	(2.32)	(2.95)	(2.72)	(4.65)	(5.56)	(3.37)	(4.59)	(4.77)	(6.55)
07/2011	Stdev	18.3%	18.0%	19.8%	11.5%	9.7%	20.6%	17.5%	19.0%	14.7%	13.1%	6.8%	5.8%
	Sharpe	0.64	0.81	0.84	0.42	0.54	0.44	0.76	0.91	0.55	0.75	0.87	1.20
	Alpha	(−0.4%)	2.2%	3.1%	3.5%	4.0%	(−3.5%)	2.2%	6.0%	9.1%	10.7%	6.1%	7.1%
	(*t*-stat)	(−0.30)	(2.06)	(2.57)	(1.71)	(2.32)	(−2.54)	(2.39)	(4.18)	(3.88)	(5.05)	(4.88)	(6.77)
									Correlation (Val, Mom) =			−0.52	−0.55
Japan stocks	Mean	2.6%	8.2%	14.7%	12.0%	10.2%	8.4%	9.9%	10.1%	1.7%	2.2%	6.4%	5.9%
01/1974 to 07/2011	(*t*-stat)	(0.61)	(2.02)	(3.69)	(4.31)	(4.22)	(2.19)	(2.94)	(2.69)	(0.57)	(0.81)	(4.28)	(4.80)
	Stdev	23.6%	22.1%	21.8%	15.3%	13.2%	23.5%	20.6%	23.1%	18.6%	16.5%	8.1%	6.7%
	Sharpe	0.11	0.37	0.67	0.79	0.77	0.36	0.48	0.44	0.09	0.13	0.78	0.88
	Alpha	−5.6%	0.1%	7.3%	13.0%	10.7%	−1.1%	0.8%	0.5%	1.7%	2.2%	6.8%	6.1%
	(*t*-stat)	(−3.36)	(0.12)	(3.95)	(4.71)	(4.47)	(−0.59)	(0.73)	(0.31)	(0.54)	(0.84)	(4.63)	(5.05)
									Correlation (Val, Mom) =			−0.60	−0.64
Global stocks	Mean	8.1%	11.0%	14.6%	6.2%	5.8%	8.5%	11.1%	14.1%	5.6%	7.1%	6.3%	6.8%
01/1972 to 07/2011	(*t*-stat)	(3.17)	(4.54)	(5.84)	(3.60)	(3.18)	(3.10)	(4.82)	(5.46)	(2.94)	(3.73)	(6.52)	(8.04)
	Stdev	16.6%	15.2%	15.7%	10.9%	11.4%	17.1%	14.5%	16.2%	12.0%	12.0%	6.1%	5.3%
	Sharpe	0.50	0.72	0.93	0.57	0.51	0.49	0.77	0.87	0.47	0.59	1.04	1.28
	Alpha	−2.3%	0.7%	4.2%	6.6%	6.1%	−3.3%	0.5%	3.1%	6.4%	8.1%	6.8%	7.5%
	(*t*-stat)	(−1.70)	(0.69)	(3.49)	(3.79)	(3.37)	(−3.00)	(1.00)	(2.78)	(3.37)	(4.31)	(7.09)	(8.98)
									Correlation (Val, Mom) =			−0.52	−0.60
Panel B: Other Asset Class Portfolios													
Country indices	Mean	3.1%	6.6%	9.1%	6.0%	5.7%	2.3%	5.8%	11.0%	8.7%	7.4%	7.3%	10.6%
01/1978 to	(*t*-stat)	(1.10)	(2.40)	(3.20)	(3.45)	(3.40)	(0.81)	(2.13)	(3.72)	(4.14)	(3.57)	(6.62)	(5.72)
07/2011	Stdev	16.2%	15.7%	16.2%	9.8%	9.5%	16.2%	15.4%	16.8%	11.9%	11.8%	6.3%	10.6%
	Sharpe	0.19	0.42	0.56	0.61	0.60	0.14	0.37	0.65	0.73	0.63	1.16	1.00
	Alpha	−3.2%	0.5%	2.7%	5.9%	5.3%	−3.9%	−0.3%	4.4%	8.2%	7.1%	7.1%	10.0%
	(*t*-stat)	(−3.24)	(0.48)	(2.76)	(3.45)	(3.24)	(−3.41)	(−0.40)	(4.00)	(4.00)	(3.47)	(6.49)	(5.47)
									Correlation (Val, Mom) =			−0.34	−0.37

(*Continued*)

Table I (*Continued*)

		Value Portfolios					Momentum Portfolios					50/50 Combination	
		P1	P2	P3	P3–P1	Factor	P1	P2	P3	P3–P1	Factor	P3–P1	Factor
Currencies	Mean	−0.5%	0.3%	2.8%	3.3%	3.9%	−0.7%	0.3%	2.8%	3.5%	3.0%	3.4%	5.6%
01/1979 to	(*t*-stat)	(−0.30)	(0.23)	(1.98)	(1.89)	(2.47)	(−0.40)	(0.20)	(1.91)	(1.90)	(1.77)	(3.51)	(3.89)
07/2011	Stdev	9.2%	8.3%	7.9%	9.7%	9.0%	9.4%	8.0%	8.2%	10.3%	9.6%	5.4%	8.0%
	Sharpe	−0.05	0.04	0.35	0.34	0.44	−0.07	0.04	0.34	0.34	0.32	0.63	0.69
	Alpha	−1.4%	−0.6%	2.0%	3.4%	4.1%	−1.6%	−0.6%	2.0%	3.6%	3.1%	3.5%	5.7%
	(*t*-stat)	(−1.53)	(−0.94)	(2.25)	(2.04)	(2.63)	(−1.58)	(−1.01)	(2.18)	(1.99)	(1.84)	(3.83)	(4.11)
									Correlation (Val, Mom) =			−0.42	−0.43
Fixed income	Mean	3.0%	4.0%	4.2%	1.1%	0.5%	3.8%	3.8%	4.2%	0.4%	1.0%	0.8%	0.7%
01/1982 to 07/2011	(*t*-stat)	(2.31)	(3.58)	(3.76)	(0.97)	(0.39)	(3.42)	(3.49)	(3.28)	(0.35)	(0.88)	(1.03)	(1.08)
	Stdev	7.0%	5.9%	5.9%	6.3%	6.4%	5.9%	5.9%	6.8%	6.0%	5.8%	4.0%	3.5%
	Sharpe	0.43	0.67	0.71	0.18	0.07	0.64	0.66	0.61	0.06	0.17	0.19	0.20
	Alpha	−1.3%	0.3%	0.7%	1.9%	1.4%	0.2%	0.3%	−0.1%	−0.3%	0.1%	0.8%	0.7%
	(*t*-stat)	(−1.87)	(0.51)	(1.03)	(1.68)	(1.21)	(0.34)	(0.48)	(−0.17)	(−0.29)	(0.08)	(1.10)	(1.15)
									Correlation (Val, Mom) =			−0.17	−0.35
Commodities	Mean	4.2%	4.1%	10.5%	6.3%	7.3%	0.7%	5.8%	13.1%	12.4%	11.5%	9.4%	17.1%
01/1972 to 07/2011	(*t*-stat)	(1.21)	(1.34)	(3.50)	(1.61)	(1.92)	(0.22)	(2.27)	(3.73)	(3.29)	(3.14)	(4.42)	(4.78)
	Stdev	21.5%	18.8%	18.5%	24.2%	23.7%	19.0%	15.9%	21.8%	23.4%	22.8%	13.1%	22.2%
	Sharpe	0.19	0.22	0.57	0.26	0.31	0.04	0.37	0.60	0.53	0.51	0.71	0.77
	Alpha	−2.9%	−2.4%	4.8%	7.7%	8.2%	−5.6%	0.4%	5.8%	11.4%	10.5%	9.5%	17.1%
	(*t*-stat)	(−1.39)	(−1.39)	(2.34)	(2.02)	(2.19)	(−2.95)	(0.26)	(2.71)	(3.06)	(2.89)	(4.57)	(4.82)
									Correlation (Val, Mom) =			−0.39	−0.46
Global other asset	Mean	2.2%	3.1%	5.7%	3.4%	3.6%	1.6%	3.4%	6.3%	4.6%	4.4%	4.0%	6.8%
classes 01/1972 to	(*t*-stat)	(1.88)	(2.97)	(5.63)	(3.08)	(3.42)	(1.49)	(3.61)	(5.30)	(3.88)	(3.83)	(6.39)	(7.03)
07/2011	Stdev	7.3%	6.4%	6.2%	6.4%	6.6%	6.7%	5.9%	7.3%	7.4%	7.1%	3.9%	6.0%
	Sharpe	0.30	0.48	0.91	0.50	0.55	0.24	0.58	0.85	0.63	0.62	1.03	1.14
	Alpha	−1.0%	−0.1%	3.0%	4.0%	3.9%	−1.4%	−0.7%	3.1%	4.5%	4.1%	4.2%	6.8%
	(*t*-stat)	(−1.26)	(−0.18)	(4.10)	(3.67)	(3.71)	(−1.85)	(−1.07)	(3.65)	(3.80)	(3.62)	(6.83)	(7.12)
											Correlation (Val, Mom) =	−0.40	−0.49
Global all asset	Mean	4.5%	6.1%	9.1%	4.6%	4.6%	4.2%	6.4%	9.2%	5.0%	5.4%	5.0%	6.8%
classes 01/1972	(*t*-stat)	(3.00)	(4.42)	(6.47)	(4.55)	(4.47)	(2.74)	(4.88)	(6.09)	(4.18)	(4.59)	(8.77)	(9.83)
to 07/2011	Stdev	9.3%	8.5%	8.7%	6.3%	6.4%	9.5%	8.1%	9.4%	7.5%	7.4%	3.5%	4.3%
	Sharpe	0.48	0.71	1.04	0.73	0.72	0.44	0.79	0.98	0.67	0.74	1.42	1.59
	Alpha	−2.0%	−0.2%	2.9%	4.8%	4.8%	−2.6%	0.3%	2.6%	5.2%	5.6%	5.0%	6.9%
	(*t*-stat)	(−2.29)	(−0.35)	(3.71)	(4.81)	(4.69)	(−3.07)	(0.43)	(3.04)	(4.34)	(4.76)	(9.00)	(10.03)
									Correlation (Val, Mom) =			−0.53	−0.60

(*Continued*)

Table I (*Continued*)

Panel C: Alternative Value Measures for Fixed Income

		Value Portfolios					50-50 Combination	
Fixed income 01/1983 to 07/2011		**P1**	**P2**	**P3**	**P3–P1**	**Factor**	**P3–P1**	**Factor**
	Mean	3.0%	4.0%	4.2%	1.1%	0.5%	0.8%	0.7%
Value = 5-year yield change	(*t*-stat)	(2.31)	(3.58)	(3.76)	(0.97)	(0.39)	(1.03)	(1.08)
(yield to yield 5 years ago)	Stdev	7.0%	5.9%	5.9%	6.3%	6.4%	4.0%	3.5%
	Sharpe	0.43	0.67	0.71	0.18	0.07	0.19	0.20
	Alpha	−1.3%	0.3%	0.7%	1.9%	1.4%	0.8%	0.7%
	(*t*-stat)	(−1.87)	(−0.51)	(1.03)	(1.68)	(1.21)	(1.10)	(1.15)
					Correlation (Val, Mom) =		− 0.17	−0.35
Value = real bond yield (10-year	Mean	0.9%	2.1%	2.9%	2.0%	1.9%	0.9%	1.4%
yield to 5-year inflation forecast)	(*t*-stat)	(1.84)	(3.36)	(4.17)	(3.78)	(3.58)	(2.44)	(2.63)
	Stdev	2.6%	3.2%	3.6%	2.7%	2.8%	1.9%	2.9%
	Sharpe	0.36	0.65	0.81	0.73	0.70	0.46	0.49
	Alpha	−0.6%	−0.1%	0.4%	1.0%	0.9%	0.4%	0.6%
	(*t*-stat)	(−2.02)	(−0.15)	(1.09)	(2.36)	(2.09)	(1.17)	(1.25)
					Correlation (Val, Mom) =		− 0.09	−0.03
Value = term spread (10-year	Mean	0.8%	1.5%	2.3%	1.5%	1.7%	0.6%	1.1%
yield to short rate)	(*t*-stat)	(1.25)	(2.37)	(3.60)	(2.69)	(3.22)	(1.47)	(1.96)
	Stdev	3.2%	3.2%	3.2%	2.8%	2.5%	2.2%	2.9%
	Sharpe	0.25	0.250.48	0.73	0.55	0.65	0.28	0.37
	Alpha	−1.2%	−0.7%	0.1%	1.3%	1.4%	0.5%	0.8%
	(*t*-stat)	(−3.03)	(−1.79)	(0.28)	(2.78)	(3.31)	(1.22)	(1.55)
					Correlation (Val, Mom) =		0.22	0.28
Value = composite average of	Mean	0.3%	1.6%	3.0%	2.6%	2.9%	1.3%	2.2%
all three measures	(*t*-stat)	(0.58)	(2.56)	(4.63)	(4.88)	(5.89)	(3.17)	(4.30)
	Stdev	3.2%	3.4%	3.4%	2.9%	2.6%	2.1%	2.7%
	Sharpe	0.11	0.48	0.87	0.91	1.10	0.59	0.81
	Alpha	−1.5%	−0.3%	0.8%	2.3%	2.6%	1.0%	1.8%
	(*t*-stat)	(−4.01)	(−0.80)	(2.25)	(4.40)	(5.48)	(2.55)	(3.66)
					Correlation (Val, Mom) =		0.03	0.04

P3–P1 spread and signal-weighted factors (following equations (2) and (3)), and the last row for each asset class reports the correlation of returns between value and momentum for both the P3-P1 zero-cost spread portfolio and the zero-cost signal-weighted factor returns.

Panel A of Table I reports results for each of the individual stock strategies. Consistent with results in the literature, there is a significant return premium for value in every stock market, with the strongest performance in Japan. Momentum premia are also positive in every market, especially in Europe, but are statistically insignificant in Japan. As the last row for each market indicates, the correlation between value and momentum returns is strongly negative, averaging about –0.60. Combining two positive return strategies with such strong negative correlation to each other increases Sharpe ratios significantly. In every market, the value/momentum combination outperforms either value or momentum by itself. Hence, many theories attempting to explain the observed Sharpe ratio for value or momentum have a higher hurdle to meet if considering a simple linear combination of the two.

In addition, the combination of value and momentum is much more stable across markets. For instance, previous research attempting to explain why momentum does not work very well in Japan (see Chui, Titman, and Wei (2010) for a behavioral explanation related to cultural biases) needs to confront the fact that value has performed exceptionally well in Japan during the same time period, as well as the fact that the correlation between value and momentum in Japan is –0.64 over this period. So, rather than explain why momentum did not work in Japan, it would be nearly equally appropriate to ask why value did so well (see Asness (2011)). Moreover, an equal combination of value and momentum in Japan realizes an even higher Sharpe ratio than value alone suggesting that a positive weight on momentum in Japan improves the efficient frontier, which is also confirmed from a static portfolio optimization.

The last set of rows of Table I, Panel A show the power of combining value and momentum portfolios across markets. We report an average of value, momentum, and their combination across all four regions ("Global stocks") by weighting each market by the inverse of their ex post sample standard deviation.[11] Value applied globally generates an annualized Sharpe ratio not much larger than the average of the Sharpe ratios across each market, indicating strong covariation among value strategies across markets. Likewise, momentum applied globally does not produce a Sharpe ratio much larger than the average Sharpe ratio across markets, indicating strong correlation structure among momentum portfolios globally, too.

Panel B of Table I reports the same statistics for the nonstock asset classes. There are consistent value and momentum return premia in these asset classes as well, including some not previously examined (e.g., bonds, value in currencies and commodities).[12] While value and momentum returns vary somewhat across the asset classes, the combination of value and momentum is quite robust due to a consistent negative correlation between value and momentum within each asset class that averages –0.49. We also examine a diversified portfolio of value, momentum, and their combination across all asset classes. Since the volatilities of the portfolios are vastly different across asset classes—for example, commodity strategies have about four times the volatility of bond strategies—we weight each asset class by the inverse of its ex post sample volatility, so that each asset class contributes roughly an equal amount to the ex post volatility of the diversified portfolio.[13] The diversified portfolio across all asset classes yields small improvements in Sharpe ratios, which suggests the presence of correlation structure in value and momentum returns across

these different asset classes. Models that give rise to value and momentum returns in equities, such as the production- or investment-based theories of Berk, Green, and Naik (1999), Johnson (2002), Gomes, Kogan, and Zhang (2003), Zhang (2005), Sagi and Seasholes (2007), Liu, Whited, and Zhang (2009), Li, Livdan, and Zhang (2009), Belo (2010), Li and Zhang (2010), and Liu and Zhang (2008), may not easily apply to other asset classes, yet we find similar value and momentum effects that are correlated to those found in equities, suggesting at least part of these premia are not captured by these models. Likewise, theories of investor behavior, which largely rely on individual investors in equities, will also have difficulty accommodating these facts.

Combining the stock (Panel A) and nonstock (Panel B) value and momentum strategies across all asset classes produces even larger Sharpe ratios. We combine the global stock strategies with the global nonstock other asset class strategies by weighting each by the inverse of their in-sample volatility, where we weight the average stock strategy by its volatility and the average non-stock strategy by its volatility, rather than weighting each individual market or asset class by its own volatility. The 50/50 value and momentum combination portfolio produces an annual Sharpe ratio of 1.45, which presents an even greater challenge for asset pricing models that already struggle to explain the magnitude of the U.S. equity premium, which is about one third as large. Considering value and momentum together and applying them globally across all asset classes, the Sharpe ratio hurdle that these pricing models need to explain is several times larger than those found in U.S. equity data alone.

B. Alternative Measures

We use a single measure for value and a single measure for momentum for all eight markets we study. We choose the most studied or simplest measure in each case and attempt to maintain uniformity across asset classes to minimize the potential for data mining. Using these simple, uniform measures results in positive risk premia for value and momentum in every asset class we study, though some of the results are statistically insignificant. In particular, our weakest results pertain to bonds, which do not produce statistically reliable premia. However, data mining worries may be weighed against the potential improvements from having better measures of value and momentum. For example, value strategies among bonds can be markedly improved with more thoughtful measures. Using our current measure of value, the 5-year change in yields of 10-year maturity bonds, we are only able to produce a Sharpe ratio of 0.18 and an alpha of 1.9% that is not statistically significant (t-statistic of 1.68). However, Panel C of Table I reports results for value strategies among bonds that use alternative measures, such as the real bond yield, which is the yield on 10-year bonds minus the 5-year forecast in inflation, and the term spread, which is the yield on 10-year bonds minus the short rate. As Panel C of Table I shows, these alternative value measures produce Sharpe ratios of 0.73 and 0.55, respectively, and the t-statistics of their alphas are significant at 2.36 and 2.78.

Moreover, we are able to produce even more reliable risk premia when using multiple measures of value simultaneously that diversify away measurement error and noise across the variables.[14] Creating a composite average index of value measures using all three measures above produces even stronger results, where value strategies generate Sharpe ratios

of 0.91 and 1.10 with *t*-statistics on their alphas of 4.40 and 5.48. These alternative measures of value also blend nicely with our original measure for momentum, where, in each case, the 50/50 value/momentum combination portfolios also improve with these alternative measures.

Hence, our use of single, simple, and uniform value and momentum measures may understate the true returns to these strategies in each asset class. Nevertheless, we stick with these simple measures to be conservative and to mitigate data mining concerns, even though, in the case of bonds, the results appear to be insignificant with such simple measures.

C. Comovement across Asset Classes

Table II reports the correlations of value and momentum returns across diverse asset classes to identify their common movements. The strength of comovement may support or challenge various theoretical explanations for value and momentum, and may ultimately point to underlying economic drivers for their returns. The correlations are computed from the returns of the signal-weighted zero-cost factor portfolios from equation (2), but results are similar using the top third minus bottom third P3–P1 portfolio returns.

Panel A of Table II reports the correlations among value strategies and among momentum strategies globally across asset markets. We first compute the average return series for value and momentum across all stock markets and across all nonstock asset classes separately. For example, we compute the volatility-weighted average of all the individual stock value strategies across the four equity markets—the United States, the United Kingdom, Europe, and Japan—and the weighted average of the value strategies across the nonequity asset classes—index futures, currencies, bonds, and commodities. We do the same for momentum. We then compute the correlation matrix between these average return series. The diagonal of the correlation matrix is computed as the average correlation between each individual market's return series and the average of all *other* return series in other markets. For instance, the first entry in the covariance matrix is the average of the correlations between each equity market's value strategy and a portfolio of all other equity market value strategies: an average of the correlation of U.S. value with a diversified value strategy in all other individual equity markets (United Kingdom, Europe, and Japan); the correlation of U.K. value with a diversified value strategy in the United States, Europe, and Japan; the correlation of Europe value with a diversified value strategy in the United States, the United Kingdom, and Japan; and the correlation of Japan value with a diversified value strategy in the United States, United Kingdom, and Europe. We then take an equal weighted average of these four correlations to get the first element of the correlation matrix in Panel A of Table II. In general, we obtain more powerful statistical findings when looking at the correlations of the average return series rather than the average of individual correlations, since the former better diversifies away random noise from each market, a theme we emphasize throughout the paper.[15] Correlations are computed from quarterly returns to help mitigate any nonsynchronous trading issues across markets, due to illiquid assets that do not trade continuously or time zone differences. An *F*-test on the joint significance of the correlations is also performed.

Table II Correlation of Value and Momentum Strategies across Markets and Asset Classes

Reported are the average correlations among all value and momentum strategies across markets and asset classes. Panel A reports the correlations of the average return series, where we first compute the average return series for a group (e.g., all individual stock value strategies across all markets and all nonstock value strategies across all nonstock asset classes) and then compute the correlation between the two average return series. The diagonal elements in Panel A are computed as the average correlation between each market's return series and the average of all other return series in other markets, excluding the market itself. For example, we compute the correlation between U.S. stock value returns and the average returns to value strategies across stocks in the United Kingdom, Europe, and Japan. We then do the same for U.K. value returns with the average of value returns across the United States, Europe, and Japan, and repeat this for Europe value and Japan value strategies as well. We then take the average of these correlations and report them as the first diagonal element of Panel A. Correlations are computed from quarterly returns to mitigate the influence of nonsynchronous trading across markets. Panel B breaks down the correlations of the individual stock value and momentum strategies series with each of the nonstock value and momentum strategies. An F-test for the joint significance of the individual correlations within each group is performed, where * indicates the correlations are significantly different from zero at the 5% significance level.

Panel A: Correlation of Average Return Series

	Stock Value	Nonstock Value	Stock Momentum	Nonstock Momentum
Stock value	0.68*	0.15*	−0.53*	−0.26*
Nonstock value		0.07	−0.16*	−0.13*
Stock momentum			0.65*	0.37*
Nonstock momentum				0.21*

Panel B: Correlation of Average Stock Series with Each Nonstock Series

	Country Index Value	Currency Value	Fixed Income Value	Commodity Value	Country Index Momentum	Currency Momentum	Fixed Income Momentum	Commodity Momentum
Global Stock value	0.27*	0.13*	−0.03	0.01	−0.28*	−0.20*	−0.01	−0.17*
Global Stock momentum	−0.19*	−0.12*	−0.05	−0.06	0.40*	0.28*	0.09	0.20*

Panel A of Table II shows a consistent pattern, where value in one market or asset class is positively correlated with value elsewhere, momentum in one market or asset class is positively correlated with momentum elsewhere, and value and momentum are negatively correlated everywhere across markets and asset classes. The average individual stock value strategy has a correlation of 0.68 with the average value strategy in other stock markets, and of 0.15 with the average nonstock value strategy. The average individual stock momentum strategy has a correlation of 0.65 with the average momentum strategy in other stock markets and a correlation of 0.37 with the average nonstock momentum strategy. The strong correlation structure among value and momentum strategies across such different assets is interesting since these asset classes have different types of investors, institutional and market structures, and information environments.

Value and momentum are also negatively correlated across asset classes. The correlation between a value strategy in one stock market and a portfolio of momentum strategies in other stock markets is –0.53. In addition, value in one asset class is negatively correlated with momentum in another asset class. For example, the correlation between the average stock value strategy and the average nonstock momentum strategy is –0.26, the correlation between nonstock value strategies and stock momentum strategies is –0.16, and the correlation between nonstock value and nonstock momentum in other asset classes is –0.13 on average. This correlation structure—value being positively correlated across assets, momentum being positively correlated across assets, and value and momentum being negatively correlated within and across asset classes—cannot be explained by the correlation of passive exposure to the asset classes themselves. The value and momentum strategies we examine are long–short and market neutral with respect to each asset class, and yet exhibit stronger correlation across asset classes than do passive exposures to these asset classes.

Panel B of Table II breaks down the correlations of the average stock strategies with each of the nonstock strategies. Nearly all of the value strategies across asset classes are consistently positively correlated, all of the momentum strategies are consistently positively correlated, all of the correlations between value and momentum are consistently negatively correlated, and most of these correlations are statistically different from zero.

For robustness, we also show that defining value differently produces similar negative correlation numbers between value and momentum strategies. Our value measure for equities, *BE*/*ME*, uses the most recent market value in the denominator, which yields a –0.53 correlation between value and momentum in Table II, Panel A. However, lagging prices by 1 year in the *BE*/*ME* measure (i.e., using *ME* from 1 year prior) so that the value measure uses price data that do not overlap with the momentum measure, still produces a negative correlation between value and momentum of –0.28, which is highlighted in the Internet Appendix. While these correlations are smaller in magnitude, they are still significantly negative.

In addition, using the negative of the past 5-year return of a stock as a value measure for equities, which is what we use for the nonequity asset classes, also generates negative correlations between value and momentum of similar magnitude (–0.48 as highlighted in the Internet Appendix). This provides more evidence that past 5-year returns capture similar effects as *BE*/*ME* (Gerakos and Linnainmaa (2012) reach a similar conclusion). Hence, simply using recent prices or using past 5-year returns as a value measure does not appear to be driving the negative correlation between value and momentum returns, which appears to be robust across different value measures.

Figure 1 examines the first principal component of the covariance matrix of the value and momentum returns. The top panel of the figure plots the eigenvector weights associated with the largest eigenvalue from the covariance matrix of the individual stock value and momentum strategies in each stock market. The bottom panel of the figure plots the eigenvector weights for all asset classes, which include a global individual stock value and momentum factor across all countries. Both panels show that the first principal component loads in one direction on all value strategies and loads in exactly the opposite direction on all momentum strategies, highlighting the strong and ubiquitous negative correlation between value and momentum across asset classes as well as the positive correlation among value strategies and among momentum strategies across asset classes. The first principal component, which is essentially long momentum and short value (or vice versa) in every asset class, accounts for 54% of the individual stock strategies' covariance matrix and 23% of the all-asset-class covariance matrix. The commonality among value and momentum strategies across vastly different assets and markets with widely varying information, structures, and investors points to common global factor structure among these phenomena.

The Internet Appendix also shows that correlations across markets and asset classes for the value/momentum combination strategies are lower than they are for value or momentum alone, indicating that the negative correlation between value and momentum offsets some of the common variation when combined together in a portfolio. In other words, it appears that value and momentum load oppositely on some common sources of risk.

Figure 2 illustrates succinctly the return and correlation evidence on value and momentum globally by plotting the cumulative returns to value, momentum, and their combination in each asset market and across all asset markets. The consistent positive returns and strong correlation structure across assets, as well as the negative correlation between value and momentum in every market, is highlighted in the graphs.

III. Relation to Macroeconomic and Liquidity Risk

In this section we investigate possible sources driving the common variation of value and momentum strategies across markets and asset classes.

A. Macroeconomic Risk Exposure

Table III reports results from time-series regressions of value and momentum returns for U.S. stocks, global stocks, nonstock asset classes, and all asset classes combined on various measures of macroeconomic risks.[16]

The first two columns of Table III report the time series regression coefficients of U.S. value and momentum returns on U.S. macroeconomic variables: long-run consumption growth, a recession indicator, GDP growth, as well as the U.S. stock market return in excess of the T-bill rate and the Fama and French (1993) bond market factor returns TERM and DEF. Consumption growth is the real per capita growth in nondurable and service consumption obtained quarterly and long-run consumption growth is the future 3-year growth rate in consumption, measured as the sum of log quarterly consumption growth 12 quarters ahead as in Parker and Julliard (2005) and Malloy, Moskowitz, and Vissing-Jorgensen (2009). GDP growth is real per capita growth in GDP. These macroeconomic data are

Figure 1. First principal component for value and momentum strategies. Plotted are the eigenvector values associated with the largest eigenvalue of the covariance matrix of returns to value and momentum strategies. The top graph plots the first principal component of value and momentum strategies in individual stocks in four international markets—the United States, the United Kingdom, Europe (excluding the United Kingdom), and Japan—and the bottom graph plots the first principal component for value and momentum strategies in five asset classes—individual stocks globally, country equity index futures, currencies, sovereign bonds, and commodities. Also reported is the percentage of the covariance matrix explained by the first principal component.

Figure 2. Cumulative returns to value and momentum strategies across markets and asset classes. Plotted are the cumulative (sum of log) returns to value, momentum, and their 50/50 combination strategies in each of the eight asset markets considered: equities in the United States, the United Kingdom, Europe, and Japan; equity index futures; currencies; bonds; and commodities. Returns are plotted for the rank weighted factor portfolios, which are zero-investment portfolios that weight each asset in proportion to its rank based on either value or momentum, following equation (2). Results are also reported for an average of all individual stock strategies across all stock markets ("Global stocks"), across all nonstock asset classes ("Global other asset classes"), and across all markets and asset classes ("Global all asset classes"), where average return series are computed using equal volatility weights across the markets and asset classes to account for differences in volatility across asset classes. All return series are scaled to 10% annual volatility for ease of comparison. Reported on each graph are the annualized Sharpe ratios for each strategy as well as the correlation between value and momentum in each market.

Figure 2. (*Continued*)

Figure 2. (*Continued*)

obtained from the National Income and Product Accounts (NIPA). The recession indicator is defined using ex post peak (=0) and trough dates (=1) from the NBER.

As Table III shows, U.S. stock value strategies are positively related to long-run consumption growth in U.S. data, consistent with the findings of Parker and Julliard (2005), Bansal and Yaron (2004), Malloy, Moskowitz, and Vissing-Jorgensen (2009), and Hansen, Heaton, and Li (2008). U.S. stock momentum strategy returns are not related to long-run consumption growth. Value and momentum are slightly negatively related to recessions and GDP growth, but none of these relationships are statistically significant. TERM and DEF are positively related to value and the default spread is negatively related to momentum.

The next six columns of Table III report regression results for value and momentum in global stocks, all nonstock asset classes, and all asset classes on global macroeconomic variables. Here, we use global long-run consumption growth, which is a GDP-weighted average of 12-quarter-ahead nondurable and service per capita consumption growth in the United States, the United Kingdom, Europe, and Japan. Global macroeconomic data are obtained from Economic Cycle Research Institute (ECRI), which covers production and consumption data as well as business cycle dates using the same methodology as the NBER for approximately 50 countries over time. Similarly, our global recession variable is the GDP-weighted average of recession indicators in each country and global GDP growth is the average across countries weighted by beginning-of-year GDP. For the market return, we use the MSCI World Index in excess of the U.S. T-bill rate. Finally, since we do not have data to construct TERM and DEF internationally, we use the U.S. versions.

As Table III shows, the global macroeconomic variables are generally not significantly related to value and momentum returns, with a couple of exceptions. Momentum is significantly negatively related to recessions, especially among nonstock asset classes. The default spread is positively related to global stock value, but is insignificantly negatively related to value returns in other asset classes. DEF is consistently negatively related to momentum returns in all asset classes.

B. Liquidity Risk Exposure

Table IV reports results from regressions that add various liquidity risk proxies to the macroeconomic variables above.

B.1. Measuring Funding and Market Liquidity Risk

To measure liquidity risk exposure, we regress value and momentum returns on shocks to liquidity. We follow Moskowitz and Pedersen (2012) to define our liquidity shocks. We consider both funding liquidity shocks (e.g., Brunnermeier and Pedersen (2009)) and market liquidity shocks. The funding liquidity variables are the Treasury-Eurodollar (TED) spread (the average over the month of the daily local 3-month interbank LIBOR interest rate minus the local 3-month government rate), the LIBOR minus term repo spread (the spread between the local 3-month LIBOR rate and the local term repurchase rate), and the spread between interest rate swaps and local short-term government rates (Swap-T-bill) in each of the four markets. We sign every variable so that it represents liquidity. Hence, we

Table III Macroeconomic Risk Exposures

Reported are coefficient estimates, t-statistics (in parentheses), and R^2s from time-series regressions of the value and momentum strategy returns in U.S. individual stocks, global individual stocks (across the United States, the United Kingdom, Europe, and Japan), nonstock asset classes, and all asset classes (stock and nonstock) on various measures of macroeconomic risks. The macroeconomic variables are a measure of long-run consumption growth, which is the 3-year future growth rate in per capita nondurable real consumption (quarterly), a recession dummy (0 = peak, 1 = trough) obtained from NBER dates for the United States and ECRI dates outside of the United States, contemporaneous GDP growth rates (from NIPA for the United States and from ECRI outside of the United States), the MSCI world equity index return in excess of the U.S. T-bill rate, and the bond factor returns of Fama and French (1993) TERM and DEF, which represent the term spread on U.S. government bonds and the default spread between U.S. corporate bonds and U.S. Treasuries, respectively. For U.S. stock return regressions, only U.S. macroeconomic variables are used as independent variables. For the global, nonstock, and all-asset-class return regressions, the macroeconomic variables are averaged across all countries, weighting each country in proportion to its GDP. The intercepts from the regressions are not reported for brevity.

		U.S. Stocks			Global Stocks		Nonstock Assets		All Asset Classes	
		Value	Momentum		Value	Momentum	Value	Momentum	Value	Momentum
U.S. values for independent variables	Long-run consumption growth	0.0004	0.0001	Global values for independent variables (GDP-weighted)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
		(2.06)	(0.33)		(0.93)	(0.92)	(0.68)	(0.03)	(1.01)	(0.43)
	Recession dummy	−0.0068	−0.0056		0.0037	−0.0044	0.0045	−0.0081	0.0043	−0.0072
		(−1.06)	(−0.73)		(0.66)	(−0.75)	(1.48)	(−2.44)	(1.55)	(−2.26)
	GDP growth	−0.0050	0.0019		−0.0011	0.0023	−0.0005	−0.0034	−0.0006	−0.0020
		(−1.75)	(0.57)		(−0.39)	(0.80)	(−0.32)	(−2.08)	(−0.45)	(−1.29)
	Market	−0.3435	0.0219		−0.0615	−0.0709	0.0101	−0.0083	−0.0068	−0.0231
		(−7.46)	(0.40)		(−1.41)	(−1.55)	(0.44)	(−0.32)	(−0.32)	(−0.93)
	TERM	0.2038	−0.0234		0.0523	0.0141	−0.0885	0.0370	−0.0551	0.0316
		(2.64)	(−0.25)		(1.04)	(0.27)	(−3.30)	(1.25)	(−2.23)	(1.11)
	DEF	0.7439	−0.7733		0.2650	−0.3752	−0.0510	−0.0787	0.0240	−0.1490
		(5.25)	(−4.57)		(2.86)	(−3.87)	(−1.03)	(−1.44)	(0.53)	(−2.84)
	R-square	13.1%	5.9%		2.3%	6.4%	3.4%	2.9%	2.9%	4.7%

Table IV Liquidity Risk Exposures

Reported are coefficient estimates and *t*-statistics (in parentheses) from time-series regressions of the value and momentum strategy returns across all asset classes on a host of liquidity shocks to measure liquidity risk exposure. The liquidity shocks are estimated as residuals from an AR(2) of a set of funding liquidity variables and market liquidity variables. The funding liquidity variables are the Treasury-Eurodollar (TED) spread, the LIBOR minus term repo spread, and the interest rate swap minus T-bill spread. We also compute a principal component weighted average index of the funding liquidity shocks ("Funding liquidity PC") from the correlation matrix of the liquidity shocks and use this as another regressor. The market liquidity variables are the on-the-run minus off-the-run 10-year government Treasury note spread, the Pástor and Stambaugh (2003) and liquidity measure, and the illiquidity measure of Acharya and Pedersen (2005). All variables are signed so that they represent liquidity, and hence we take the negative of the Acharya and Pedersen (2005) measure. A principal component-weighted average index of the market liquidity shocks from the correlation matrix of the liquidity shocks is also used. Finally, we use a principal component-weighted average index of all liquidity shocks (funding and market) from the correlation matrix of those liquidity shocks as a regressor, where every variable is signed to represent liquidity. Panel A reports results using only U.S. liquidity risk variables and Panel B reports results using global liquidity risk measures, where the global liquidity risks are estimated by taking the average of all the liquidity measures across countries—the United States, the United Kingdom, Europe, and Japan—weighted by the principal component of each country's contribution to the correlation matrix of each liquidity measure across the four markets. TED spreads, LIBOR—term repo rates, swap—T-bill rates, and on-the-run minus off-the-run spreads for each country are quoted using each country's government bond rates. The Pástor and Stambaugh (2003) and Acharya and Pedersen (2005) measures are computed outside of the United States following the same methodology outlined in those papers to individual stocks in each of the other markets—the United Kingdom, Europe, and Japan. All regressions include the set of macroeconomic variables from Table III as controls (coefficient estimates not reported). The intercepts from the regressions are not reported for brevity.

Panel A: U.S. Liquidity Risk Measures

		Value	Momentum	50/50 Combination	Val – Mom
Funding liquidity risk	TED spread	−0.0052	0.0129	0.0061	−0.0180
		(−1.44)	(3.07)	(2.13)	(−2.62)
	LIBOR-term repo	−0.0137	0.0087	−0.0058	−0.0223
		(−2.15)	(1.11)	(−1.26)	(−1.71)
	Swap-T-bill	−0.0002	0.0141	0.0104	−0.0143
		(−0.05)	(3.34)	(3.67)	(−2.04)
	Funding liquidity PC	−0.0111	0.0153	0.0042	−0.0264
		(−2.89)	(3.31)	(1.49)	(−3.41)
Market liquidity risk	On-the-run – off-the-run	0.0063	−0.0053	−0.0043	0.0115
	Pástor-Stambaugh	(0.53)	(−0.38)	(−0.50)	(0.49)
		0.0034	0.0107	0.0159	−0.0074
		(0.32)	(0.89)	(1.93)	(−0.37)
	Acharya-Pedersen	0.0010	0.0005	0.0013	0.0004
		(2.02)	(1.44)	(3.05)	(0.70)
	Market liquidity PC	−0.0080	0.0222	0.0200	−0.0302
		(−0.44)	(0.94)	(1.06)	(−0.97)
All liquidity risk	All PC	−0.0154	0.0195	0.0043	−0.0349
		(−2.84)	(2.96)	(1.09)	(−3.17)

(*Continued*)

Table IV (*Continued*)

Panel B: Global Liquidity Risk Measures

		Value	Momentum	50/50 Combination	Val - Mom
Funding liquidity risk	TED spread	−0.0067	0.0094	0.0023	−0.0161
		(−1.69)	(2.00)	(0.74)	(−2.05)
	LIBOR−term repo	−0.0177	0.0139	−0.0005	−0.0316
		(−2.87)	(1.66)	(−0.08)	(−2.36)
	Swap−T−bill	−0.0076	0.0055	−0.0012	−0.0131
		(−2.15)	(1.31)	(−0.46)	(−1.86)
	Funding liquidity PC	−0.0094	0.0112	0.0013	−0.0206
		(−4.74)	(3.58)	(0.58)	(−4.67)
Market liquidity risk	On-the-run	0.0108	−0.0001	0.0037	0.0109
	—off-the-run	(0.68)	(−0.01)	(0.32)	(0.34)
	Pástor-Stambaugh	0.0010	−0.0002	0.0003	0.0011
		(1.06)	(−0.15)	(0.43)	(0.61)
	Acharya-Pedersen	0.0009	0.0008	0.0020	0.0001
		(0.39)	(0.28)	(1.30)	(0.02)
	Market liquidity PC	−0.0009	0.0016	0.0012	−0.0025
		(−0.74)	(1.21)	(1.00)	(−1.45)
All liquidity risk	All PC	−0.0079	0.0093	0.0016	−0.0172
		(−3.25)	(4.43)	(0.82)	(−4.63)

take the negative of the TED spread and the other spreads so that they capture liquidity, since a wider spread represents worse liquidity.

The funding series are available for the common period January 1987 to July 2011. We define shocks to these variables as the residuals from an AR(2) model, following Korajczyk and Sadka (2008) and Moskowitz and Pedersen (2012).[17] The market liquidity variables are the on-the-run minus off-the-run 10-year government Treasury note spread (see Krishnamurthy (2002)) in each of the four markets (the United States, the United Kingdom, Japan, and Europe, using Germany as a proxy); the Pástor and Stambaugh (2003) liquidity measure (their factor, not their factor mimicking portfolio; specifically, their innovations obtained from CRSP); and the illiquidity measure of Acharya and Pedersen (2005), motivated by Amihud's (2002) measure. We construct the Pástor and Stambaugh (2003) and Acharya and Pedersen (2005) measures in other countries by following their methodologies applied to stocks in those markets. Once again, these variables are signed so that they represent *liquidity*, and hence we take the negative of the Acharya and Pedersen (2005) measure, which is based on Amihud's (2002) illiquidity measure.

In addition, we take the first principal component of the correlation matrix of all funding liquidity shocks, all market liquidity shocks, and all liquidity shocks and construct an index of shocks for funding, market, and all liquidity.[18] The principal component of the correlation, rather than covariance, matrix is used because the liquidity variables have significantly different volatilities and units.

Figure 3 plots the time series of the index of all global liquidity shocks monthly from January 1987 to July 2011. The plot shows that our constructed global liquidity

Figure 3. Time series of global liquidity shocks. The time series of global liquidity shocks is plotted from January 1987 to June 2010, where global liquidity shocks are as defined in Section III. Global liquidity shocks are the residuals from an AR(2) of the global liquidity index, which is a principal component weighted average of all market and funding liquidity variables across all markets (the United States, the United Kingdom, Europe, and Japan) as described in Section III. Also highlighted on the graph are episodes known to have generated movements in aggregate liquidity.

shocks capture a dozen of the largest known liquidity events in global markets over the last 25 years, including the 1987 stock market crash, decimalization, September 11, 2001, the quant meltdown of August 2007, Bear Stearns, and the Lehman Brothers bankruptcy.

B.2. Value and Momentum Returns and Liquidity Risk

Table IV reports regression results of value and momentum returns on the liquidity shocks, controlling for the macro variables in Table III. We only report the coefficient estimates on the liquidity shocks for brevity and because the coefficient estimates on the macro variables do not change much with the addition of the liquidity variables. We examine each liquidity shock in isolation in separate regressions. Panel A of Table IV reports results using the U.S. liquidity shock measures. The dependent variables are the global value and momentum "everywhere" factor returns, the 50/50 combination between them, and the difference between value and momentum returns to test for differences in liquidity exposure between value and momentum. The first four rows of Panel A of Table IV show that funding liquidity risk is consistently negatively related to value returns and significantly

positively related to momentum returns. Value performs poorly when funding liquidity rises, which occurs during times when borrowing is easier, while momentum performs well during these times. The opposite exposure to funding liquidity shocks for value and momentum contributes partly to their negative correlation.[19]

The next four rows examine market liquidity shocks in the U.S. market. Here, we find little relation between market liquidity shocks and value and momentum returns. The Acharya and Pedersen (2005) liquidity measure is marginally negatively related to value and positively related to momentum, but overall the relation between market liquidity shocks and value and momentum returns is weak. Pástor and Stambaugh (2003) (and Sadka (2006)) find a positive and significant relation between U.S. equity momentum returns and their market liquidity shocks. We find the same sign as Pástor and Stambaugh (2003) for our global momentum returns across asset classes over our sample period, but do not detect a significant relation. In addition to our momentum returns covering a wider set of asset classes and a different time period from Pástor and Stambaugh (2003), we also use their factor and *not* their factor mimicking portfolio. They show that the latter exhibits a much stronger relation to momentum, while the former exhibits a weak relation to momentum, consistent with our global results.

Panel B of Table IV reports the regression results using the global funding and market liquidity shocks. Global funding liquidity shocks negatively impact value returns and positively affect momentum returns, but global market liquidity shocks do not seem to have much impact, consistent with the U.S. liquidity measures. Furthermore, the global measures, especially the funding liquidity index, seem to provide more statistical significance. The opposite signed loadings on liquidity risk for value and momentum may partially explain why the two strategies are negatively correlated.

However, the opposite signed loadings on a single factor, such as liquidity risk, cannot explain why *both* value and momentum earn positive risk premia. On the one hand, part of the returns to momentum can be explained as compensation for liquidity risk exposure since momentum loads positively on liquidity shocks and liquidity risk carries a positive risk premium. On the other hand, value loads negatively on liquidity risk, which makes its positive return an even deeper puzzle.

Why does momentum load positively and value load negatively on liquidity risk? One simple and intuitive story might be that momentum captures the most popular trades, being long the assets whose prices have recently appreciated as fickle investors flocked to these assets. Value, on the other hand, expresses a contrarian view, where assets have experienced price declines over several years. When a liquidity shock occurs, investor liquidations (from cash needs, redemptions, risk management, "running for the exit" at the same time; see Pedersen (2009)) puts more price pressure on the more "crowded" trades. These liquidations may affect crowded high momentum securities more than the less popular contrarian/value securities. Further investigation into the opposite signed exposure of value and momentum to liquidity risk is an interesting research question, but beyond the scope of this paper.

Finally, as Table IV shows, because of the opposite signed exposure of value and momentum to funding liquidity shocks, the 50/50 equal combination of value and momentum is essentially immune to funding shocks, and yet, as we have shown, generates huge positive returns. Thus, while exploring liquidity risk's relation to value and momentum more deeply may be interesting, liquidity risk by itself cannot explain why a combination

of value and momentum is so profitable, and hence can only partially explain part of the cross-sectional variation in returns.

B.3. The Power of Averaging Across Markets

A key feature of the analysis in Tables III and IV is that we examine the average returns to value and momentum across a wide set of markets and asset classes simultaneously. The power of looking at the universal average return to value and momentum greatly improves our ability to identify common factor exposure. For example, if we examine each individual value and momentum strategy's exposure to liquidity risk separately, we do not find nearly as strong a pattern and, in fact, might conclude there exists little evidence of any reliable relation to liquidity risk.

Figure 4 depicts the *t*-statistics of the liquidity betas of each of our individual market and asset class value and momentum strategies. The average *t*-statistic of the liquidity betas for value is –0.95 and for momentum is 1.81—hardly convincing. In contrast, when we regress the average value and momentum return series across all markets and asset classes on global liquidity shocks, we get a *t*-statistic for the liquidity beta of –3.25 for value and 4.43 for momentum. The average liquidity beta among the individual strategies is not nearly as strong as the liquidity beta of the average. Averaging across all markets and asset

Figure 4. Liquidity risk beta *t*-statistics. Plotted are the t-statistics of the liquidity risk beta estimates of value and momentum strategies in each asset class using shocks to the global liquidity index as described in Section III. Also reported is the cross-sectional average t-statistic of value and momentum strategies across the asset classes ("average") as well as the t-statistic of the average return series across all asset classes for value and momentum ("all asset classes").

classes mitigates much of the noise not related to value or momentum, such as idiosyncratic regional or asset-specific noise, allowing for better identification of a common factor such as liquidity risk to emerge. When we restrict attention to one asset class at a time, or to one strategy within an asset class, the patterns above are difficult to detect. The scope and uniformity of studying value and momentum everywhere at once is what allows these patterns to be identified.

IV. Comovement and Asset Pricing Tests

The strong common factor structure evidenced in Section II and the link to liquidity risk in Section III suggest that we formally examine asset pricing tests to assess the economic significance of these patterns and how much of the return premia to value and momentum can be captured by this common variation.

A. Explaining Value/Momentum in One Market with Value/Momentum in Other Markets

We first examine how well value and momentum in one market or asset class are explained by value and momentum returns in other asset classes. This test is not a formal asset pricing test, but a test of comovement across markets and asset classes. In the next subsection, we examine formal asset pricing tests. Specifically, we run the regression

$$R_{i,t}^{p} - r_{f,t} = \alpha_i^p + \beta_i^p MKT_t + v_i^p \sum_{j \neq 1} w_j VAL_{j,t} + m_i^p \sum_{j \neq 1} w_j MOM_{j,t} + \varepsilon_{i,t}^p, \tag{4}$$

where $R_{i,t}^{p}$ is the time t return to portfolio p among the six high, middle, and low value and momentum portfolios in one of the eight asset markets i, for a total of 48 test assets. The time series of excess returns (in excess of the U.S. T-bill rate) of each portfolio is regressed on the excess returns of the market portfolio *MKT* (proxied by the MSCI World Index) and the returns to value and momentum factors in *all other* markets and asset classes. The latter two variables are constructed as the equal volatility-weighted average of the zero-cost signal-weighted value and momentum factors in all other markets (where w_j represents the equal volatility weight for each asset class), excluding the market whose test assets are being used as the dependent variable.

We estimate equation (4) for each market and asset class separately. Figure 5 plots the actual average return of each of the test assets against the predicted expected return from the regression. The plot shows how much of the average returns to value and momentum portfolios in one market or asset class can be explained by value and momentum returns from other markets and asset classes. A 45° line passing through the origin is also plotted to highlight both the cross-sectional fit and the magnitude of the pricing errors across test assets.

Figure 5. Explaining value and momentum in one market with value and momentum in other markets. Plotted are the actual average returns (in excess of the U.S. 1-month T-bill rate) of the 48 value and momentum low, middle, and high portfolios in each market and asset class against their predicted expected returns using their betas with respect to value and momentum strategies in all other markets globally. Specifically, for each of the eight markets we consider (U.S. stocks, U.K. stocks, Europe stocks, Japan stocks, country index futures, currencies, government bonds, commodities) we estimate the betas of value and momentum low, middle, and high portfolios in each market with respect to a value and momentum factor across all other markets by running a time-series regression of each value and momentum portfolio in one market on the (equal volatility-weighted) average of the value and momentum factors across all other markets, excluding the market being analyzed. The predicted value from this regression is the predicted expected return of the strategy that we plot against the average actual average return over the sample period. The average absolute value of the alphas from these regressions and the cross-sectional R^2 of the actual average returns against the predicted expected returns are also reported. To highlight both the alphas and cross-sectional fit, a 45° line is plotted through the origin.

As Figure 5 shows, the average returns line up well with the predicted expected returns. The cross-sectional R^2 is 0.55 and the average absolute value of the pricing errors (alpha) is 22.6 basis points per month. A formal statistical test of the joint significance of the pricing errors is not possible since the independent variables change across test assets for each market and asset class (which is why this is not a formal asset pricing test).

The results indicate that value and momentum returns in one market are strongly related to value and momentum returns in other markets and asset classes. Unlike many asset pricing tests conducted in a single market, here there is no overlap of securities between the test assets used as the dependent variable and the factors used as regressors. The dependent variable contains securities from a completely separate market or asset class from those used to construct the factors on the right-hand side of the regression. Hence, the evidence in Figure 5 makes a compelling case for common global factor structure in value and

momentum returns and suggests that this common variation is economically meaningful since it captures a significant fraction of the cross section of average returns.

B. A Global Three-Factor Model

To conduct a more formal asset pricing test, and to compare across various asset pricing models, we construct a three-factor model similar to equation (4), but where the regressors are the same for every asset. This three-factor model is similar in spirit to those of Fama and French (1993) and Carhart (1997), but applied globally to all markets and asset classes we study. Specifically, we estimate the following time-series regression for each of the 48 high, middle, and low value and momentum portfolios across asset classes:

$$R_{i,t}^{p} - r_{f,t} = \alpha_i^p + \beta_i^p \, MKT_t + v_i^p \, VAL_t^{everywhere} + m_i^p \, MOM_t^{everywhere} + \varepsilon_{i,t}^p, \tag{5}$$

where $VAL_t^{everywhere}$ and $MOM_t^{everywhere}$ are the equal-volatility-weighted across-asset-class value and momentum factors.

The first graph in Panel A of Figure 6 plots the actual sample average returns of the 48 test assets versus their predicted expected returns from equation (5) along with a 45° line through the origin to highlight the magnitude of the pricing errors. The cross-sectional R^2 is 0.71 and the average absolute value of the alpha is 18 basis points, indicating slightly better fit than equation (4), which is not surprising, since, unlike equation (4), equation (5) contains some of the same securities on the left- and right-hand side of the regression. However, the fit is similar to equation (4), and equation (5) also allows for a formal joint test of the significance of the alphas, since the explanatory variables are the same for each test asset. Hence, we report the Gibbons, Ross, and Shanken (GRS, 1989) F-statistic and p-value for a joint test of the pricing errors.

The remaining graphs in Panel A of Figure 6 plot the pricing errors of the 48 test assets under alternative asset pricing models: the global CAPM, using the MSCI World Index as the market proxy; a four-factor model inspired by Carhart (1997), which is the Fama–French three-factor model consisting of the U.S. stock market *RMRF*, the U.S. size factor *SMB*, and the U.S. value factor *HML* augmented with the U.S. stock momentum factor *UMD* obtained from Ken French's website, that we refer to as the "Fama–French four-factor model"; and a six-factor model that adds the Fama and French (1993) bond return factors *DEF* and *TERM*, which capture the default and term spread for U.S. bonds, that we refer to as the "Fama–French six-factor model." As Figure 6 shows, the global CAPM does not do a very good job fitting the cross section of value and momentum returns across markets and asset classes, producing the largest absolute pricing errors and smallest R^2. The Fama–French four- and six-factor specifications explain the returns a little better than the CAPM, but not nearly as well as the global three-factor model. The Fama–French factors generate twice the absolute magnitude of pricing errors as the three-factor global model and have much lower R^2s.

Panel B of Figure 6 repeats the same plots for test assets derived only from U.S. stocks. Here, we use the Fama–French 25 size-value and 25 size-momentum portfolios from Ken French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) as test assets. These are, respectively, 5×5 double-sorted portfolios of U.S. stocks

Figure 6. Asset pricing tests of the cross section of expected returns. Plotted are the actual average returns versus model-implied expected returns of the 48 value and momentum low, middle, and high portfolios in each market and asset class under the global CAPM (MSCI World Index), Fama and French four-factor model consisting of U.S. market, size, value, and momentum factors (RMRF, SMB, HML, and UMD), Fama and French six-factor model that adds TERM and DEF to the four-factor model to capture term and default risk premia, and the AMP three-factor model consisting of a global market factor (MSCI World Index), a value everywhere factor, and a momentum everywhere factor, which are value

Figure 6. (*Continued*)
and momentum long-short portfolios diversified across the eight asset classes we consider, where each asset class is weighted by the inverse of its in-sample volatility. A 45° line that passes through the origin is added to highlight the pricing errors (vertical distances to the 45° line), where each model is forced to also price the equity premium. Also reported on each graph are the average absolute value of the alphas, the F-statistic and p-value from the GRS test, and the cross-sectional R^2 under each asset pricing model.

based on size and *BE*/*ME* and 5×5 portfolios sorted on size and past 2- to 12-month returns. Our three-factor model derived from other markets and asset classes does a reasonable job explaining the returns to these 50 U.S. equity portfolios. The cross-sectional R^2 is 0.64 and the average absolute pricing error is only 18 basis points. While the Fama–French factors, which are derived from the same U.S. stocks as the test assets, obviously do slightly better, our three-factor model, which is derived from other asset classes, captures a significant fraction of the cross-sectional variation in U.S. equity returns. In addition, our three-factor model does not contain a size factor, which is important for pricing the Fama-French U.S. stock portfolios. If we exclude the two smallest quintiles of stocks from the Fama-French portfolios, then our three-factor model does as well as the Fama-French model in pricing the remaining Fama-French U.S. portfolios.

Taken together, Panel A of Figure 6 shows that our global three-factor model can explain the returns to value and momentum across markets and asset classes much better than local U.S. factors can and Panel B shows that our global factors can explain the local returns to value and momentum in U.S. stocks almost as well as the U.S. factors can. These results suggest that global value and momentum portfolios across markets and asset classes are closer to the efficient frontier than U.S. stock-only value and momentum portfolios, and therefore provide a more robust set of asset pricing factors that can be used more broadly.

C. Further Pricing Tests and Economic Magnitudes

To further investigate the economic importance of the commonality among value and momentum strategies across asset markets, we examine their relation to macroeconomic and liquidity risks through cross-sectional and time-series asset pricing tests.

C.1. Cross-Sectional Pricing Tests

Table V reports Fama–MacBeth cross-sectional regressions of returns of the 48 value and momentum test portfolios on their betas with respect to funding liquidity risk, GDP growth, long-run consumption growth, TERM, and DEF. Regressions are run in the style of Fama and MacBeth (1973), where the cross section of monthly returns are regressed on the betas (estimated univariately using rolling windows of the past 60 months of returns) each month, and the time-series mean and *t*-statistic of the cross-sectional regression coefficients are reported in Table V. As the first row of Table V shows, liquidity risk betas capture part of the cross-sectional variation in average returns across the 48 portfolios, as indicated by the positive and significant coefficient on the liquidity beta. That coefficient also represents the risk premium for liquidity risk among the 48 test assets, which is 24 basis points per month or about 3% per year. The Fama–MacBeth regressions not only test the cross-sectional relation between average returns and betas with respect to a factor, but the time series of the coefficient estimates represents the return series to a minimum variance portfolio with a unit exposure to that factor (see Fama and MacBeth (1973) and Fama (1976)). Hence, the time series of monthly coefficient estimates represents a factor mimicking portfolio for liquidity risk, which we call $FP_{liq\ risk}$. Likewise, the coefficients on the other variable represent the returns to factor-mimicking portfolios for those factors, each orthogonalized to the other factors, which we will use in time-series asset-pricing tests to follow.

Table V Cross-Sectional Asset Pricing Tests of Global Value and Momentum Strategies

Reported are Fama and MacBeth (1973) regression coefficient estimates and *t*-statistics of the cross section of average returns to the 48 value and momentum portfolios across the eight markets and asset classes we consider. The dependent variable is the cross section of returns on the low, middle, and high value and momentum portfolios of individual stocks in the United States, the United Kingdom, Europe, Japan, country index futures, currencies, government bonds, and commodities. The regressors are beta estimates of these portfolios with respect to the "All" liquidity risk measure from Table IV (the principal component-weighted average index of all liquidity shock measures across all markets globally); GDP growth; long-run consumption growth; the MSCI world index ("market"); the value everywhere factor, consisting of an equal volatility-weighted average of value strategies across all markets and asset classes; and a momentum everywhere factor defined similarly. The last four rows report results using only funding and market liquidity variables to measure liquidity risk, where the principal component-weighted average index of funding and market liquidity shocks are used separately to measure funding liquidity risk and market liquidity risk. Betas are estimated in a univariate regression with respect to each of the factors using a rolling window of the past 60 months of returns. For the market, a Dimson correction is used to account for possible nonsynchronous trading effects, where each portfolio's returns are estimated on the contemporaneous value of the market plus 2 month lags of market realizations and the beta is the sum of the three coefficients on the contemporaneous and 1- and 2-month lags of the market. The cross-sectional regressions are estimated in the style of Fama and MacBeth (1973), where the cross section of returns on the 48 portfolios are regressed each month on the cross-section of beta estimates and the time-series mean and *t*-statistics of the monthly regression coefficients are reported.

Fama–MacBeth Cross-Sectional Regressions							
$\beta_{Liquidity\ risk}$	$\beta_{GDP\ growth}$	β_{LRCG}	β_{TERM}	β_{DEF}	β_{Mkt}	β_{Value}	$\beta_{Momentum}$
Dependent variable: Cross-section of 48 value and momentum portfolios							
0.0024							
(3.05)							
	0.0003	0.0005	0.0021	0.0023			
	(0.43)	(0.42)	(2.19)	(2.18)			
0.0023	−0.0001	0.0012	0.0014	0.0001			
(2.29)	(−0.13)	(1.01)	(1.59)	(0.11)			
0.0005			0.0015	0.0020	0.0029		
(0.56)			(1.75)	(2.22)	(2.58)		
0.0016	0.0018	− 0.0001	0.0033	0.0014	−0.0006	0.0031	0.003
(1.38)	(2.87)	(− 0.03)	(2.87)	(1.12)	(−0.38)	(3.96)	(3.53)
Funding liquidity variables only							
0.0022	−0.0002	0.0019	0.0011	0.0015			
(2.06)	(−0.30)	(1.45)	(1.05)	(1.30)			
0.0012	0.0019	0.0003	0.0027	0.0011	0.0008	0.0034	0.0031
(1.80)	(3.28)	(0.17)	(2.19)	(0.84)	(0.58)	(4.40)	(3.50)
Market liquidity variables only							
0.0001	0.0003	0.0005	0.0021	0.0022			
(0.07)	(0.48)	(0.44)	(2.67)	(2.63)			
−0.0013	0.0019	− 0.0010	0.0045	0.0030	0.0004	0.0038	0.0040
(−0.88)	(3.19)	(−0.77)	(5.05)	(2.73)	(0.31)	(5.94)	(5.82)

The second row of Table V shows that neither GDP growth nor long-run consumption growth captures much cross-sectional variation in returns, but TERM and DEF do, exhibiting

a risk premium of 21 and 23 basis points, respectively. However, the third row of Table V adds liquidity betas to the regression, where we find that the significance of TERM and DEF are subsumed by liquidity risk. Finally, we add betas with respect to the global three-factor model—the MSCI World Index, and the value and momentum everywhere factors. It is perhaps not too surprising that betas with respect to value and momentum factors capture average returns to value and momentum portfolios and that they subsume a significant portion of the explanatory power of other factors such as liquidity risk.

The next two rows of regression results in Table V repeat the regressions using only funding liquidity variables to capture liquidity risk and the last two rows use only market liquidity variables to measure liquidity risk. As Table V shows, only funding liquidity risk appears to be priced in the cross section of our global assets, and exposure to value and momentum common factors seems to capture part of funding liquidity risk exposure.

C.2. Time-Series Pricing Tests

To gain more insight into the economic magnitudes that liquidity risk and the other factors explain, we use the factor mimicking portfolios created from the Fama–MacBeth regressions to conduct time-series asset pricing tests. Specifically, we regress each of the 48 portfolios' time series of monthly returns on the factor mimicking portfolio returns for liquidity risk, GDP growth, and long-run consumption growth, as well as TERM, DEF, and the value and momentum everywhere factors. Because we use factor mimicking portfolios as regressors, both the dependent and independent variables are measured in returns, and hence we can conduct formal pricing tests.

Panel A of Table VI reports the results for the 48 value and momentum portfolios across markets and asset classes. We also include the market portfolio in every regression. For each factor model, we report the GRS F-statistic and p-value for testing the joint significance of the alphas under each model. We also report the average absolute value of the alphas to gauge the magnitude of the pricing errors under each model, the cross-sectional R^2 of the average returns on the test assets against the predicted expected returns from each model, and the *Eig%* metric from Moskowitz (2003), which is the sum of eigenvalues from the covariance matrix of the test assets implied by the model divided by the sum of eigenvalues of the sample covariance matrix. The *Eig%* measure captures how much of the covariance matrix of returns among the test assets each model can explain.

As the first row of Panel A of Table VI shows, the market portfolio alone (global CAPM) generates substantial pricing errors—an average absolute alpha of 35 basis points per month that is easily rejected by the GRS test—and leaves a lot of time-series and cross-sectional variation unexplained. The market portfolio captures about 57% of the covariation among the returns. The second row adds the liquidity risk factor mimicking portfolio as a regressor, and although the GRS test is still rejected, the average absolute alpha declines to 31 basis points, the cross-sectional R^2 increases, and the amount of covariation captured increases. Hence, liquidity risk adds some additional explanatory power for both pricing and common variation of value and momentum portfolios globally across asset classes.

Table VI Time-Series Asset Pricing Tests of Global Value and Momentum Strategies

Panel A reports results from time-series asset pricing tests of the 48 value and momentum portfolios across markets and asset classes on a set of global and U.S.-only asset pricing models. The global models include a global CAPM (MSCI World index); a two-factor model of the global market portfolio plus a factor-mimicking portfolio for a global liquidity risk factor (estimated from the cross-sectional regression coefficients in Table IV); a six-factor model that includes the global market and factor-mimicking portfolios for global liquidity risk, GDP growth, long-run consumption growth (each estimated from the cross-sectional regression coefficients in Table IV), and TERM and DEF; the three-factor AMP model of the global market plus all-asset-class value and momentum factors, as well as two-factor versions that include the market and just one of either all-asset-class value or momentum. The U.S.-only factor models include the U.S. CAPM (U.S. CRSP value-weighted index), Fama and French three-factor model that adds size and value factors to the market, Fama–French four-factor model that adds a momentum factor, and Fama–French six-factor model that adds TERM and DEF as additional factors. Panel B reports results for time-series tests using the same asset pricing models on the 25 Fama–French size and *BE*/*ME* portfolios and 25 size and momentum portfolios that pertain only to U.S. individual stocks. Panel C reports results for 13 hedge fund indices obtained from DJCS and HFRI. All panels report the GRS (1989) *F*-statistic on the joint significance of the alphas under each model from the time-series regressions, the *p*-value of the *F*-statistic, the average absolute alpha, the average time-series R^2, the cross-sectional R^2 of average returns on the predicted expected return from each model, and the percentage of the covariance matrix captured by each model using the *Eig%* metric of Moskowitz (2003) and described in Section IV. Regressions are estimated from monthly returns.

	Asset Pricing Models	GRS *F* Stat	*p*-Value	Average |α|	Average Time-Series R^2	Cross-Sectional R^2	% of Covariances
Panel A: 48 Value and Momentum Portfolios Globally across Asset Classes							
Global Asset Pricing Factors	Mkt (Global CAPM)	6.02	0.000	0.0035	0.40	0.52	57%
	Mkt, $FP_{liq\ risk}$	5.02	0.000	0.0031	0.48	0.54	64%
	Mkt, $FP_{liq\ risk}$, FP_{GDPg}, FP_{LRCG}, TERM, DEF	4.09	0.000	0.0027	0.59	0.56	80%
	Mkt, $VAL_{everywhere}$, $MOM_{everywhere}$	2.66	0.000	0.0018	0.68	0.72	74%
	Mkt, $VAL_{everywhere}$	3.72	0.000	0.0028	0.42	0.43	68%
	Mkt, $MOM_{everywhere}$	3.80	0.000	0.0022	0.42	0.57	70%
U.S. Asset Pricing Factors	Mkt (U.S. CAPM)	6.59	0.000	0.0039	0.30	0.44	47%
	FF 3-Factor	7.18	0.000	0.0036	0.31	0.50	53%
	FF 4-Factor	6.70	0.000	0.0035	0.33	0.55	63%
	FF 6-factor	7.11	0.000	0.0035	0.39	0.62	64%

(*Continued*)

Table VI (*Continued*)

Panel B: Fama–French 25 Size-Value and 25 Size-Momentum Portfolios							
Global Asset Pricing Factors	Mkt (Global CAPM)	4.09	0.000	0.0030	0.41	0.20	48%
	Mkt, $FP_{liq\ risk}$	4.12	0.000	0.0030	0.42	0.20	49%
	Mkt, $FP_{liq\ risk}$, FP_{GDPg}, F_{LRCG}, TERM, DEF	4.76	0.000	0.0035	0.57	0.38	61%
	Mkt, $VAL_{everywhere}$, $MOM_{everywhere}$	3.22	0.000	0.0019	0.70	0.68	66%
U.S. Asset Pricing Factors	Mkt (U.S. CAPM)	4.25	0.000	0.0032	0.73	0.17	82%
	FF 3-Factor	3.81	0.000	0.0023	0.87	0.30	93%
	FF 4-Factor	3.28	0.000	0.0011	0.91	0.77	97%
	FF 6-factor	3.35	0.000	0.0011	0.91	0.77	97%
Panel C: 13 Hedge Fund Indices							
Global Asset Pricing Factors	Mkt (Global CAPM)	12.14	0.000	0.0032	0.30	0.20	43%
	Mkt, $FP_{liq\ risk}$	12.36	0.000	0.0025	0.34	0.30	47%
	Mkt, $FP_{liq\ risk}$, FP_{GDPg}, F_{LRCG}, TERM, DEF	11.86	0.000	0.0022	0.46	0.17	53%
	Mkt, $VAL_{everywhere}$, $MOM_{everywhere}$	7.26	0.000	0.0018	0.41	0.47	54%
U.S. Asset Pricing Factors	Mkt (U.S. CAPM)	12.14	0.000	0.0028	0.30	0.18	43%
	FF 3-Factor	12.64	0.000	0.0026	0.35	0.19	47%
	FF 4-Factor	13.03	0.000	0.0022	0.37	0.36	49%
	FF 6-factor	12.25	0.000	0.0021	0.44	0.36	51%

While there is a link between value and momentum and liquidity risk, only a small fraction of the return premia and covariation is captured by our proxies for these risks. We view these findings as an important starting point for possible theories related to value and momentum phenomena, but emphasize that we are far from a full explanation of these effects. We also recognize that measurement error in liquidity risk may limit what we can explain. In addition, a single liquidity risk factor alone cannot explain value and momentum since they are negatively correlated with each other but both produce positive returns, unless there is substantial time variation in liquidity risk betas and in the liquidity risk premium. Thus, it is not surprising that the pricing errors from this model specification remain large.

The third row of Table VI, Panel A adds factor mimicking portfolio returns for GDP growth, long-run consumption growth, and TERM and DEF. Pricing errors decline further while R^2s and the amount of covariation explained increase. The fourth row uses our three-factor model, which provides the best fit. Here, the average absolute alpha is only 18 basis points, the cross-sectional R^2 is 72%, and 84% of the covariation among the test assets is captured by these factors. The next two rows further show that having both value and momentum in the model is important, since having only value or momentum by itself increases pricing errors and decreases the fit considerably. This further underscores the difficulty of using a single factor to explain both value and momentum.

The last four rows of Table VI, Panel A examine models of U.S. stock factors: the U.S. market portfolio in excess of the U.S. T-bill rate, the Fama–French three-factor model, the Fama–French four-factor model that adds the momentum factor, and the Fama–French six-factor model that also adds TERM and DEF. As Table VI shows, the U.S. factors do not do a great job of describing the global value and momentum portfolio returns, leaving larger pricing errors and lower R^2s, and capturing a smaller fraction of their covariance matrix.

Panel B of Table VI repeats the same exercise as Panel A, but uses the 25-size-*BE/ME* and 25 size-momentum U.S. equity portfolios from Ken French's website as test assets. Not surprisingly, the Fama–French U.S. factors do a good job of capturing these returns, though the GRS test is still rejected. However, the global value and momentum everywhere factors, which consist primarily of non-U.S. equities and other asset classes, also do a good job explaining the 50 U.S. equity-based test assets—the average absolute alpha is only 19 basis points, the cross-sectional R^2 is 68%, and the percentage of covariation captured is 66%. This is better than the Fama–French three factor model does and only slightly worse than the Fama–French four- or six-factor models, which are specifically designed to capture these portfolios and are constructed from the same set of securities as the test assets themselves.

Finally, Panel C of Table VI considers how well these factor models can explain hedge fund returns. Using the returns of 13 hedge fund indices from Dow Jones Credit Suisse (DJCS) and Hedge Fund Research Institute (HFRI) that include from DJCS the Market Neutral, Long-Short, Multi Strategy, Macro, Managed Futures, Currency, Emerging Markets, and Overall hedge fund indices and from HFRI the Equity Hedge, Fund of Funds, Macro, Emerging Markets, and Overall hedge fund indices, Panel C of Table VI shows that the global three-factor model has smaller pricing errors than the Fama–French model and its extensions with the momentum, TERM, and DEF factors. These results are consistent with Boyson, Stahel, and Stulz (2010), Sadka (2012), and Bali, Brown, and Caglayan (2011, 2012), who find that the Fama and French U.S. stock factors do not explain the cross

section of hedge fund returns very well. However, our simple value and momentum factors applied globally across asset classes do appear to capture a sizable fraction of the returns to hedge funds.

The evidence in Table VI suggests that the global across-asset three-factor model does a good job of capturing not only the returns to value and momentum globally across asset classes, but also the returns to size and value and size and momentum in U.S. equities, as well as the cross section of hedge fund returns, providing additional testing grounds that are created from a completely different set of securities. Conversely, while local U.S. factors capture U.S. equity returns well, they do not explain a lot of value and momentum returns globally or across asset classes, nor do they capture the returns to various hedge fund strategies well. While our three-factor global model performs better in explaining all of these different test assets, the GRS test still rejects our model in all cases, suggesting that more work needs to be done to fully describe the cross section of returns.

V. Robustness and Implementation

Finally, we examine the robustness of our findings to implementation issues. A reader convinced of the efficacy of value and momentum strategies, particularly in combination, may be concerned with real world implementation issues. Though well beyond the scope of this paper, in this section we briefly discuss some practical concerns, including implementation costs and portfolio construction, as well as opportunities to improve upon our admittedly but intentionally simple approach.

A. Transaction Costs

Like most academic studies, we focus on gross returns, which are most suitable to illuminating the relation between risk and returns. However, gross returns overstate the profits earned by pursuing the strategies we examine in practice. A few papers try to examine the transaction costs and capacity of these strategies, especially momentum, perhaps due to its higher turnover. For example, Korajczyk and Sadka (2004) and Lesmond, Schill, and Zhou (2003) argue that the real world returns and capacity of equity momentum strategies are considerably lower than the theoretical results would imply. Their conclusions are based on aggregate trade data and theoretical models of transactions costs. Using live trading data, Frazzini, Israel, and Moskowitz (2012) challenge these results and show that the real world trading costs of value, momentum, and a combination of the two in equities are orders of magnitude lower for a large institution than those implied by the calibrated models of Korajczyk and Sadka (2004) and Lesmond, Schill, and Zhou (2003). As a result, Frazzini, Israel, and Moskowitz (2012) conclude that these strategies can be scaled considerably and still generate strong net returns. In addition, we focus on an extremely large and liquid set of equities in each market (approximately the largest 17% of firms), where trading costs, price impact, and capacity constraints are minimized.

Studies on trading costs also focus exclusively on individual stocks, but half of the markets that we examine are implemented with futures contracts, which typically have much lower trading costs than stocks. Hence, although our equity strategies outperform our nonequity strategies in gross returns, net of trading cost returns are likely to be much closer.

Furthermore, Garleanu and Pedersen (2012) model how portfolios can be optimally rebalanced to mitigate transaction costs and demonstrate how this improves the net performance of commodity momentum strategies, for example. In a similar spirit, Frazzini, Israel, and Moskowitz (2012) demonstrate how equity portfolios can benefit from several practical steps taken to reduce transactions costs that, while having a cost in terms of gross returns (from style drift), can improve net returns. For instance, the strategies we study here are all naively rebalanced exactly monthly no matter what the expected gain per amount traded. Varying the rebalance frequency, optimizing the portfolios for expected trading costs, and extending or occasionally contracting the trade horizon can all improve the basic implementation of these strategies.

B. Shorting

Our paper is, of course, as much about shorting assets as it is about going long. While going long versus short is symmetric for futures, shorting involves special costs in stock markets. If our results are completely dependent on shorting and if shorting is too costly or not implementable, this would certainly raise questions about the real world efficacy of these strategies. However, Israel and Moskowitz (2012) provide evidence that the return contributions of both value and momentum strategies across the same asset classes we study here are roughly equal from the long and short sides of the portfolio and that long-only portfolios of value and momentum still produce abnormal returns. Thus, these strategies are still effective even if shorting is restricted. In addition, Frazzini, Israel, and Moskowitz (2012) provide some evidence that the trading costs of shorting stocks are not materially different from the costs of buying or selling stocks, and that real-world shorting costs for a large institutional investor are not prohibitive to running sizable funds in these strategies.

C. Portfolio Formation

In this paper we intentionally keep everything as simple as possible, both for clarity and as a precaution against the pernicious effects of data mining. In fact, one of the paper's objectives is to provide a robust out-of-sample test of ideas that have been largely tested in individual, particularly U.S., stocks and extend them to other asset classes. However, when faced with real world implementation, there are many choices to consider. For example, we look at two simple portfolio implementations in the paper: top 1/3 minus bottom 1/3, and a linear weighting scheme based on ranking securities. These are far from the only possibilities, and the choice of weighting scheme can impact not only gross returns, but also transactions costs. While we do not claim that either of these choices is optimal in either a gross or net return sense, we also explore more extreme sorts of securities into deciles and find that doing so does not materially affect the results. In the Internet Appendix we replicate our main results for individual equity markets in the United States, the United Kingdom, Europe, and Japan for decile portfolios and find very similar results. In the Internet Appendix we also plot the pricing errors of our three-factor model for these 80 decile portfolios of value and momentum in each of the four equity markets (the United States, the United Kingdom, Europe, and Japan). As the accompanying figure shows and the asset

pricing statistics verify, our three-factor model does a good job of capturing these more extreme portfolio returns, too.

We also value weight stocks within our portfolios and equal weight the securities in other asset classes. However, other weighting schemes yield similar results and, because we focus on the largest, most liquid securities, trading costs are unlikely to be affected much by such changes. Hence, our main findings are robust to a variety of perturbations and portfolio formations.

D. Volatility Scaling

When we aggregate our strategies across asset classes, we ensure that the different asset classes are scaled to have similar volatility. To do so, we scale each asset class by the inverse of its realized volatility over the full sample. However, since the full sample is not known in advance, a real world portfolio would need to scale by a measure of volatility that is estimated ex ante. For robustness, in the Internet Appendix we report results for all of our value and momentum strategies scaled to the same ex ante volatility of 2% per annum using a rolling 3-year estimate of each portfolio's volatility from daily returns. The results are reported together with the original unadjusted returns. The Sharpe ratios and correlations of the strategies are very similar and yield identical conclusions.

E. Dollar Neutral vs. Beta Neutral

As is standard in academic studies, our strategies are constructed to be $1 long and $1 short, but they need not have a zero market beta exposure (at the local or global level). However, real world portfolios can, and often do, attempt to create long-short portfolios that are ex ante beta neutral (in addition to, or instead of, being dollar neutral). We find that our inferences based on the strategies' alpha from factor regressions are not affected by market hedging.

F. Value and Momentum Measures

We use one measure for value and one for momentum for all eight markets we study. We choose the most studied or simplest measure in each case and attempt to maintain uniformity across asset classes to minimize the potential for data mining. In real world implementations, data mining worries may be weighed against the potential improvements from having multiple (and perhaps better) measures of value and momentum, if for no other reason than to diversify away measurement error or noise across variables. Israel and Moskowitz (2012) show, for instance, how other measures of value and momentum can improve the stability of returns to these styles in equities. Most practical implementations use a variety of measures for a given style. In fact, we set out to examine value and momentum in eight different markets and asset classes using a single uniform measure for each. Although we find positive returns to value and momentum in each asset class, these returns are not always significant. In particular, our weakest results using the current measures of value and momentum pertain to bonds, which do not produce statistically significant

premia. However, as shown in Table I, Panel C, the returns can be vastly improved using other measures of value and momentum, and taking a composite average index of measures for value and momentum produces even more stable and reliable results. Hence, our use of simple, uniform value and momentum measures may understate the true returns to these strategies.

The literature on realistic implementation of these strategies is still young, and the list of choices to make when moving from an academic study like ours to implementing these strategies in practice is long. But current evidence, research, and practical experience point to the effects we study being highly applicable to real world portfolios. Consistent with this conjecture, as shown in Table VI, our simple value and momentum global factors capture a sizable fraction of the returns to hedge fund indices, which suggests that hedge funds are engaged in similar or highly correlated strategies globally.

G. Evolution over Time

As the hedge fund industry has grown and more capital has been devoted to these strategies, it is interesting to consider what effect, if any, such activity has on the efficacy of value and momentum strategies. While a complete analysis of this question is beyond the scope of this paper, we offer a couple of results perhaps worthy of future investigation.

Table VII reports the Sharpe ratios and correlations among the value and momentum strategies over the first and second halves of the sample period—1972 to 1991 and 1992 to 2011. As the first row of Table VII, Panel A shows, the Sharpe ratios to both value and momentum have declined slightly over time. In addition, their correlations across markets have increased over time—the average correlation among value strategies has risen from 0.31 to 0.71 and among momentum strategies has risen from 0.46 to 0.77. However, the correlation between value and momentum has declined from –0.44 to –0.63, and, as a result, the Sharpe ratio of the combination of value and momentum has not changed much over time, since the increased correlation across markets is being offset by the more negative correlation between value and momentum. These results may be consistent with increased participation of arbitrageurs driving up correlations among value and momentum strategies globally.

The next row of Table VII, Panel A repeats the same analysis, splitting the sample prior to and after August 1998, which is roughly when the funding crisis peaked following the collapse of Long Term Capital Management (LTCM). The correlation among value strategies is much higher after August 1998 (0.16 pre-1998 vs. 0.64 post-1998), and the correlation among momentum strategies is also higher after 1998 (0.43 vs. 0.71). The next three rows of Table VII, Panel A report the same statistics for periods of worsening and improving funding liquidity, defined as negative and positive funding liquidity shocks, and are split separately into pre- and post-1998. Consistent with our previous regression results in Table IV, value strategies do worse when liquidity improves and momentum strategies do worse when liquidity declines, but these patterns appear only after 1998. Prior to the financial crisis of 1998, funding liquidity shocks seem to have little impact on value or momentum strategies. After 1998, however, value generates a Sharpe ratio of 0.85 during periods of worsening liquidity, but only 0.28 when liquidity improves. Conversely, momentum produces a Sharpe ratio of 0.19 when liquidity worsens, but a Sharpe ratio of 0.99

when liquidity improves. The 50/50 value/momentum combination is immune to liquidity risk, even after 1998.

Panel B of Table VII examines more formally how value and momentum correlations change over time and with liquidity shocks by running time-series regressions in which the dependent variable is the cross product of monthly returns on the various strategies to proxy for time-varying correlations. We estimate the time t correlation among value strategies globally, $\rho(Val, Val)_t$, as the average across asset-classes at time t of $r_{i,t}^{val} \times r_{k,t}^{val}$, where $r_{i,t}^{val}$ is the return to the value strategy in market or asset class i at time t. We define $\rho(Mom, Mom)_t$, and $\rho(Val, Mom)_t$ similarly. The time series of these correlations is regressed on a linear time trend, a global recession indicator, and the time series of liquidity shocks. The first three columns of Table VII, Panel B show that the average correlation among value and momentum strategies across markets and asset classes has been significantly increasing over time and the correlation between value and momentum is significantly more negative over time. Recessions increase the correlation among both value and momentum strategies globally, controlling for the time trend. Liquidity shocks also appear to significantly increase correlations among momentum strategies, controlling for the time trend and recessions. However, the last three columns repeat the regressions adding a post-1998 dummy variable and an interaction between the post-1998 dummy and liquidity shocks. Rather than a time trend, the post-1998 dummy seems to be driving any correlation changes, and the impact of liquidity shocks on correlations also appears to be exclusively a post-1998 phenomenon. These results are consistent with an increase in the importance of liquidity risk on the efficacy of these strategies following the events of August 1998 that appear to be more important than any time trend on the increasing popularity of value and momentum strategies among leveraged arbitrageurs.[20] Hence, funding liquidity risk and limits to arbitrage activity may be a progressively more crucial feature of these strategies and future work may consider these issues in understanding the returns to value and momentum.

VI. Conclusion

We provide comprehensive evidence on the return premia to value and momentum strategies globally across asset classes, and uncover strong common factor structure among their returns. The strong correlation structure among value and momentum strategies across such diverse asset classes is difficult to reconcile under existing behavioral theories, while the high return premium and Sharpe ratio of a global across-asset-class diversified value and momentum portfolio presents an even more daunting hurdle for rational risk-based models to accommodate than the more traditional approach of considering value or momentum separately in a single asset market. Although both behavioral and rational theories for value and momentum focus predominantly on equities, the existence of correlated value and momentum effects in other asset classes—with their different investors, institutional structures, and information environments—argues for a more general framework.

We further find that exposure to funding liquidity risk provides a partial explanation for this correlation structure, especially following the funding crisis of 1998, but leaves much to be explained. While the relation to funding liquidity risk could imply that limited arbitrage activity may contribute to the prevalence and dynamics of these phenomena, we leave the ubiquitous evidence on the efficacy of value and momentum across the diverse asset classes we study, its strong correlation structure, and intriguing dynamics related to funding risk as a challenge for future theory and empirical work to address.

Table VII Dynamics of Value and Momentum Returns

Panel A reports Sharpe ratios and correlations among the value, momentum, and 50/50 value/momentum combination strategies across different economic environments. The first three columns report the Sharpe ratios of the all-asset-class value, momentum, and 50/50 combination strategies and the last three columns report the average correlations among value strategies globally; among momentum strategies globally, and among 50/50 value/momentum combinations globally. These statistics are reported for the two halves of the sample period, prior to and after August 1998 for the top and bottom half of observations based on our global index of liquidity shocks from Table III ("improving" and "worsening" liquidity, respectively), and the same split of improving versus worsening liquidity pre- and post-August 1998. Panel B reports time-series regressions of conditional correlations among value strategies globally; among momentum strategies globally; and between value and momentum strategies globally on a time trend, a global recession indicator (as defined in Table III), global liquidity shocks (as defined in Table III), a post-August 1998 dummy, and an interaction between the post-August 1998 dummy and liquidity shocks. The conditional correlations used as the dependent variables are estimated as the average pairwise correlations among the strategies each month using the cross product of monthly returns to each strategy as described in Section V.G.

Panel A: Dynamics of Sharpe Ratios and Correlations

	Sharpe Ratios			Correlations		
	Value	Momentum	50/50 Combination	*ρ(Val, Val)*	*ρ(Mom, Mom)*	*ρ(Val, Mom)*
1st half—1972 to 1991	0.78	0.90	1.40	0.31	0.46	−0.44
2nd half—1992 to 2010	0.68	0.71	1.43	0.71	0.77	−0.63
Pre-08/1998	0.68	1.02	1.49	0.16	0.43	−0.51
Post-08/1998	0.75	0.72	1.39	0.64	0.71	−0.55
Worsening liquidity	0.95	0.57	1.36	0.54	0.72	−0.53
Improving liquidity	0.59	0.87	1.45	0.77	0.79	−0.56
Worsening liquidity (pre-1998)	1.10	1.00	1.76	0.40	0.59	−0.30
Improving liquidity (pre-1998)	1.09	1.27	2.04	0.36	0.49	−0.29
Worsening liquidity (post-1998)	0.85	0.19	0.88	0.65	0.81	−0.71
Improving liquidity (post-1998)	0.28	0.77	1.07	0.87	0.87	−0.65

Panel B: Dynamics of Value and Momentum Correlations

Dependent variable	$\rho(Val, Val)_t$	$\rho(Mom, Mom)_t$	$\rho(Val, Mom)_t$	$\rho(Val, Val)_t$	$\rho(Mom, Mom)_t$	$\rho(Val, Mom_t$
Time trend	0.0067	0.0181	−0.0320	−0.0011	0.0045	−0.0197
	(2.21)	(3.26)	(−4.22)	(0.20)	(0.52)	(−1.37)
Recession	0.0828	0.0971	0.0195	0.0823	0.0938	0.0206
	(1.88)	(2.31)	(0.31)	(2.05)	(2.30)	(0.34)
Liquidity shocks	0.0131	0.0519	−0.0303	0.0458	−0.0048	−0.0717
	(0.98)	(2.58)	(−1.62)	(1.80)	(−0.11)	(−1.64)
Post-08/1998				0.1212	0.2136	−0.1928
				(1.70)	(1.82)	(−0.99)
Liquidity shocks × post-08/1998				−0.0379	0.0929	0.0161
				(−1.20)	(2.11)	(0.34)

Finally, we provide a simple global three-factor model that describes a new set of 48 global across-asset-class test assets, the Fama–French portfolios, and a variety of hedge fund indices. In further investigating the underlying economic sources driving value and

momentum returns, we hope this simple three-factor framework can be useful for future research that is becoming increasingly concerned with pricing global assets across markets.

Endnotes

[1]Early evidence on U.S. equities finds that value stocks on average outperform growth stocks (Stattman (1980), Rosenberg, Reid, and Lanstein (1985), and Fama and French (1992)) and stocks with high positive momentum (high 6- to 12-month past returns) outperform stocks with low momentum (Jegadeesh and Titman (1993), Asness (1994)). Similar effects are found in other equity markets (Fama and French (1998), Rouwenhorst (1998), Liew and Vassalou (2000), Griffin, Ji, and Martin (2003), Chui, Wei, and Titman (2010)), and in country equity indices (Asness, Liew, and Stevens (1997) and Bhojraj and Swaminathan (2006)). Momentum is also found in currencies (Shleifer and Summers (1990), Kho (1996), LeBaron (1999)) and commodities (Erb and Harvey (2006), Gorton, Hayashi, and Rouwenhorst (2008)).

[2]A single factor would require significant time variation in betas and/or risk premia to accommodate these facts. We remain agnostic as to whether our factors capture such dynamics or represent separate unconditional factors.

[3]See Gomes, Kogan, and Zhang (2003), Zhang (2005), Li, Livdan, and Zhang (2009), Belo (2010), Li and Zhang (2010), Liu and Zhang (2008), Berk, Green, and Naik (1999), Johnson (2002), Sagi and Seasholes (2007), and Liu, Whited, and Zhang (2009).

[4]This procedure is similar to how MSCI defines its universe of stocks for its global stock indices.

[5]Hong, Lim, and Stein (2000), Grinblatt and Moskowitz (2004), Fama and French (2012), and Israel and Moskowitz (2012) show that value and momentum returns are inversely related to the size of securities over the time period studied here, though Israel and Moskowitz (2012) show this relation is not robust for momentum in other sample periods. Value and momentum returns have also been shown to be stronger in less liquid emerging markets (Rouwenhorst (1998), Erb and Harvey (2006), Griffin, Ji, and Martin (2003)). A previous version of this paper used a broader and less liquid set of stocks that exhibited significantly stronger value and momentum returns.

[6]Austria, Belgium, Denmark, Norway, and Portugal are not index futures but are constructed from the returns of an equity index swap instrument using the respective local market index from MSCI.

[7]While research has shown that other value measures are more powerful for predicting stock returns (e.g., Lakonishok, Shleifer, and Vishny (1994), Asness, Porter, and Stevens (2000), Piotroski (2000)), we maintain a basic and simple approach that is somewhat consistent across asset classes.

[8]Novy-Marx (2012) shows that the past 7- to 12-month return is a better momentum predictor in U.S. stocks than the past 2- to 6-month return, though the past 2- to 6-month return is still a positive predictor. We use the more standard momentum measure based on the past 2- to 12-month return for several reasons. First, as Novy-Marx (2012) shows, the benefit of using returns from the past 7 to 12 months as opposed to the entire 2- to 12-month past return is negligible in U.S. stocks. Second, Goyal and Wahal (2012) examine the power of past 7- to 12-month versus past 2- to 6-month returns across 36 countries and find that there is no significant difference between these past return predictors in 35 out of 36 countries—the exception being the United States. Third, *MOM2–12* is the established momentum signal that has worked well out of sample over time and across geography. While we believe using *MOM2–12* is the most prudent and reasonable measure to use for these reasons, using other momentum signals, such as *MOM7–12*, should not alter any of our conclusions.

[9]An Internet Appendix may be found in the online version of this article.

[10]Weighting the nonstock asset classes by their ex ante volatility gives similar results. In addition, rebalancing back to equal weights annually rather than monthly produces similar results.

[11]We compute the monthly standard deviation of returns for each passive benchmark in each market and weight each market by the inverse of this number, rescaled to sum to one, to form a global portfolio across all markets. Each market's dollar contribution to the global portfolio is therefore proportional to the reciprocal of its measured volatility, but each market contributes an equal fraction to the total volatility of the portfolio, ignoring correlations. We weight every portfolio (low, middle, high, and value and momentum) and factor within each market by the same number based on the volatility of the total market index for that market. For the nonstock asset classes we do the same, where the benchmark portfolio is simply an equal weighted average of all the securities in that asset class. Weighting by total market cap or equal weighting produces nearly identical results, but we use the equal volatility weighting scheme to be consistent with our procedure for the nonequity asset classes, where market cap has no meaning and where volatility differs greatly across different asset clasees.

[12]The somewhat weaker returns for the nonstock asset classes would be partially attenuated if transactions costs were considered, since trading costs are typically higher for individual stocks than the futures contracts we examine outside of equities. Therefore, net-of-trading-cost returns would elevate the relative importance of the nonstock strategies. We discuss implementation issues briefly in Section V.

[13]Using ex ante rolling measures of volatility and covariances yields similar results.

[14]Israel and Moskowitz (2012) show how other measures of value and momentum can improve the stability of returns to these styles among individual equities.

[15]In the Internet Appendix to the paper, we report the average of the individual correlations among the stock and nonstock value and momentum strategies, where we first compute the pairwise correlations of all individual strategies (e.g., U.S. value with Japan value) and then take the average for each group. We exclude the correlation of each strategy with itself (removing the 1s) when averaging and also exclude the correlation of each strategy with all other strategies within the same market (i.e., exclude U.S. momentum when examining U.S. value's correlation with other momentum strategies). While these individual correlations are consistently weaker than those obtained from taking averages first and then computing correlations, the average pairwise correlations also exhibit strong comovement among value and momentum across markets.

[16]Chordia and Shivakumar (2002) claim that a conditional forecasting model of macroeconomic risks can explain momentum profits in U.S. stocks, but Griffin, Ji, and Martin (2003) show that neither an unconditional or conditional model of macroeconomic risks can explain momentum in equities globally across 40 countries, including the United States. We examine the relation between macroeconomic risks and value and momentum strategies globally across asset classes to potentially shed new light on this question.

[17]There is no special or theoretical reason to use an AR(2). An AR(3), AR(1), and first differences model yield similar results.

[18]A previous version of this paper also included the liquidity measures of Sadka (2006) and Adrian and Shin (2010) and found similar results. However, because the Sadka (2006) and Adrian and Shin (2009) measures require data not available in other equity markets, such as tick and trade data and balance sheet information from prime brokers, we cannot compute them internationally and hence omit them. See Amihud, Mendelson, and Pedersen (1994) for a survey of liquidity and liquidity risk measures.

[19]Another interpretation of these funding shocks is that they proxy for changes in risk aversion or risk premia. So, in addition to funding liquidity being tight when spreads are wide, it may also be the case that risk aversion or risk premia in the economy are particularly high. Under this alternative

view, however, it would seem that *both* value and momentum returns would decline with rising spreads, whereas we find that value and momentum returns move in opposite directions with respect to these shocks. In addition, the market portfolio and macroeconomic variables are included in the regression, which may partly capture changing risk or risk premia.

[20]Israel and Moskowitz (2012) examine the relation between size, value, and momentum profitability and aggregate trading costs and institutional investment over time. They find little evidence that the returns to these strategies vary with either of these variables.

References

Acharya, Viral, and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Adrian, Tobias, and Hyun Song Shin, 2010, Liquidity and leverage, *Journal of Financial Intermediation*, 19, 418–437.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Yakov, Haim Mendelson, and Lasse Heje Pedersen, 2005, Liquidity and asset prices, *Foundations and Trends in Finance* 1, 269–364.

Asness, Clifford S., 1994, *Variables that explain stock returns*, Ph.D. Dissertation, University of Chicago.

Asness, Clifford S., 2011, Momentum in Japan: The exception that proves the rule, *Journal of Portfolio Management* 37, 67–75.

Asness, Clifford S., and Andrea Frazzini, 2012, *The devil in HML's details*, Working paper, AQR Capital Management.

Asness, Clifford S., John M. Liew, and Ross L. Stevens, 1997, Parallels between the cross-sectional predictability of stock and country returns, *Journal of Portfolio Management* 23, 79–87.

Asness, Clifford S., R. Burt Porter, and Ross Stevens, 2000, *Predicting stock returns using industry-relative firm characteristics*, Working paper, AQR Capital Management.

Bali, Tarun G., Stephen J. Brown, and Mustafa O. Caglayan, 2011, Do hedge funds' exposures to risk factors predict their future returns? *Journal of Financial Economics* 101, 36–68.

Bali, Tarun G., Stephen J. Brown, and Mustafa O. Caglayan, 2012, Systematic risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 106, 114–131.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 35, 54–56.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Belo, Federico, 2010, Production-based measures of risk for asset pricing, *Journal of Monetary Economics* 57, 146–163.

Berk, Jonathan, Richard Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1153–1607.

Bessembinder, Henrik, 1992, Systematic risk, hedging pressure, and risk premiums in futures markets, *Review of Financial Studies* 5, 637–667.

Bhojraj, Sanjeev, and Bhaskaran Swaminathan, 2006, Macromomentum: Returns predictability in international equity indices, *The Journal of Business* 79, 429–451.

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, 1994, Industry returns and the Fisher effect, *Journal of Finance* 49, 1595–1615.

Boyson, Nicole, Christof W. Stahel, and René M. Stulz, 2010, Hedge fund contagion and liquidity shocks, *Journal of Finance* 65, 1789–1816.

Brunnermeier, Markus, and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Carhart, M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle and time-varying expected returns, *Journal of Finance* 57, 985–1019.

Chui, Andy, Sheridan Titman, and K.C. John Wei, 2010, Individualism and momentum around the world, *Journal of Finance* 65, 361–392.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A theory of overconfidence, self-attribution, and security market under and over-reactions, *Journal of Finance* 53, 1839–1885.

DeBondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

Erb, Claude, and Campbell Harvey, 2006, The strategic and tactical value of commodity futures, *Financial Analysts Journal* 62, 69–97.

Fama, Eugene F., 1976, *Foundations of Finance* (Basic Books, New York, NY).

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 1998, Value versus growth: The international evidence, *Journal of Finance* 53, 1975–1999.

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

Fama, Eugene F., and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2012, Trading costs of asset pricing anomalies Working paper, University of Chicago.

Frazzini, Andrea, and Lasse H. Pedersen, 2010, *Betting against beta*, Working paper, AQR Capital Management and New York University.

Garleanu, Nicolae, and Lasse H. Pedersen, 2012, Dynamic trading with predictable returns and transaction costs, *Journal of Finance*, forthcoming.

Gerakos, Joseph, and Juhani Linnainmaa, 2012, *Decomposing value*, Working paper, University of Chicago.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gomes, Joao F., Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst, 2008, *The fundamentals of commodity futures returns*, Working paper, University of Pennsylvania.

Goyal, Amit, and Sunil Wahal, 2012, *Is momentum an echo*? Working paper, Arizona State University.

Griffin, John, Susan Ji, and Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 1515–1547.

Grinblatt, Mark, and Tobias J. Moskowitz, 2004, Predicting stock price movements from past returns: The role of consistency and tax-loss selling, *Journal of Financial Economics* 71, 541–579.

Hansen, Lars P., John Heaton, and Nan Li, 2008, Consumption strikes back?: Measuring long run risk, *Journal of Political Economy* 116, 260–302.

Hansen, Lars P., and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557–590.

Hong, Harrison, Terrence Lim, and Jeremy Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and Jeremy Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Israel, Ronen, and Tobias J. Moskowitz, 2012, The role of shorting, firm size, and time on market anomalies, *Journal of Financial Economics* (forthcoming).

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Johnson, Timothy, 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Kho, Bong C., 1996, Time-varying risk premia, volatility, and technical trading rule profits: Evidence from foreign currency futures markets, *Journal of Financial Economics* 41, 249–290.

Koijen, Ralph, Tobias Moskowitz, Lasse H. Pedersen, and Evert Vrugt, 2012, Carry, Working paper, University of Chicago.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs? *Journal of Finance* 59, 1039–1082.

Korajczyk, Robert A., and Ronnie Sadka, 2008, Pricing the commonality across alternative measures of liquidity, *Journal of Financial Economics* 87, 45–72.

Krishnamurthy, A., 2002, The bond/old-bond spread, *Journal of Financial Economics* 66, 463–506.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lebaron, B., 1999, Technical trading rule profitability and foreign exchange intervention, *Journal of International Economics*, 49, 125–143.

Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2003, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349–380.

Li, Dongmei, and Lu Zhang, 2010, Does *q*-theory with investment frictions explain anomalies in the cross-section of returns? *Journal of Financial Economics* 98, 297–314.

Li, Erica X. N., Dmitry Livdan, and Lu Zhang, 2009, Anomalies, *Review of Financial Studies* 22, 4301–4334.

Liew, Jimmy, and Maria Vassalou, 2000, Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics* 57, 221–245.

Liu, Laura Xiaolei, Toni M. Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105–1139.

Liu, Laura Xiaolei, and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Lo, Andrew W., and A. Craig MacKinaly, 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3, 175–205.

Malloy, Christopher, Tobias J. Moskowitz, and Annette Vissing-Jorgensen, 2009, Long-run stockholder consumption risk and asset returns, *Journal of Finance* 64, 2427–2479.

Moskowitz, Tobias J., 2003, An analysis of covariance risk and pricing anomalies, *Review of Financial Studies* 16, 417–457.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

Moskowitz, Tobias J., and Lasse Heje Pedersen, 2012, Global market and funding liquidity, Working paper, University of Chicago.

Novy-Marx, Robert, 2012, Is momentum really momentum? *Journal of Financial Economics* 103, 429–453.

Parker, Jonathan A., and Christian Julliard, 2005, Consumption risk and the cross section of expected returns, *Journal of Political Economy* 113, 185–222.

Pástor, Ľuboš , and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Pedersen, Lasse, 2009, When everyone runs for the exit, *International Journal of Central Banking* 5, 177–199.

Piotroski, Joseph, 2000, Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research*, 38, 1–41.

Roon, Frank A., Theo de Nijman, and Christian H. Veld, 2000, Hedging pressure effects in futures markets, *Journal of Finance*, 55, 1437–1456.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–16.

Rouwenhorst, K. Geert, 1998, International momentum strategies, *Journal of Finance* 53, 267–284.

Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80, 309–349.

Sadka, Ronnie, 2012, Hedge-fund performance and liquidity risk, *Journal of Investment Management* 10, 60–72.

Sagi, Jacob, and Mark Seasholes, 2007, Firm-specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389–434.

Shleifer, Andrei, and Lawrence H. Summers, 1990, The noise trader approach to finance, *Journal of Economic Perspectives* 4, 19–33.

Stattman, Dennis, 1980, Book values and stock returns, *Chicago MBA: A Journal of Selected Papers* 5, 25–45.

Vayanos, Dimitri, and Paul Woolley, 2012, An institutional theory of momentum and reversal, Working paper, London School of Economics.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

This page intentionally left blank

Betting Against Beta*

Andrea Frazzini and Lasse Heje Pedersen

We present a model with leverage and margin constraints that vary across investors and time. We find evidence consistent with each of the model's five central predictions: (1) Because constrained investors bid up high-beta assets, high beta is associated with low alpha, as we find empirically for US equities, 20 international equity markets, Treasury bonds, corporate bonds, and futures. (2) A betting against beta (BAB) factor, which is long leveraged low-beta assets and short high-beta assets, produces significant positive risk-adjusted returns. (3) When funding constraints tighten, the return of the BAB factor is low. (4) Increased funding liquidity risk compresses betas toward one. (5) More constrained investors hold riskier assets.

1. Introduction

A basic premise of the capital asset pricing model (CAPM) is that all agents invest in the portfolio with the highest expected excess return per unit of risk (Sharpe ratio) and leverage or de-leverage this portfolio to suit their risk preferences. However, many investors, such as individuals, pension funds, and mutual funds, are constrained in the leverage that they can take, and they therefore overweight risky securities instead of using leverage.

* We thank Cliff Asness, Aaron Brown, John Campbell, Josh Coval (discussant), Kent Daniel, Gene Fama, Nicolae Garleanu, John Heaton (discussant), Michael Katz, Owen Lamont, Juhani Linnainmaa (discussant), Michael Mendelson, Mark Mitchell, Lubos Pastor (discussant), Matt Richardson, William Schwert (editor), Tuomo Vuolteenaho, Robert Whitelaw and two anonymous referees for helpful comments and discussions as well as seminar participants at AQR Capital Management, Columbia University, New York University, Yale University, Emory University, University of Chicago Booth School of Business, Northwestern University Kellogg School of Management, Harvard University, Boston University, Vienna University of Economics and Business, University of Mannheim, Goethe University Frankfurt, the American Finance Association meeting, NBER conference, Utah Winter Finance Conference, Annual Management Conference at University of Chicago Booth School of Business, Bank of America and Merrill Lynch Quant Conference, and Nomura Global Quantitative Investment Strategies Conference. Lasse Heje Pedersen gratefully acknowledges support from the European Research Council (ERC Grant no. 312417) and the FRIC Center for Financial Frictions (Grant no. DNRF102).

Reprinted from the *Journal of Financial Economics*, January 2014, Vol. 111 (1) 1–25, with permission.

Recipient of the 2015 Fama/DFA Prize (first place), the 2011 Swiss Finance Institute Outstanding Paper Award, and the 2011 Roger F. Murray Prize (third place).

For instance, many mutual fund families offer balanced funds in which the "normal" fund may invest around 40% in long-term bonds and 60% in stocks, whereas the "aggressive" fund invests 10% in bonds and 90% in stocks. If the "normal" fund is efficient, then an investor could leverage it and achieve a better trade-off between risk and expected return than the aggressive portfolio with a large tilt toward stocks. The demand for exchange-traded funds (ETFs) with embedded leverage provides further evidence that many investors cannot use leverage directly.

This behavior of tilting toward high-beta assets suggests that risky high-beta assets require lower risk-adjusted returns than low-beta assets, which require leverage. Indeed, the security market line for US stocks is too flat relative to the CAPM (Black, Jensen, and Scholes, 1972) and is better explained by the CAPM with restricted borrowing than the standard CAPM [see Black (1972, 1993), Brennan (1971), and Mehrling (2005) for an excellent historical perspective].

Several questions arise: How can an unconstrained arbitrageur exploit this effect, i.e., how do you bet against beta? What is the magnitude of this anomaly relative to the size, value, and momentum effects? Is betting against beta rewarded in other countries and asset classes? How does the return premium vary over time and in the cross section? Who bets against beta?

We address these questions by considering a dynamic model of leverage constraints and by presenting consistent empirical evidence from 20 international stock markets, Treasury bond markets, credit markets, and futures markets.

Our model features several types of agents. Some agents cannot use leverage and, therefore, overweight high-beta assets, causing those assets to offer lower returns. Other agents can use leverage but face margin constraints. Unconstrained agents underweight (or short-sell) high-beta assets and buy low-beta assets that they lever up. The model implies a flatter security market line (as in Black (1972)), where the slope depends on the tightness (i.e., Lagrange multiplier) of the funding constraints on average across agents (Proposition 1).

One way to illustrate the asset pricing effect of the funding friction is to consider the returns on market-neutral betting against beta (BAB) factors. A BAB factor is a portfolio that holds low-beta assets, leveraged to a beta of one, and that shorts high-beta assets, de-leveraged to a beta of one. For instance, the BAB factor for US stocks achieves a zero beta by holding \$1.4 of low-beta stocks and shortselling \$0.7 of high-beta stocks, with offsetting positions in the risk-free asset to make it self-financing.[1] Our model predicts that BAB factors have a positive average return and that the return is increasing in the ex ante tightness of constraints and in the spread in betas between high- and low-beta securities (Proposition 2).

When the leveraged agents hit their margin constraint, they must de-leverage. Therefore, the model predicts that, during times of tightening funding liquidity constraints, the BAB factor realizes negative returns as its expected future return rises (Proposition 3). Furthermore, the model predicts that the betas of securities in the cross section are compressed toward one when funding liquidity risk is high (Proposition 4). Finally, the model implies that more-constrained investors overweight high-beta assets in their portfolios and less-constrained investors overweight low-beta assets and possibly apply leverage (Proposition 5).

Our model thus extends the Black (1972) insight by considering a broader set of constraints and deriving the dynamic time series and cross-sectional properties arising from the equilibrium interaction between agents with different constraints.

We find consistent evidence for each of the model's central predictions. To test Proposition 1, we first consider portfolios sorted by beta within each asset class. We find that alphas and Sharpe ratios are almost monotonically declining in beta in each asset class. This finding provides broad evidence that the relative flatness of the security market line is not isolated to the US stock market but that it is a pervasive global phenomenon. Hence, this pattern of required returns is likely driven by a common economic cause, and our funding constraint model provides one such unified explanation.

To test Proposition 2, we construct BAB factors within the US stock market and within each of the 19 other developed MSCI stock markets. The US BAB factor realizes a Sharpe ratio of 0.78 between 1926 and March 2012. To put this BAB factor return in perspective, note that its Sharpe ratio is about twice that of the value effect and 40% higher than that of momentum over the same time period. The BAB factor has highly significant risk-adjusted returns, accounting for its realized exposure to market, value, size, momentum, and liquidity factors (i.e., significant one-, three-, four-, and five-factor alphas), and it realizes a significant positive return in each of the four 20-year subperiods between 1926 and 2012.

We find similar results in our sample of international equities. Combining stocks in each of the non-US countries produces a BAB factor with returns about as strong as the US BAB factor.

We show that BAB returns are consistent across countries, time, within deciles sorted by size, and within deciles sorted by idiosyncratic risk and are robust to a number of specifications. These consistent results suggest that coincidence or data mining are unlikely explanations. However, if leverage constraints are the underlying drivers as in our model, then the effect should also exist in other markets.

Hence, we examine BAB factors in other major asset classes. For US Treasuries, the BAB factor is a portfolio that holds leveraged low-beta (i.e., short-maturity) bonds and shortsells de-leveraged high-beta (i.e., long-term) bonds. This portfolio produces highly significant risk-adjusted returns with a Sharpe ratio of 0.81. This profitability of shortselling long-term bonds could seem to contradict the well-known "term premium" in fixed income markets. There is no paradox, however. The term premium means that investors are compensated on average for holding long-term bonds instead of T-bills because of the need for maturity transformation. The term premium exists at all horizons, however. Just as investors are compensated for holding ten-year bonds over T-bills, they are also compensated for holding one-year bonds. Our finding is that the compensation per unit of risk is in fact larger for the one-year bond than for the ten-year bond. Hence, a portfolio that has a leveraged long position in one-year (and other short-term) bonds and a short position in long-term bonds produces positive returns. This result is consistent with our model in which some investors are leverage-constrained in their bond exposure and, therefore, require lower risk-adjusted returns for long-term bonds that give more "bang for the buck." Indeed, short-term bonds require tremendous leverage to achieve similar risk or return as long-term bonds. These results complement those of Fama (1984, 1986) and Duffee

(2010), who also consider Sharpe ratios across maturities implied by standard term structure models.

We find similar evidence in credit markets: A leveraged portfolio of highly rated corporate bonds outperforms a de-leveraged portfolio of low-rated bonds. Similarly, using a BAB factor based on corporate bond indices by maturity produces high risk-adjusted returns.

We test the time series predictions of Proposition 3 using the TED spread as a measure of funding conditions. Consistent with the model, a higher TED spread is associated with low contemporaneous BAB returns. The lagged TED spread predicts returns negatively, which is inconsistent with the model if a high TED spread means a high tightness of investors' funding constraints. This result could be explained if higher TED spreads meant that investors' funding constraints would be tightening as their banks reduce credit availability over time, though this is speculation.

To test the prediction of Proposition 4, we use the volatility of the TED spread as an empirical proxy for funding liquidity risk. Consistent with the model's beta-compression prediction, we find that the dispersion of betas is significantly lower when funding liquidity risk is high.

Lastly, we find evidence consistent with the model's portfolio prediction that more-constrained investors hold higher-beta securities than less-constrained investors (Proposition 5). We study the equity portfolios of mutual funds and individual investors, which are likely to be constrained. Consistent with the model, we find that these investors hold portfolios with average betas above one. On the other side of the market, we find that leveraged buyout (LBO) funds acquire firms with average betas below 1 and apply leverage. Similarly, looking at the holdings of Warren Buffett's firm Berkshire Hathaway, we see that Buffett bets against beta by buying low-beta stocks and applying leverage (analyzed further in Frazzini, Kabiller, and Pedersen (2012)).

Our results shed new light on the relation between risk and expected returns. This central issue in financial economics has naturally received much attention. The standard CAPM beta cannot explain the cross section of unconditional stock returns (Fama and French, 1992) or conditional stock returns (Lewellen and Nagel, 2006). Stocks with high beta have been found to deliver low risk-adjusted returns (Black, Jensen, and Scholes, 1972; Baker, Bradley, and Wurgler, 2011); thus, the constrained-borrowing CAPM has a better fit (Gibbons, 1982; Kandel, 1984; Shanken, 1985). Stocks with high idiosyncratic volatility have realized low returns (Falkenstein, 1994; Ang, Hodrick, Xing, Zhang, 2006, 2009), but we find that the beta effect holds even when controlling for idiosyncratic risk.[2] Theoretically, asset pricing models with benchmarked managers (Brennan, 1993) or constraints imply more general CAPM-like relations (Hindy, 1995; Cuoco, 1997). In particular, the margin-CAPM implies that high-margin assets have higher required returns, especially during times of funding illiquidity (Garleanu and Pedersen, 2011; Ashcraft, Garleanu, and Pedersen, 2010). Garleanu and Pedersen (2011) show empirically that deviations of the law of one price arise when high-margin assets become cheaper than low-margin assets, and Ashcraft, Garleanu, and Pedersen (2010) find that prices increase when central bank lending facilities reduce margins. Furthermore, funding liquidity risk is linked to market liquidity risk (Gromb and Vayanos, 2002; Brunnermeier and Pedersen, 2009), which also affects required returns (Acharya and Pedersen, 2005). We complement the literature by deriving new cross-sectional and time series predictions in a simple

dynamic model that captures leverage and margin constraints and by testing its implications across a broad cross section of securities across all the major asset classes. Finally, Asness, Frazzini, and Pedersen (2012) report evidence of a low-beta effect across asset classes consistent with our theory.

The rest of the paper is organized as follows. Section 2 lays out the theory, Section 3 describes our data and empirical methodology, Sections 4–7 test Propositions 1–5, and Section 8 concludes. Appendix A contains all proofs, Appendix B provides a number of additional empirical results and robustness tests, and Appendix C provides a calibration of the model. The calibration shows that, to match the strong BAB performance in the data, a large fraction of agents must face severe constraints. An interesting topic for future research is to empirically estimate agents' leverage constraints and risk preferences and study whether the magnitude of the BAB returns is consistent with the model or should be viewed as a puzzle.

2. Theory

We consider an overlapping-generations (OLG) economy in which agents $i = 1,\ldots,I$ are born each time period t with wealth W_t^i and live for two periods. Agents trade securities $s = 1,\ldots,S$, where security s pays dividends δ_t^s and has χ^{*s} shares outstanding.[3] Each time period t, young agents choose a portfolio of shares $\chi = (\chi^1,\ldots,\chi^S)'$, investing the rest of their wealth at the risk-free return r^f, to maximize their utility:

$$\max \chi'\left(E_t\left(P_{t+1}+\delta_{t+1}\right)-\left(1+r^f\right)P_t\right)-\frac{\gamma^i}{2}\chi'\Omega_t\chi, \tag{1}$$

where P_t is the vector of prices at time t, Ω_t is the variance–covariance matrix of $P_{t+1} + \delta_{t+1}$, and γ^i is agent i's risk aversion. Agent i is subject to the portfolio constraint

$$m_t^i \sum_s \chi^s P_t^s \leq W_t^i \tag{2}$$

This constraint requires that some multiple m_t^i of the total dollars invested, the sum of the number of shares x^s times their prices P^s, must be less than the agent's wealth.

The investment constraint depends on the agent i. For instance, some agents simply cannot use leverage, which is captured by $m^i = 1$ [as Black (1972) assumes]. Other agents not only could be precluded from using leverage but also must have some of their wealth in cash, which is captured by m^i greater than one. For instance, $m^i = 1/(1 - 0.20) = 1.25$ represents an agent who must hold 20% of her wealth in cash. For instance, a mutual fund could need some ready cash to be able to meet daily redemptions, an insurance company needs to pay claims, and individual investors may need cash for unforeseen expenses.

Other agents could be able to use leverage but could face margin constraints. For instance, if an agent faces a margin requirement of 50%, then his m^i is 0.50. With this margin requirement, the agent can invest in assets worth twice his wealth at most. A smaller margin requirement m^i naturally means that the agent can take greater positions. Our formulation assumes for simplicity that all securities have the same margin requirement, which may be true when comparing securities within the same asset class (e.g., stocks), as we do empirically. Garleanu and Pedersen (2011) and Ashcraft, Garleanu, and Pedersen (2010) consider assets with different margin requirements and show theoretically and

empirically that higher margin requirements are associated with higher required returns (Margin CAPM).

We are interested in the properties of the competitive equilibrium in which the total demand equals the supply:

$$\sum_i \chi^i = \chi^* \tag{3}$$

To derive equilibrium, consider the first order condition for agent i:

$$0 = E_t\left(P_{t+1} + \delta_{t+1}\right) - \left(1 + r^f\right)P_t - \gamma^i \Omega \chi^i - \psi_t^i P_t \tag{4}$$

where ψ^i is the Lagrange multiplier of the portfolio constraint. Solving for x^i gives the optimal position:

$$\chi^i = \frac{1}{\gamma^i}\Omega^{-1}\left(E_t\left(P_{t+1} + \delta_{t+1}\right) - \left(1 + r^f + \psi_t^i\right)P_t\right). \tag{5}$$

The equilibrium condition now follows from summing over these positions:

$$\chi^* = \frac{1}{\gamma}\Omega^{-1}\left(E_t\left(P_{t+1} + \delta_{t+1}\right) - \left(1 + r^f + \psi_t\right)P_t\right), \tag{6}$$

where the aggregate risk aversion γ is defined by $1/\gamma = \Sigma_i 1/\gamma^i$ and $\psi_t = \sum_i \left(\gamma/\gamma^i\right)\psi_t^i$ is the weighted average Lagrange multiplier. (The coefficients γ/γ^i sum to one by definition of the aggregate risk aversion γ.) The equilibrium price can then be computed:

$$P_t = \frac{E_t\left(P_{t+1} + \delta_{t+1}\right) - \gamma\Omega\chi^*}{1 + r^f + \psi_t}, \tag{7}$$

Translating this into the return of any security $r_{t+1}^i = \left(P_{t+1}^i + \delta_{t+1}^i\right)/P_t^i - 1$, the return on the market r_{t+1}^M, and using the usual expression for beta, $\beta_t^s = \mathrm{cov}_t\left(r_{t+1}^s, r_{t+1}^M\right)/\mathrm{var}_t\left(r_{t+1}^M\right)$, we obtain following results. (All are in Appendix A, which also illustrates the portfolio choice with leverage constraints in a mean-standard deviation diagram.)

Proposition 1 (high beta is low alpha).

(i) *The equilibrium required return for any security s is*

$$E_t\left(r_{t+1}^s\right) = r^f + \psi_t + \beta_t^s \lambda_t \tag{8}$$

where the risk premium is $\lambda_t = E_t\left(r_{t+1}^M\right) - r^f - \psi_t$ *and* ψ_t *is the average Lagrange multiplier, measuring the tightness of funding constraints.*

(ii) *A security's alpha with respect to the market is* $\alpha_t^s = \psi_t\left(1 - \beta_t^s\right)$. *The alpha decreases in the beta,* β_t^s.

(iii) *For an efficient portfolio, the Sharpe ratio is highest for an efficient portfolio with a beta less than one and decreases in* β_t^s *for higher betas and increases for lower betas.*

As in Black's CAPM with restricted borrowing (in which $m^i = 1$ for all agents), the required return is a constant plus beta times a risk premium. Our expression shows explicitly how risk premia are affected by the tightness of agents' portfolio constraints, as measured by the average Lagrange multiplier ψ_t. Tighter portfolio constraints (i.e., a larger ψ_t) flatten the security market line by increasing the intercept and decreasing the slope λ_t.

Whereas the standard CAPM implies that the intercept of the security market line is r^f, the intercept here is increased by binding funding constraints (through the weighted average of the agents' Lagrange multipliers). One could wonder why zero-beta assets require returns in excess of the risk-free rate. The answer has two parts. First, constrained agents prefer to invest their limited capital in riskier assets with higher expected return. Second, unconstrained agents do invest considerable amounts in zero-beta assets so, from their perspective, the risk of these assets is not idiosyncratic, as additional exposure to such assets would increase the risk of their portfolio. Hence, in equilibrium, zero-beta risky assets must offer higher returns than the risk-free rate.

Assets that have zero covariance to the Tobin (1958) "tangency portfolio" held by an unconstrained agent do earn the risk-free rate, but the tangency portfolio is not the market portfolio in our equilibrium. The market portfolio is the weighted average of all investors' portfolios, i.e., an average of the tangency portfolio held by unconstrained investors and riskier portfolios held by constrained investors. Hence, the market portfolio has higher risk and expected return than the tangency portfolio, but a lower Sharpe ratio.

The portfolio constraints further imply a lower slope λ_t of the security market line, i.e., a lower compensation for a marginal increase in systematic risk. The slope is lower because constrained agents need high unleveraged returns and are, therefore, willing to accept less compensation for higher risk.[4]

We next consider the properties of a factor that goes long low-beta assets and shortsells high-beta assets. To construct such a factor, let w_L be the relative portfolio weights for a portfolio of low-beta assets with return $r^L_{t+1} = w'_L r_{t+1}$ and consider similarly a portfolio of high-beta assets with return r^H_{t+1}. The betas of these portfolios are denoted β^L_t and β^H_t, where $\beta^L_t < \beta^H_t$. We then construct a betting against beta (BAB) factor as

$$r^{BAB}_{t+1} = \frac{1}{\beta^L_t}\left(r^L_{t+1} - r^f\right) - \frac{1}{\beta^H_t}\left(r^H_{t+1} - r^f\right) \tag{9}$$

this portfolio is market-neutral; that is, it has a beta of zero. The long side has been leveraged to a beta of one, and the short side has been de-leveraged to a beta of one. Furthermore, the BAB factor provides the excess return on a self-financing portfolio, such as HML (high minus low) and SMB (small minus big), because it is a difference between excess returns. The difference is that BAB is not dollar-neutral in terms of only the risky securities because this would not produce a beta of zero.[5] The model has several predictions regarding the BAB factor.

Proposition 2 (positive expected return of BAB). The expected excess return of the self-financing BAB factor is positive

$$E_t\left(r^{BAB}_{t+1}\right) = \frac{\beta^H_t - \beta^L_t}{\beta^L_t \beta^H_t}\psi_t \geq 0 \tag{10}$$

and increasing in the ex ante beta spread $\left(\beta^H_t - \beta^L_t\right) / \left(\beta^L_t \beta^H_t\right)$ *and funding tightness* ψ_t.

Proposition 2 shows that a market-neutral BAB portfolio that is long leveraged low-beta securities and short higher-beta securities earns a positive expected return on average. The size of the expected return depends on the spread in the betas and how binding the portfolio constraints are in the market, as captured by the average of the Lagrange multipliers ψ_t.

Proposition 3 considers the effect of a shock to the portfolio constraints (or margin requirements), m^k, which can be interpreted as a worsening of funding liquidity, a credit crisis in the extreme. Such a funding liquidity shock results in losses for the BAB factor as its required return increases. This happens because agents may need to de-leverage their bets against beta or stretch even further to buy the high-beta assets. Thus, the BAB factor is exposed to funding liquidity risk, as it loses when portfolio constraints become more binding.

Proposition 3 (funding shocks and BAB returns). A tighter portfolio constraint, that is, an increase in m_t^k for some of k, leads to a contemporaneous loss for the BAB factor

$$\frac{\partial r_t^{BAB}}{\partial m_t^k} \leq 0 \tag{11}$$

and an increase in its future required return:

$$\frac{\partial E_t\left(r_{t+1}^{BAB}\right)}{\partial m_t^k} \geq 0 \tag{12}$$

Funding shocks have further implications for the cross section of asset returns and the BAB portfolio. Specifically, a funding shock makes all security prices drop together (that is, $\left(\partial P_t^s / \partial \psi_t\right) / P_t^s$ is the same for all securities s). Therefore, an increased funding risk compresses betas toward one.[6] If the BAB portfolio construction is based on an information set that does not account for this increased funding risk, then the BAB portfolio's conditional market beta is affected.

Proposition 4 (beta compression). Suppose that all random variables are identically and independently distributed (i.i.d.) over time and δ_t is independent of the other random variables. Further, at time $t-1$ after the BAB portfolio is formed and prices are set, the conditional variance of the discount factor $1/(1+r^f+\psi_t)$ rises (falls) due to new information about m_t and W_t. Then,

(i) *The conditional return betas β_{t-1}^i of all securities are compressed toward one (more dispersed), and*
(ii) *The conditional beta of the BAB portfolio becomes positive (negative), even though it is market neutral relative to the information set used for portfolio formation.*

In addition to the asset-pricing predictions that we derive, funding constraints naturally affect agents' portfolio choices. In particular, more-constrained investors tilt toward riskier securities in equilibrium and less-constrained agents tilt toward safer securities with higher reward per unit of risk. To state this result, we write next period's security payoffs as

$$P_{t+1}+\delta_{t+1} = E_t\left(P_{t+1}+\delta_{t+1}\right) + b\left(P_{t+1}^M+\delta_{t+1}^M - E_t\left(P_{t+1}^M+\delta_{t+1}^M\right)\right) + e \tag{13}$$

where b is a vector of market exposures, and e is a vector of noise that is uncorrelated with the market. We have the following natural result for the agents' positions.

Proposition 5 (constrained investors hold high betas). Unconstrained agents hold a portfolio of risky securities that has a beta less than one; constrained agents hold portfolios

Table 1 Summary Statistics: Equities
This table shows summary statistics as of June of each year. The sample includes all commons stocks on the Center for Research in Security Prices daily stock files (shrcd equal to 10 or 11) and Xpressfeed Global security files (tcpi equal to zero). Mean ME is the average market value of equity, in billions of US dollars. Means are pooled averages as of June of each year.

Country	Local market index	Number of stocks, total	Number of stocks, mean	Mean ME (firm, billion of US dollars)	Mean ME (market, billion of US dollars)	Start year	End year
Australia	MSCI Australia	3,047	894	0.57	501	1989	2012
Austria	MSCI Austria	211	81	0.75	59	1989	2012
Belgium	MSCI Belgium	425	138	1.79	240	1989	2012
Canada	MSCI Canada	5,703	1,180	0.89	520	1984	2012
Denmark	MSCI Denmark	413	146	0.83	119	1989	2012
Finland	MSCI Finland	293	109	1.39	143	1989	2012
France	MSCI France	1,815	589	2.12	1,222	1989	2012
Germany	MSCI Germany	2,165	724	2.48	1,785	1989	2012
Hong Kong	MSCI Hong Kong	1,793	674	1.22	799	1989	2012
Italy	MSCI Italy	610	224	2.12	470	1989	2012
Japan	MSCI Japan	5,009	2,907	1.19	3,488	1989	2012
Netherlands	MSCI Netherlands	413	168	3.33	557	1989	2012
New Zealand	MSCI New Zealand	318	97	0.87	81	1989	2012
Norway	MSCI Norway	661	164	0.76	121	1989	2012
Singapore	MSCI Singapore	1,058	375	0.63	240	1989	2012
Spain	MSCI Spain	376	138	3.00	398	1989	2012
Sweden	MSCI Sweden	1,060	264	1.30	334	1989	2012
Switzerland	MSCI Switzerland	566	210	3.06	633	1989	2012
United Kingdom	MSCI UK	6,126	1,766	1.22	2,243	1989	2012
United States	CRSP value-weighted index	23,538	3,182	0.99	3,215	1926	2012

of risky securities with higher betas. If securities s and k are identical except that s has a larger market exposure than k, $b^s > b^k$, then any constrained agent j with greater than average Lagrange multiplier, $\psi_t^j > \psi_t$, holds more shares of s than k. The reverse is true for any agent with $\psi_t^j < \psi_t$.

We next provide empirical evidence for Propositions 1–5. Beyond matching the data qualitatively, Appendix C illustrates how well a calibrated model can quantitatively match the magnitude of the estimated BAB returns.

3. Data and Methodology

The data in this study are collected from several sources. The sample of US and international equities has 55,600 stocks covering 20 countries, and the summary statistics for stocks are reported in Table 1. Stock return data are from the union of the Center for Research in Security Prices (CRSP) tape and the Xpressfeed Global database. Our US equity data include all available common stocks on CRSP between January 1926 and March 2012, and betas are computed with respect to the CRSP value-weighted market index. Excess returns are above the US Treasury bill rate. We consider alphas with respect to the market factor and factor returns based on size (SMB), book-to-market (HML), momentum (up minus down, UMD), and (when available) liquidity risk.[7]

The international equity data include all available common stocks on the Xpressfeed Global daily security file for 19 markets belonging to the MSCI developed universe between January 1989 and March 2012. We assign each stock to its corresponding market based on the location of the primary exchange. Betas are computed with respect to the corresponding MSCI local market index.[8]

All returns are in US dollars, and excess returns are above the US Treasury bill rate. We compute alphas with respect to the international market and factor returns based on size (SMB), book-to-market (HML), and momentum (UMD) from Asness and Frazzini (2013) and (when available) liquidity risk.[9]

We also consider a variety of other assets. Table 2 contains the list of instruments and the corresponding ranges of available data. We obtain US Treasury bond data from the CRSP US Treasury Database, using monthly returns (in excess of the one-month Treasury bill) on the Fama Bond portfolios for maturities ranging from one to ten years between January 1952 and March 2012. Each portfolio return is an equal-weighted average of the unadjusted holding period return for each bond in the portfolio. Only non-callable, non-flower notes and bonds are included in the portfolios. Betas are computed with respect to an equally weighted portfolio of all bonds in the database.

We collect aggregate corporate bond index returns from Barclays Capital's Bond.Hub database.[10] Our analysis focuses on the monthly returns (in excess of the one-month Treasury bill) of four aggregate US credit indices with maturity ranging from one to ten years and nine investment-grade and high-yield corporate bond portfolios with credit risk ranging from AAA to Ca-D and Distressed.[11] The data cover the period between January 1973 and March 2012, although the data availability varies depending on the individual bond series. Betas are computed with respect to an equally weighted portfolio of all bonds in the database.

We also study futures and forwards on country equity indexes, country bond indexes, foreign exchange, and commodities. Return data are drawn from the internal pricing data

Table 2 Summary Statistics: Other Asset Classes

This table reports the securities included in our data sets and the corresponding date range.

Asset class	Instrument	Frequency	Start year	End year
Equity indices	Australia	Daily	1977	2012
	Germany	Daily	1975	2012
	Canada	Daily	1975	2012
	Spain	Daily	1980	2012
	France	Daily	1975	2012
	Hong Kong	Daily	1980	2012
	Italy	Daily	1978	2012
	Japan	Daily	1976	2012
	Netherlands	Daily	1975	2012
	Sweden	Daily	1980	2012
	Switzerland	Daily	1975	2012
	United Kingdom	Daily	1975	2012
	United States	Daily	1965	2012
Country bonds	Australia	Daily	1986	2012
	Germany	Daily	1980	2012
	Canada	Daily	1985	2012
	Japan	Daily	1982	2012
	Norway	Daily	1989	2012
	Sweden	Daily	1987	2012
	Switzerland	Daily	1981	2012
	United Kingdom	Daily	1980	2012
	United States	Daily	1965	2012
Foreign exchange	Australia	Daily	1977	2012
	Germany	Daily	1975	2012
	Canada	Daily	1975	2012
	Japan	Daily	1976	2012
	Norway	Daily	1989	2012
	New Zealand	Daily	1986	2012
	Sweden	Daily	1987	2012
	Switzerland	Daily	1975	2012
	United Kingdom	Daily	1975	2012
US Treasury bonds	Zero to one year	Monthly	1952	2012
	One to two years	Monthly	1952	2012
	Two to three years	Monthly	1952	2012
	Three to four years	Monthly	1952	2012
	Four to five years	Monthly	1952	2012
	Four to ten years	Monthly	1952	2012
	More than ten years	Monthly	1952	2012

(*Continued*)

Table 2 (*Continued*)

Asset class	Instrument	Frequency	Start year	End year
Credit indices	One to three years	Monthly	1976	2012
	Three to five year	Monthly	1976	2012
	Five to ten years	Monthly	1991	2012
	Seven to ten years	Monthly	1988	2012
Corporate bonds	Aaa	Monthly	1973	2012
	Aa	Monthly	1973	2012
	A	Monthly	1973	2012
	Baa	Monthly	1973	2012
	Ba	Monthly	1983	2012
	B	Monthly	1983	2012
	Caa	Monthly	1983	2012
	Ca-D	Monthly	1993	2012
	Distressed	Monthly	1986	2012
Commodities	Aluminum	Daily	1989	2012
	Brent oil	Daily	1989	2012
	Cattle	Daily	1989	2012
	Cocoa	Daily	1984	2012
	Coffee	Daily	1989	2012
	Copper	Daily	1989	2012
	Corn	Daily	1989	2012
	Cotton	Daily	1989	2012
	Crude	Daily	1989	2012
	Gasoil	Daily	1989	2012
	Gold	Daily	1989	2012
	Heat oil	Daily	1989	2012
	Hogs	Daily	1989	2012
	Lead	Daily	1989	2012
	Nat gas	Daily	1989	2012
	Nickel	Daily	1984	2012
	Platinum	Daily	1989	2012
	Silver	Daily	1989	2012
	Soymeal	Daily	1989	2012
	Soy oil	Daily	1989	2012
	Sugar	Daily	1989	2012
	Tin	Daily	1989	2012
	Unleaded	Daily	1989	2012
	Wheat	Daily	1989	2012
	Zinc	Daily	1989	2012

maintained by AQR Capital Management LLC. The data are collected from a variety of sources and contain daily return on futures, forwards, or swap contracts in excess of the relevant financing rate. The type of contract for each asset depends on availability or the

relative liquidity of different instruments. Prior to expiration, positions are rolled over into the next most-liquid contract. The rolling date's convention differs across contracts and depends on the relative liquidity of different maturities. The data cover the period between January 1963 and March 2012, with varying data availability depending on the asset class. For more details on the computation of returns and data sources, see Moskowitz, Ooi, and Pedersen (2012), Appendix A. For equity indexes, country bonds, and currencies, the betas are computed with respect to a gross domestic product (GDP)-weighted portfolio, and for commodities, the betas are computed with respect to a diversified portfolio that gives equal risk weight across commodities.

Finally, we use the TED spread as a proxy for time periods when credit constraints are more likely to be binding [as in Garleanu and Pedersen (2011) and others]. The TED spread is defined as the difference between the three-month Eurodollar LIBOR and the three-month US Treasuries rate. Our TED data run from December 1984 to March 2012.

3.1. Estimating Ex Ante Betas

We estimate pre-ranking betas from rolling regressions of excess returns on market excess returns. Whenever possible, we use daily data, rather than monthly data, as the accuracy of covariance estimation improves with the sample frequency (Merton, 1980).[12] Our estimated beta for security i is given by

$$\hat{\beta}_i^{ts} = \hat{\rho}\frac{\hat{\sigma}_i}{\hat{\sigma}_m} \tag{14}$$

where $\hat{\sigma}_i$ and $\hat{\sigma}_m$ are the estimated volatilities for the stock and the market and $\hat{\rho}$ is their correlation. We estimate volatilities and correlations separately for two reasons. First, we use a one-year rolling standard deviation for volatilities and a five-year horizon for the correlation to account for the fact that correlations appear to move more slowly than volatilities.[13] Second, we use one-day log returns to estimate volatilities and overlapping three-day log returns, $r_{i,t}^{3d} = \sum_{k=0}^{2} \ln\left(1 + r_{t+k}^{i}\right)$, for correlation to control for nonsynchronous trading (which affects only correlations). We require at least six months (120 trading days) of non-missing data to estimate volatilities and at least three years (750 trading days) of non-missing return data for correlations. If we have access only to monthly data, we use rolling one- and five-year windows and require at least 12 and 36 observations.

Finally, to reduce the influence of outliers, we follow Vasicek (1973) and Elton, Gruber, Brown, and Goetzmann (2003) and shrink the time series estimate of beta $\left(\beta_i^{TS}\right)$ toward the cross-sectional mean (β^{XS}):

$$\hat{\beta}_i = w_i\hat{\beta}_i^{TS} + \left(1 - w_i\right)\hat{\beta}^{XS} \tag{15}$$

for simplicity, instead of having asset-specific and time-varying shrinkage factors as in Vasicek (1973), we set $w = 0.6$ and $\beta^{XS} = 1$ for all periods and across all assets. However, our results are very similar either way.[14]

Our choice of the shrinkage factor does not affect how securities are sorted into portfolios because the common shrinkage does not change the ranks of the security betas.

However, the amount of shrinkage affects the construction of the BAB portfolios because the estimated betas are used to scale the long and short sides of portfolio as seen in Eq. (9).

To account for the fact that noise in the ex ante betas affects the construction of the BAB factors, our inference is focused on realized abnormal returns so that any mismatch between ex ante and (ex post) realized betas is picked up by the realized loadings in the factor regression. When we regress our portfolios on standard risk factors, the realized factor loadings are not shrunk as above because only the ex ante betas are subject to selection bias. Our results are robust to alternative beta estimation procedures as we report in Appendix B.

We compute betas with respect to a market portfolio, which is either specific to an asset class or the overall world market portfolio of all assets. While our results hold both ways, we focus on betas with respect to asset class-specific market portfolios because these betas are less noisy for several reasons. First, this approach allows us to use daily data over a long time period for most asset classes, as opposed to using the most diversified market portfolio for which we only have monthly data and only over a limited time period. Second, this approach is applicable even if markets are segmented.

As a robustness test, Table B8 in Appendix B reports results when we compute betas with respect to a proxy for a world market portfolio consisting of many asset classes. We use the world market portfolio from Asness, Frazzini, and Pedersen (2012).[15] The results are consistent with our main tests as the BAB factors earn large and significant abnormal returns in each of the asset classes in our sample.

3.2. Constructing Betting against Beta Factors

We construct simple portfolios that are long low-beta securities and that shortsell high-beta securities (BAB factors). To construct each BAB factor, all securities in an asset class are ranked in ascending order on the basis of their estimated beta. The ranked securities are assigned to one of two portfolios: low-beta and high-beta. The low- (high-) beta portfolio is composed of all stocks with a beta below (above) its asset class median (or country median for international equities). In each portfolio, securities are weighted by the ranked betas (i.e., lower-beta securities have larger weights in the low-beta portfolio and higher-beta securities have larger weights in the high-beta portfolio). The portfolios are rebalanced every calendar month.

More formally, let z be the $n \times 1$ vector of beta ranks $z_i = rank\ (\beta_{it})$ at portfolio formation, and let $\bar{z} = 1_n' z / n$ be the average rank, where n is the number of securities and 1_n is an $n \times 1$ vector of ones. The portfolio weights of the low-beta and high-beta portfolios are given by

$$\begin{aligned} w_H &= k\left(z - \bar{z}\right)^+ \\ w_L &= k\left(z - \bar{z}\right)^- \end{aligned} \tag{16}$$

where k is a normalizing constant $k = 2 / 1_n' \left| z - \bar{z} \right|$ and x^+ and x^- indicate the positive and negative elements of a vector x. By construction, we have $1_n' w_H = 1$ and $1_n' w_L = 1$. To construct the BAB factor, both portfolios are rescaled to have a beta of one at portfolio formation.

The BAB is the self-financing zero-beta portfolio (8) that is long the low-beta portfolio and that shortsells the high-beta portfolio.

$$r_{t+1}^{BAB} = \frac{1}{\beta_t^L}\left(r_{t+1}^L - r^f\right) - \frac{1}{\beta_t^H}\left(r_{t+1}^H - r^f\right), \tag{17}$$

where $r_{t+1}^L = r_{t+1}^{'} w_L$, $r_{t+1}^H = r_{t+1}^{'} w_H$, $\beta_t^L = \beta_t^{'} w_L$, and $\beta_t^H = \beta_t^{'} w_H$.

For example, on average, the US stock BAB factor is long \$1.4 of low-beta stocks (financed by shortselling \$1.4 of risk-free securities) and shortsells \$0.7 of high-beta stocks (with \$0.7 earning the risk-free rate).

3.3. Data Used to Test the Theory's Portfolio Predictions

We collect mutual fund holdings from the union of the CRSP Mutual Fund Database and Thomson Financial CDA/Spectrum holdings database, which includes all registered domestic mutual funds filing with the Securities and Exchange Commission. The holdings data run from March 1980 to March 2012. We focus our analysis on open-end, actively managed, domestic equity mutual funds. Our sample selection procedure follows that of Kacperczyk, Sialm, and Zheng (2008), and we refer to their Appendix for details about the screens that were used and summary statistics of the data.

Our individual investors' holdings data are collected from a nationwide discount brokerage house and contain trades made by about 78 thousand households in the period from January 1991 to November 1996. This data set has been used extensively in the existing literature on individual investors. For a detailed description of the brokerage data set, see Barber and Odean (2000).

Our sample of buyouts is drawn from the mergers and acquisitions and corporate events database maintained by AQR/CNH Partners.[16] The data contain various items, including initial and subsequent announcement dates, and (if applicable) completion or termination date for all takeover deals in which the target is a US publicly traded firm and where the acquirer is a private company. For some (but not all) deals, the acquirer descriptor also contains information on whether the deal is a leveraged buyout (LBO) or management buyout (MBO). The data run from January 1963 to March 2012.

Finally, we download holdings data for Berkshire Hathaway from Thomson-Reuters Financial Institutional (13f) Holding Database. The data run from March 1980 to March 2012.

4. Betting against Beta in Each Asset Class

We now test how the required return varies in the cross-section of beta-sorted securities (Proposition 1) and the hypothesis that the BAB factors have positive average returns (Proposition 2). As an overview of these results, the alphas of all the beta-sorted portfolios considered in this paper are plotted in Fig. 1. We see that declining alphas across beta-sorted portfolios are general phenomena across asset classes. (Fig. B1 in Appendix B plots the Sharpe ratios of beta-sorted portfolios and also shows a consistently declining pattern.)

Figure 1. Alphas of beta-sorted portfolios. This figure shows monthly alphas. The test assets are beta-sorted portfolios. At the beginning of each calendar month, securities are ranked in ascending order on the basis of their estimated beta at the end of the previous month. The ranked securities are assigned to beta-sorted portfolios. This figure plots alphas from low beta (left) to high beta (right). Alpha is the intercept in a regression of monthly excess return. For equity portfolios, the explanatory variables are the monthly returns from Fama and French (1993), Asness and Frazzini (2013), and Carhart (1997) portfolios. For all other portfolios, the explanatory variables are the monthly returns of the market factor. Alphas are in monthly percent.

Fig. 2 plots the annualized Sharpe ratios of the BAB portfolios in the various asset classes. All the BAB portfolios deliver positive returns, except for a small insignificantly negative return in Austrian stocks. The BAB portfolios based on large numbers of securities (US stocks, international stocks, Treasuries, credits) deliver high risk-adjusted returns relative to the standard risk factors considered in the literature.

4.1. Stocks

Table 3 reports our tests for US stocks. We consider ten beta-sorted portfolios and report their average returns, alphas, market betas, volatilities, and Sharpe ratios. The average

Figure 2. Betting against beta (BAB) Sharpe ratios by asset class. This figure shows annualized Sharpe ratios of BAB factors across asset classes. To construct the BAB factors, all securities are assigned to one of two portfolios: low beta and high beta. Securities are weighted by the ranked betas and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The BAB factor is a self-financing portfolio that is long the low-beta portfolio and shorts the high-beta portfolio. Sharpe ratios are annualized.

returns of the different beta portfolios are similar, which is the well-known relatively flat security market line. Hence, consistent with Proposition 1 and with Black (1972), the alphas decline almost monotonically from the low-beta to high-beta portfolios. The alphas decline when estimated relative to a one-, three-, four-, and five-factor model. Moreover, Sharpe ratios decline monotonically from low-beta to high-beta portfolios.

The rightmost column of Table 3 reports returns of the betting against beta factor, i.e., a portfolio that is long leveraged low-beta stocks and that shortsells de-leveraged high-beta stocks, thus maintaining a beta-neutral portfolio. Consistent with Proposition 2, the BAB factor delivers a high average return and a high alpha. Specifically, the BAB factor has Fama and French (1993) abnormal returns of 0.73% per month (t-statistic = 7.39). Further adjusting returns for the Carhart (1997) momentum factor, the BAB portfolio earns abnormal returns of 0.55% per month (t-statistic = 5.59). Last, we adjust returns using a five-factor model by adding the traded liquidity factor by Pastor and Stambaugh (2003), yielding an abnormal BAB return of 0.55% per month (t-statistic = 4.09, which is lower in part because the liquidity factor is available during only half of our sample). While the

Table 3 US Equities: Returns, 1926–2012

This table shows beta-sorted calendar-time portfolio returns. At the beginning of each calendar month, stocks are ranked in ascending order on the basis of their estimated beta at the end of the previous month. The ranked stocks are assigned to one of ten deciles portfolios based on NYSE breakpoints. All stocks are equally weighted within a given portfolio, and the portfolios are rebalanced every month to maintain equal weights. The right-most column reports returns of the zero-beta betting against beta (BAB) factor. To construct the BAB factor, all stocks are assigned to one of two portfolios: low beta and high beta. Stocks are weighted by the ranked betas (lower beta security have larger weight in the low-beta portfolio and higher beta securities have larger weights in the high-beta portfolio), and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The betting against beta factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. This table includes all available common stocks on the Center for Research in Security Prices database between January 1926 and March 2012. Alpha is the intercept in a regression of monthly excess return. The explanatory variables are the monthly returns from Fama and French (1993) mimicking portfolios, Carhart (1997) momentum factor and Pastor and Stambaugh (2003) liquidity factor. CAPM = Capital Asset Pricing Model. Regarding the five-factor alphas the Pastor and Stambaugh (2003) liquidity factor is available only between 1968 and 2011. Returns and alphas are in monthly percent, *t*-statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Beta (ex ante) is the average estimated beta at portfolio formation. Beta (realized) is the realized loading on the market portfolio. Volatilities and Sharpe ratios are annualized.

Portfolio	P1 (low beta)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (high beta)	BAB
Excess return	**0.91**	**0.98**	**1.00**	**1.03**	**1.05**	**1.10**	**1.05**	**1.08**	**1.06**	**0.97**	**0.70**
	(6.37)	(5.73)	(5.16)	(4.88)	(4.49)	(4.37)	(3.84)	(3.74)	(3.27)	(2.55)	(7.12)
CAPM alpha	**0.52**	**0.48**	**0.42**	**0.39**	**0.34**	**0.34**	0.22	0.21	0.10	−0.10	**0.73**
	(6.30)	(5.99)	(4.91)	(4.43)	(3.51)	(3.20)	(1.94)	(1.72)	(0.67)	(−0.48)	(7.44)
Three-factor alpha	**0.40**	**0.35**	**0.26**	**0.21**	**0.13**	0.11	−0.03	−0.06	**−0.22**	**−0.49**	**0.73**
	(6.25)	(5.95)	(4.76)	(4.13)	(2.49)	(1.94)	(−0.59)	(−1.02)	(−2.81)	(−3.68)	(7.39)
Four-factor alpha	**0.40**	**0.37**	**0.30**	**0.25**	**0.18**	**0.20**	0.09	0.11	0.01	−0.13	**0.55**
	(6.05)	(6.13)	(5.36)	(4.92)	(3.27)	(3.63)	(1.63)	(1.94)	(0.12)	(−1.01)	(5.59)
Five-factor alpha	**0.37**	**0.37**	**0.33**	**0.30**	**0.17**	**0.20**	0.11	0.14	0.02	0.00	**0.55**
	(4.54)	(4.66)	(4.50)	(4.40)	(2.44)	(2.71)	(1.40)	(1.65)	(0.21)	(−0.01)	(4.09)
Beta (ex ante)	0.64	0.79	0.88	0.97	1.05	1.12	1.21	1.31	1.44	1.70	0.00
Beta (realized)	0.67	0.87	1.00	1.10	1.22	1.32	1.42	1.51	1.66	1.85	−0.06
Volatility	15.70	18.70	21.11	23.10	25.56	27.58	29.81	31.58	35.52	41.68	10.75
Sharpe ratio	0.70	0.63	0.57	0.54	0.49	0.48	0.42	0.41	0.36	0.28	0.78

Table 4 International Equities: Returns, 1984–2012

This table shows beta-sorted calendar-time portfolio returns. At the beginning of each calendar month, stocks are ranked in ascending order on the basis of their estimated beta at the end of the previous month. The ranked stocks are assigned to one of ten deciles portfolios. All stocks are equally weighted within a given portfolio, and the portfolios are rebalanced every month to maintain equal weights. The rightmost column reports returns of the zero-beta betting against beta (BAB) factor. To construct the BAB factor, all stocks are assigned to one of two portfolios: low beta and high beta. The low- (high-) beta portfolio is composed of all stocks with a beta below (above) its country median. Stocks are weighted by the ranked betas (lower beta security have larger weight in the low-beta portfolio and higher beta securities have larger weights in the high-beta portfolio), and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The betting against beta factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. This table includes all available common stocks on the Xpressfeed Global database for the 19 markets listed in Table 1. The sample period runs from January 1984 to March 2012. Alpha is the intercept in a regression of monthly excess return. The explanatory variables are the monthly returns of Asness and Frazzini (2013) mimicking portfolios and Pastor and Stambaugh (2003) liquidity factor. CAPM = Capital Asset Pricing Model. Regarding the five-factor alphas the Pastor and Stambaugh (2003) liquidity factor is available only between 1968 and 2011. Returns are in US dollars and do not include any currency hedging. Returns and alphas are in monthly percent, *t*-statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Beta (ex-ante) is the average estimated beta at portfolio formation. Beta (realized) is the realized loading on the market portfolio. Volatilities and Sharpe ratios are annualized.

Portfolio	P1 (low beta)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (high beta)	BAB
Excess return	**0.63**	**0.67**	**0.69**	**0.58**	**0.67**	0.63	0.54	0.59	0.44	0.30	**0.64**
	(2.48)	(2.44)	(2.39)	(1.96)	(2.19)	(1.93)	(1.57)	(1.58)	(1.10)	(0.66)	(4.66)
CAPM alpha	**0.45**	**0.47**	**0.48**	**0.36**	**0.44**	**0.39**	0.28	0.32	0.15	0.00	**0.64**
	(2.91)	(3.03)	(2.96)	(2.38)	(2.86)	(2.26)	(1.60)	(1.55)	(0.67)	(–0.01)	(4.68)
Three-factor alpha	**0.28**	**0.30**	**0.29**	0.16	0.22	0.11	0.01	–0.03	–0.23	–0.50	**0.65**
	(2.19)	(2.22)	(2.15)	(1.29)	(1.71)	(0.78)	(0.06)	(–0.17)	(–1.20)	(–1.94)	(4.81)
Four-factor alpha	0.20	0.24	0.20	0.10	0.19	0.08	0.04	0.06	–0.16	–0.16	**0.30**
	(1.42)	(1.64)	(1.39)	(0.74)	(1.36)	(0.53)	(0.27)	(0.35)	(–0.79)	(–0.59)	(2.20)
Five-factor alpha	0.19	0.23	0.19	0.09	0.20	0.07	0.05	0.05	–0.19	–0.18	**0.28**
	(1.38)	(1.59)	(1.30)	(0.65)	(1.40)	(0.42)	(0.33)	(0.30)	(–0.92)	(–0.65)	(2.09)
Beta (ex ante)	0.61	0.70	0.77	0.83	0.88	0.93	0.99	1.06	1.15	1.35	0.00
Beta (realized)	0.66	0.75	0.78	0.85	0.87	0.92	0.98	1.03	1.09	1.16	–0.02
Volatility	14.97	16.27	17.04	17.57	18.08	19.42	20.42	22.05	23.91	27.12	8.07
Sharpe ratio	0.50	0.50	0.48	0.40	0.44	0.39	0.32	0.32	0.22	0.13	0.95

Table 5 International Equities: Returns by Country, 1984–2012
This table shows calendar-time portfolio returns. At the beginning of each calendar month, all stocks are assigned to one of two portfolios: low beta and high beta. The low- (high-) beta portfolio is composed of all stocks with a beta below (above) its country median. Stocks are weighted by the ranked betas, and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The zero-beta betting against beta (BAB) factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. This table includes all available common stocks on the Xpressfeed Global database for the 19 markets listed in Table 1. The sample period runs from January 1984 to March 2012. Alpha is the intercept in a regression of monthly excess return. The explanatory variables are the monthly returns of Asness and Frazzini (2013) mimicking portfolios. Returns are in US dollars and do not include any currency hedging. Returns and alphas are in monthly percent, and 5% statistical significance is indicated in bold. $Short (Long) is the average dollar value of the short (long) position. Volatilities and Sharpe ratios are annualized.

Country	Excess return	*t*-Statistics Excess return	Four-factor alpha	*t*-Statistics alpha	$Short	$Long	Volatility	Sharpe ratio
Australia	0.11	0.36	0.03	0.10	0.80	1.26	16.7	0.08
Austria	−0.03	−0.09	−0.28	−0.72	0.90	1.44	19.9	−0.02
Belgium	**0.71**	2.39	**0.72**	2.28	0.94	1.46	16.9	0.51
Canada	**1.23**	5.17	**0.67**	2.71	0.85	1.45	14.1	1.05
Switzerland	**0.75**	2.91	**0.54**	2.07	0.93	1.47	14.6	0.61
Germany	0.40	1.30	−0.07	−0.22	0.94	1.58	17.3	0.27
Denmark	0.41	1.47	−0.02	−0.07	0.91	1.40	15.7	0.31
Spain	**0.59**	2.12	0.23	0.80	0.92	1.44	15.6	0.45
Finland	0.65	1.51	−0.10	−0.22	1.08	1.64	24.0	0.33
France	0.26	0.63	−0.37	−0.82	0.92	1.57	23.7	0.13
United Kingdom	**0.49**	1.99	−0.01	−0.05	0.91	1.53	13.9	0.42
Hong Kong	**0.85**	2.50	**1.01**	2.79	0.83	1.38	19.1	0.54
Italy	0.29	1.41	0.04	0.17	0.91	1.35	11.8	0.30
Japan	0.21	0.90	0.01	0.06	0.87	1.39	13.3	0.19
Netherlands	**0.98**	3.62	**0.79**	2.75	0.91	1.45	15.4	0.77
Norway	0.44	1.15	0.34	0.81	0.85	1.33	21.3	0.25
New Zealand	**0.74**	2.28	0.62	1.72	0.94	1.36	18.1	0.49
Singapore	**0.66**	3.37	**0.52**	2.36	0.79	1.24	11.0	0.72
Sweden	**0.77**	2.29	0.22	0.64	0.89	1.34	19.0	0.48

alpha of the long-short portfolio is consistent across regressions, the choice of risk adjustment influences the relative alpha contribution of the long and short sides of the portfolio.

Our results for US equities show how the security market line has continued to be too flat for another four decades after Black, Jensen, and Scholes (1972). Further, our results extend internationally. We consider beta-sorted portfolios for international equities and later turn to altogether different asset classes. We use all 19 MSCI developed countries except the US (to keep the results separate from the US results above), and we do this in two ways: We consider international portfolios in which all international stocks are pooled

together (Table 4), and we consider results separately for each country (Table 5). The international portfolio is country-neutral, i.e., the low-(high)-beta portfolio is composed of all stocks with a beta below (above) its country median.[17]

The results for our pooled sample of international equities in Table 4 mimic the US results. The alpha and Sharpe ratios of the beta-sorted portfolios decline (although not perfectly monotonically) with the betas, and the BAB factor earns risk-adjusted returns between 0.28% and 0.64% per month depending on the choice of risk adjustment, with *t*-statistics ranging from 2.09 to 4.81.

Table 5 shows the performance of the BAB factor within each individual country. The BAB delivers positive Sharpe ratios in 18 of the 19 MSCI developed countries and positive four-factor alphas in 13 out of 19, displaying a strikingly consistent pattern across equity markets. The BAB returns are statistically significantly positive in six countries, while none of the negative alphas are significant. Of course, the small number of stocks in our sample in many of the countries makes it difficult to reject the null hypothesis of zero return in each individual country.

Table B1 in Appendix B reports factor loadings. On average, the US BAB factor goes long \$1.40 (\$1.40 for international BAB) and shortsells \$0.70 (\$0.89 for international BAB). The larger long investment is meant to make the BAB factor market-neutral because the stocks that are held long have lower betas. The BAB factor's realized market loading is not exactly zero, reflecting the fact that our ex ante betas are measured with noise. The other factor loadings indicate that, relative to high-beta stocks, low-beta stocks are likely to be larger, have higher book-to-market ratios, and have higher return over the prior 12 months, although none of the loadings can explain the large and significant abnormal returns. The BAB portfolio's positive HML loading is natural since our theory predicts that low-beta stocks are cheap and high-beta stocks are expensive.

Appendix B reports further tests and additional robustness checks. In Table B2, we report results using different window lengths to estimate betas and different benchmarks (local, global). We split the sample by size (Table B3) and time periods (Table B4), we control for idiosyncratic volatility (Table B5), and we report results for alternative definitions of the risk-free rate (Table B6). Finally, in Table B7 and Fig. B2 we report an out-of-sample test. We collect pricing data from DataStream and for each country in Table 1 we compute a BAB portfolio over sample period not covered by the Xpressfeed Global database.[18] All of the results are consistent: Equity portfolios that bet against betas earn significant risk-adjusted returns.

4.2. Treasury Bonds

Table 6 reports results for US Treasury bonds. As before, we report average excess returns of bond portfolios formed by sorting on beta in the previous month. In the cross section of Treasury bonds, ranking on betas with respect to an aggregate Treasury bond index is empirically equivalent to ranking on duration or maturity. Therefore, in Table 6, one can think of the term "beta," "duration," or "maturity" in an interchangeable fashion. The right-most column reports returns of the BAB factor. Abnormal returns are computed with respect to a one-factor model in which alpha is the intercept in a regression of monthly excess return on an equally weighted Treasury bond excess market return.

The results show that the phenomenon of a flatter security market line than predicted by the standard CAPM is not limited to the cross section of stock returns. Consistent with

Table 6 US Treasury Bonds: Returns, 1952–2012

This table shows calendar-time portfolio returns. The test assets are the Center for Research in Security Prices Treasury Fama bond portfolios. Only non-callable, non-flower notes and bonds are included in the portfolios. The portfolio returns are an equal-weighted average of the unadjusted holding period return for each bond in the portfolios in excess of the risk-free rate. To construct the zero-beta betting against beta (BAB) factor, all bonds are assigned to one of two portfolios: low beta and high beta. Bonds are weighted by the ranked betas (lower beta bonds have larger weight in the low-beta portfolio and higher beta bonds have larger weights in the high-beta portfolio) and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The BAB factor is a self-financing portfolio that is long the low-beta portfolio and shorts the high-beta portfolio. Alpha is the intercept in a regression of monthly excess return. The explanatory variable is the monthly return of an equally weighted bond market portfolio. The sample period runs from January 1952 to March 2012. Returns and alphas are in monthly percent, *t*-statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Volatilities and Sharpe ratios are annualized. For P7, returns are missing from August 1962 to December 1971.

Portfolio	**P1 (low beta)**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7 (high beta)**	**BAB**
Maturity (months)	one to 12	13–24	25–36	37–48	49–60	61–120	>120	
Excess return	**0.05**	**0.09**	**0.11**	**0.13**	**0.13**	**0.16**	**0.24**	**0.17**
	(5.66)	(3.91)	(3.37)	(3.09)	(2.62)	(2.52)	(2.20)	(6.26)
Alpha	**0.03**	**0.03**	0.02	0.01	−0.01	**−0.02**	−0.07	**0.16**
	(5.50)	(3.00)	(1.87)	(0.99)	(−1.35)	(−2.28)	(−1.85)	(6.18)
Beta (ex ante)	0.14	0.45	0.74	0.98	1.21	1.44	2.24	0.00
Beta (realized)	0.16	0.48	0.76	0.98	1.17	1.44	2.10	0.01
Volatility	0.81	2.07	3.18	3.99	4.72	5.80	9.26	2.43
Sharpe ratio	0.73	0.50	0.43	0.40	0.34	0.32	0.31	0.81

Proposition 1, the alphas decline monotonically with beta. Likewise, Sharpe ratios decline monotonically from 0.73 for low-beta (short-maturity) bonds to 0.31 for high-beta (long-maturity) bonds. Furthermore, the bond BAB portfolio delivers abnormal returns of 0.17% per month (*t*-statistic = 6.26) with a large annual Sharpe ratio of 0.81.

Because the idea that funding constraints have a significant effect on the term structure of interest could be surprising, let us illustrate the economic mechanism that could be at work. Suppose an agent, e.g., a pension fund, has $1 to allocate to Treasuries with a target excess return of 2.9% per year. One way to achieve this return target is to invest $1 in a portfolio of Treasuries with maturity above ten years as seen in Table 6, P7. If the agent invests in one-year Treasuries (P1) instead, then he would need to invest $11 if all maturities had the same Sharpe ratio. This higher leverage is needed because the long-term Treasures are 11 times more volatile than the short-term Treasuries. Hence, the agent would need to borrow an additional $10 to lever his investment in one-year bonds. If the agent has leverage limits (or prefers lower leverage), then he would strictly prefer the ten-year Treasuries in this case.

According to our theory, the one-year Treasuries therefore must offer higher returns and higher Sharpe ratios, flattening the security market line for bonds. Empirically, short-term Treasuries do offer higher risk-adjusted returns so the return target can be achieved

by investing about $5 in one-year bonds. While a constrained investor could still prefer an un-leveraged investment in ten-year bonds, unconstrained investors now prefer the leveraged low-beta bonds, and the market can clear.

While the severity of leverage constraints varies across market participants, it appears plausible that a five-to-one leverage (on this part of the portfolio) makes a difference for some large investors such as pension funds.

4.3. Credit

We next test our model using several credit portfolios and report results in Table 7. In Panel A, columns 1 to 5, the test assets are monthly excess returns of corporate bond indexes by maturity. We see that the credit BAB portfolio delivers abnormal returns of 0.11% per month (t-statistic = 5.14) with a large annual Sharpe ratio of 0.82. Furthermore, alphas and Sharpe ratios decline monotonically.

In columns 6 to 10, we attempt to isolate the credit component by hedging away the interest rate risk. Given the results on Treasuries in Table 6, we are interested in testing a pure credit version of the BAB portfolio. Each calendar month, we run one-year rolling regressions of excess bond returns on the excess return on Barclay's US government bond index. We construct test assets by going long the corporate bond index and hedging this position by shortselling the appropriate amount of the government bond index: $r_t^{CDS} - r_t^f = \left(r_t - r_t^f\right) - \hat{\theta}_{t-1}\left(r_t^{USGOV} - r_t^f\right)$, where $\hat{\theta}_{t-1}$ is the slope coefficient estimated in an expanding regression using data from the beginning of the sample and up to month $t-1$. One interpretation of this returns series is that it approximates the returns on a credit default swap (CDS). We compute market returns by taking the equally weighted average of these hedged returns, and we compute betas and BAB portfolios as before. Abnormal returns are computed with respect to a two-factor model in which alpha is the intercept in a regression of monthly excess return on the equally weighted average pseudo-CDS excess return and the monthly return on the Treasury BAB factor. The addition of the Treasury BAB factor on the right-hand side is an extra check to test a pure credit version of the BAB portfolio.

The results in Panel A of Table 7 columns 6 to 10 tell the same story as columns 1 to 5: The BAB portfolio delivers significant abnormal returns of 0.17% per month (t-statistic = 4.44) and Sharpe ratios decline monotonically from low-beta to high-beta assets.

Last, in Panel B of Table 7, we report results in which the test assets are credit indexes sorted by rating, ranging from AAA to Ca-D and Distressed. Consistent with all our previous results, we find large abnormal returns of the BAB portfolios (0.57% per month with t-statistic = 3.72) and declining alphas and Sharpe ratios across beta-sorted portfolios.

4.4. Equity Indexes, Country Bond Indexes, Currencies, and Commodities

Table 8 reports results for equity indexes, country bond indexes, foreign exchange, and commodities. The BAB portfolio delivers positive returns in each of the four asset classes, with an annualized Sharpe ratio ranging from 0.11 to 0.51. We are able to reject the null hypothesis of

Table 7 US Credit: Returns, 1973–2012

This table shows calendar-time portfolio returns. Panel A shows results for US credit indices by maturity. The test assets are monthly returns on corporate bond indices with maturity ranging from one to ten years, in excess of the risk-free rate. The sample period runs from January 1976–March 2012. Unhedged indicates excess returns and Hedged indicates excess returns after hedging the index's interest rate exposure. To construct hedged excess returns, each calendar month we run one-year rolling regressions of excess bond returns on the excess return on Barclay's US government bond index. We construct test assets by going long the corporate bond index and hedging this position by shorting the appropriate amount of the government bond index. We compute market excess returns by taking an equal weighted average of the hedged excess returns. Panel B shows results for US corporate bond index returns by rating. The sample period runs from January 1973 to March 2012. To construct the zero-beta betting against beta (BAB) factor, all bonds are assigned to one of two portfolios: low beta and high beta. Bonds are weighted by the ranked betas (lower beta securities have larger weight in the low-beta portfolio and higher beta securities have larger weights in the high-beta portfolio) and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of 1 at portfolio formation. The zero-beta BAB factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. Alpha is the intercept in a regression of monthly excess return. The explanatory variable is the monthly excess return of the corresponding market portfolio and, for the hedged portfolios in Panel A, the Treasury BAB factor. Distressed in Panel B indicates the Credit Suisse First Boston distressed index. Returns and alphas are in monthly percent, *t*-statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Volatilities and Sharpe ratios are annualized.

Panel A: Credit indices, 1976–2012

	Unhedged					Hedged				
Portfolios	**One to three years**	**Three to five years**	**Five to ten years**	**Seven to ten years**	**BAB**	**One to three years**	**Three to five years**	**Five to ten years**	**Seven to ten years**	**BAB**
Excess return	**0.18**	**0.22**	**0.36**	**0.36**	**0.10**	**0.11**	**0.10**	0.11	0.10	**0.16**
	(4.97)	(4.35)	(3.35)	(3.51)	(4.85)	(3.39)	(2.56)	(1.55)	(1.34)	(4.35)
Alpha	**0.03**	0.01	**−0.04**	**−0.07**	**0.11**	**0.05**	**0.03**	**−0.03**	**−0.05**	**0.17**
	(2.49)	(0.69)	(−3.80)	(−4.28)	(5.14)	(3.89)	(2.43)	(−3.22)	(−3.20)	(4.44)
Beta (ex ante)	0.71	1.02	1.59	1.75	0.00	0.54	0.76	1.48	1.57	0.00
Beta (realized)	0.61	0.85	1.38	1.49	−0.03	0.53	0.70	1.35	1.42	−0.02
Volatility	2.67	3.59	5.82	6.06	1.45	1.68	2.11	3.90	4.15	1.87
Sharpe ratio	0.83	0.72	0.74	0.72	0.82	0.77	0.58	0.35	0.30	1.02

(*Continued*)

Table 7 (***Continued***)

Panel B: Corporate bonds, 1973–2012										
Portfolios	**Aaa**	**Aa**	**A**	**Baa**	**Ba**	**B**	**Caa**	**Ca-D**	**Distressed**	**BAB**
Excess return	**0.28**	**0.31**	**0.32**	**0.37**	**0.47**	**0.38**	0.35	0.77	–0.41	**0.44**
	(3.85)	(3.87)	(3.47)	(3.93)	(4.20)	(2.56)	(1.47)	(1.42)	(–1.06)	(2.64)
Alpha	**0.23**	**0.23**	**0.20**	**0.23**	**0.27**	0.10	−0.06	−0.04	**−1.11**	**0.57**
	(3.31)	(3.20)	(2.70)	(3.37)	(4.39)	(1.39)	(−0.40)	(−0.15)	(−5.47)	(3.72)
Beta (ex ante)	0.67	0.72	0.79	0.88	0.99	1.11	1.57	2.22	2.24	0.00
Beta (realized)	0.17	0.29	0.41	0.48	0.67	0.91	1.34	2.69	2.32	−0.47
Volatility	4.50	4.99	5.63	5.78	6.84	9.04	14.48	28.58	23.50	9.98
Sharpe ratio	0.75	0.75	0.68	0.77	0.82	0.50	0.29	0.32	−0.21	0.53

zero average return only for equity indexes, but we can reject the null hypothesis of zero returns for combination portfolios that include all or some combination of the four asset classes, taking advantage of diversification. We construct a simple equally weighted BAB portfolio. To account for different volatility across the four asset classes, in month t we rescale each return series to 10% annualized volatility using rolling three-year estimates up to month $t-1$ and then we equally weight the return series and their respective market benchmark. This portfolio construction generates a simple implementable portfolio that targets 10% BAB volatility in each of the asset classes. We report results for an all futures combo including all four asset classes and a country selection combo including only equity indices, country bonds and foreign exchange. The BAB all futures and country selection deliver abnormal return of 0.25% and 0.26% per month (t-statistics = 2.53 and 2.42).

4.5. Betting against All of the Betas

To summarize, the results in Tables 3–8 strongly support the predictions that alphas decline with beta and BAB factors earn positive excess returns in each asset class. Fig. 1 illustrates the remarkably consistent pattern of declining alphas in each asset class, and Fig. 2 shows the consistent return to the BAB factors. Clearly, the relatively flat security market line, shown by Black, Jensen, and Scholes (1972) for US stocks, is a pervasive phenomenon that we find across markets and asset classes. Averaging all of the BAB factors produces a diversified BAB factor with a large and significant abnormal return of 0.54% per month (t-statistic of 6.98) as seen in Table 8, Panel B.

5. Time Series Tests

In this section, we test Proposition 3's predictions for the time series of BAB returns: When funding constraints become more binding (e.g., because margin requirements rise), the

Table 8 Equity Indices, Country Bonds, Foreign Exchange and Commodities: Returns, 1965–2012

This table shows calendar-time portfolio returns. The test assets are futures, forwards or swap returns in excess of the relevant financing rate. To construct the betting against beta (BAB) factor, all securities are assigned to one of two portfolios: low beta and high beta. Securities are weighted by the ranked betas (lower beta security have larger weight in the low-beta portfolio and higher beta securities have larger weights in the high-beta portfolio), and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The BAB factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. Alpha is the intercept in a regression of monthly excess return. The explanatory variable is the monthly return of the relevant market portfolio. Panel A reports results for equity indices, country bonds, foreign exchange and commodities. All futures and Country selection are combo portfolios with equal risk in each individual BAB and 10% ex ante volatility. To construct combo portfolios, at the beginning of each calendar month, we rescale each return series to 10% annualized volatility using rolling three-year estimate up to moth $t - 1$ and then equally weight the return series and their respective market benchmark. Panel B reports results for all the assets listed in Tables 1 and 2. All bonds and credit includes US Treasury bonds, US corporate bonds, US credit indices (hedged and unhedged) and country bonds indices. All equities includes US equities, all individual BAB country portfolios, the international stock BAB, and the equity index BAB. All assets includes all the assets listed in Tables 1 and 2. All portfolios in Panel B have equal risk in each individual BAB and 10% ex ante volatility. Returns and alphas are in monthly percent, *t*-statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. $Short (Long) is the average dollar value of the short (long) position. Volatilities and Sharpe ratios are annualized. *Denotes equal risk, 10% ex ante volatility.

BAB portfolios	Excess return	*t*-Statistics excess return	Alpha	*t*-Statistics alpha	$Short	$Long	Volatility	Sharpe ratio
Panel A: Equity indices, country bonds, foreign exchange and commodities								
Equity indices (EI)	**0.55**	2.93	**0.48**	2.58	0.86	1.29	13.08	0.51
Country bonds (CB)	0.03	0.67	0.05	0.95	0.88	1.48	2.93	0.14
Foreign exchange (FX)	0.17	1.23	0.19	1.42	0.89	1.59	9.59	0.22
Commodities (COM)	0.18	0.72	0.21	0.83	0.71	1.48	19.67	0.11
All futures (EI + CB + FX + COM)*	**0.26**	2.62	**0.25**	2.52			7.73	0.40
Country selection (EI + CB + FX)*	**0.26**	2.38	**0.26**	2.42			7.47	0.41
Panel B: All assets								
All bonds and credit*	**0.74**	6.94	**0.71**	6.74			9.78	0.90
All equities*	**0.63**	6.68	**0.64**	6.73			10.36	0.73
All assets*	**0.53**	6.89	**0.54**	6.98			8.39	0.76

required future BAB premium increases, and the contemporaneous realized BAB returns become negative.

We take this prediction to the data using the TED spread as a proxy of funding conditions. The sample runs from December 1984 (the first available date for the TED spread) to March 2012.

Table 9 reports regression-based tests of our hypotheses for the BAB factors across asset classes. The first column simply regresses the US BAB factor on the lagged level

Table 9 Regression Results

This table shows results from (pooled) time series regressions. The left-hand side is the month t return of the betting against beta (BAB) factors. To construct the BAB portfolios, all securities are assigned to one of two portfolios: low beta and high beta. Securities are weighted by the ranked betas (lower beta securities have larger weight in the low-beta portfolio and higher beta securities have larger weights in the high-beta portfolio), and the portfolios are rebalanced every calendar month. Both portfolios are rescaled to have a beta of one at portfolio formation. The BAB factor is a self-financing portfolio that is long the low-beta portfolio and short the high-beta portfolio. The explanatory variables include the TED spread and a series of controls. Lagged TED spread is the TED spread at the end of month $t-1$. Change in TED spread is equal to TED spread at the end of month t minus Ted spread at the end of month $t-1$. Short volatility return is the month t return on a portfolio that shorts at-the-money straddles on the S&P 500 index. To construct the short volatility portfolio, on index options expiration dates we write the next-to-expire closest-to-maturity straddle on the S&P 500 index and hold it to maturity. Beta spread is defined as (HBeta − LBeta)/(HBeta * LBeta) where HBeta (LBeta) are the betas of the short (long) leg of the BAB portfolio at portfolio formation. Market return is the monthly return of the relevant market portfolio. Lagged inflation is equal to the one-year US Consumer Price Index inflation rate, lagged one month. The data run from December 1984 (first available date for the TED spread) to March 2012. Columns 1 and 2 report results for US equities. Columns 3 and 4 report results for international equities. In these regressions we use each individual country BAB factor as well as an international equity BAB factor. Columns 5 and 6 report results for all assets in our data. Asset fixed effects are included where indicated, t-statistics are shown below the coefficient estimates and all standard errors are adjusted for heteroskedasticity (White, 1980). When multiple assets are included in the regression, standard errors are clustered by date and 5% statistical significance is indicated in bold.

	US equities		**International equities, pooled**		**All assets, pooled**	
Left-hand side: BAB return	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Lagged TED spread	**−0.025**	**−0.038**	**−0.009**	**−0.015**	**−0.013**	**−0.018**
	(−5.24)	(−4.78)	(−3.87)	(−4.07)	(−4.87)	(−4.65)
Change in TED spread	**−0.019**	**−0.035**	**−0.006**	**−0.010**	**−0.007**	**−0.011**
	(−2.58)	(−4.28)	(−2.24)	(−2.73)	(−2.42)	(−2.64)
Beta spread		0.011		0.001		0.001
		(0.76)		(0.40)		(0.69)
Lagged BAB return		0.011		0.035		0.044
		(0.13)		(1.10)		(1.40)
Lagged inflation		−0.177		0.003		−0.062
		(−0.87)		(0.03)		(−0.58)
Short volatility return		**−0.238**		0.021		0.027
		(−2.27)		(0.44)		(0.48)

(*Continued*)

Table 9 (*Continued*)

	US equities		International equities, pooled		All assets, pooled	
Left-hand side: BAB return	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Market return		**−0.372**		**−0.104**		**−0.097**
		(−4.40)		(−2.27)		(−2.18)
Asset fixed effects	No	No	Yes	Yes	Yes	Yes
Number of observations	328	328	5,725	5,725	8,120	8,120
Adjusted R^2	0.070	0.214	0.007	0.027	0.014	0.036

of the TED spread and the contemporaneous change in the TED spread.[19] We see that both the lagged level and the contemporaneous change in the TED spread are negatively related to the BAB returns. If the TED spread measures the tightness of funding constraints (given by ψ in the model), then the model predicts a negative coefficient for the contemporaneous change in TED [Eq. (11)] and a positive coefficient for the lagged level [Eq. (12)]. Hence, the coefficient for change is consistent with the model, but the coefficient for the lagged level is not, under this interpretation of the TED spread. If, instead, a high TED spread indicates that agents' funding constraints are worsening, then the results would be easier to understand. Under this interpretation, a high TED spread could indicate that banks are credit-constrained and that banks tighten other investors' credit constraints over time, leading to a deterioration of BAB returns over time (if investors do not foresee this).

However, the model's prediction as a partial derivative assumes that the current funding conditions change while everything else remains unchanged, but, empirically, other things do change. Hence, our test relies on an assumption that such variation of other variables does not lead to an omitted variables bias. To partially address this issue, column 2 provides a similar result when controlling for a number of other variables. The control variables are the market return (to account for possible noise in the ex ante betas used for making the BAB portfolio market neutral), the one-month lagged BAB return (to account for possible momentum in BAB), the ex ante beta spread, the short volatility returns, and the lagged inflation. The beta spread is equal to $(\beta_S - \beta_L)/(\beta_S\beta_L)$ and measures the ex ante beta difference between the long and short side of the BAB portfolios, which should positively predict the BAB return as seen in Proposition 2. Consistent with the model, Table 9 shows that the estimated coefficient for the beta spread is positive in all specifications, but not statistically significant. The short volatility returns is the return on a portfolio that shortsells closest-to-the-money, next-to-expire straddles on the S&P 500 index, capturing potential sensitivity to volatility risk. Lagged inflation is equal to the one-year US CPI inflation rate, lagged one month, which is included to account for potential effects of money illusion as studied by Cohen, Polk, and Vuolteenaho (2005), although we do not find evidence of this effect.

Columns 3–4 of Table 9 report panel regressions for international stock BAB factors and columns 5–6 for all the BAB factors. These regressions include fixed effects and standard errors are clustered by date. We consistently find a negative relation between BAB returns and the TED spread.

6. Beta Compression

We next test Proposition 4 that betas are compressed toward one when funding liquidity risk is high. Table 10 presents tests of this prediction. We use the volatility of the TED spread to proxy for the volatility of margin requirements. Volatility in month t is defined as the standard deviation of daily TED spread innovations, $\sigma_t^{TED} = \sqrt{\sum_{s \in month\ t} \left(\Delta TED_s - \overline{\Delta TED_t}\right)^2}$. Because we are computing moments, we use the monthly volatility as of the prior calendar month, which ensures that the conditioning variable is known at the beginning of the measurement period. The sample runs from December 1984–March 2012.

Panel A of Table 10 shows the cross-sectional dispersion in betas in different time periods sorted by the TED volatility for US stocks, Panel B shows the same for international stocks, and Panel C shows this for all asset classes in our sample. Each calendar month, we compute cross-sectional standard deviation, mean absolute deviation, and inter-quintile range of the betas for all assets in the universe. We assign the TED spread volatility into three groups (low, medium, and high) based on full sample breakpoints (top and bottom third) and regress the times series of the cross-sectional dispersion measure on the full set of dummies (without intercept). In Panel C, we compute the monthly dispersion measure in each asset class and average across assets. All standard errors are adjusted for heteroskedasticity and autocorrelation up to 60 months.

Table 10 shows that, consistent with Proposition 4, the cross-sectional dispersion in betas is lower when credit constraints are more volatile. The average cross-sectional standard deviation of US equity betas in periods of low spread volatility is 0.34, and the dispersion shrinks to 0.29 in volatile credit environment. The difference is statistically significant (t-statistic = 2.71). The tests based on the other dispersion measures, the international equities, and the other assets all confirm that the cross-sectional dispersion in beta shrinks at times when credit constraints are more volatile.

Appendix B contains an additional robustness check. Because we are looking at the cross-sectional dispersion of estimated betas, one could worry that our results were driven by higher beta estimation errors, instead of a higher variance of the true betas. To investigate this possibility, we run simulations under the null hypothesis of a constant standard deviation of true betas and test whether the measurement error in betas can account for the compression observed in the data. Fig. B3 shows that the compression observed in the data is much larger than what could be generated by estimation error variance alone. Naturally, while this bootstrap analysis does not indicate that the beta compression observed in Table 10 is likely due to measurement error, we cannot rule out all types of measurement error.

Panels D, E, and F report conditional market betas of the BAB portfolio returns based on the volatility of the credit environment for US equities, international equities, and the average BAB factor across all assets, respectively. The dependent variable is the monthly return of the BAB portfolio. The explanatory variables are the monthly returns of the market portfolio, Fama and French (1993) mimicking portfolios, and Carhart (1997) momentum factor. Market betas are allowed to vary across TED volatility regimes (low, neutral, and high) using the full set of TED dummies.

We are interested in testing Proposition 4(ii), studying how the BAB factor's conditional beta depends on the TED-volatility environment. To understand this test, recall first that the BAB factor is market neutral conditional on the information set used in the

Table 10 Beta Compression

This table reports results of cross-sectional and time-series tests of beta compression. Panels A, B and C report cross-sectional dispersion of betas in US equities, international equities, and all asset classes in our sample. The data run from December 1984 (first available date for the TED spread) to March 2012. Each calendar month we compute cross sectional standard deviation, mean absolute deviation, and inter quintile range of betas. In Panel C we compute each dispersion's measure for each asset class and average across asset classes. The row denoted All reports times series means of the dispersion measures. P1 to P3 report coefficients on a regression of the dispersion measure on a series of TED spread volatility dummies. TED spread volatility is defined as the standard deviation of daily changes in the TED spread in the prior calendar month. We assign the TED spread volatility into three groups (low, neutral, and high) based on full sample breakpoints (top and bottom one third) and regress the times series of the cross-sectional dispersion measure on the full set of dummies (without intercept). *t*-Statistics are shown below the coefficient estimates, and 5% statistical significance is indicated in bold. Panels D, E and F report conditional market betas of the betting against beta (BAB) portfolio based on TED spread volatility as of the prior month. The dependent variable is the monthly return of the BAB portfolios. The explanatory variables are the monthly returns of the market portfolio, Fama and French (1993), Asness and Frazzini (2013), and Carhart (1997) mimicking portfolios, but only the alpha and the market betas are reported. CAPM indicates the Capital Asset Pricing Model. Market betas are allowed to vary across TED spread volatility regimes (low, neutral, and high) using the full set of dummies. Panels D, E and F report loading on the market factor corresponding to different TED spread volatility regimes. All assets report results for the aggregate BAB portfolio of Table 9, Panel B. All standard errors are adjusted for heteroskedasticity and autocorrelation using a Bartlett kernel (Newey and West, 1987) with a lag length of sixty months.

Cross-sectional dispersion	Standard deviation	Mean absolute deviation	Interquintile range
Panel A: US equities			
All	0.32	0.25	0.43
P1 (low TED volatility)	0.34	0.27	0.45
P2	0.33	0.26	0.44
P3 (high TED volatility)	0.29	0.23	0.40
P3 minus P1	**−0.05**	**−0.04**	−0.05
t-Statistics	(−2.71)	(−2.43)	(−1.66)
Panel B: International equities			
All	0.22	0.17	0.29
P1 (low TED volatility)	0.23	0.18	0.30
P2	0.22	0.17	0.29
P3 (high TED volatility)	0.20	0.16	0.27
P3 minus P1	**−0.04**	**−0.03**	−0.03
t-Statistics	(−2.50)	(−2.10)	(−1.46)
Panel C: All assets			
All	0.45	0.35	0.61
P1 (low TED volatility)	0.47	0.37	0.63
P2	0.45	0.36	0.62
P3 (high TED volatility)	0.43	0.33	0.58
P3 minus P1	**−0.04**	**−0.03**	**−0.06**
t-Statistics	(−3.18)	(−3.77)	(−2.66)

(*Continued*)

Table 10 (*Continued*)

		Conditional market beta			
	Alpha	**P1 (low TED volatility)**	**P2**	**P3 (high TED volatility)**	**P3 − P1**
Panel D: US equities					
CAPM	**1.06**	**−0.46**	−0.19	−0.01	**0.45**
	(3.61)	(−2.65)	(−1.29)	(−0.11)	(3.01)
Control for three factors	**0.86**	**−0.40**	−0.02	0.08	**0.49**
	(4.13)	(−3.95)	(−0.19)	(0.69)	(3.06)
Control for four factors	**0.66**	**−0.28**	0.00	0.13	**0.40**
	(3.14)	(−5.95)	(0.02)	(1.46)	(4.56)
Panel E: International equities					
CAPM	**0.60**	−0.09	0.02	0.06	0.16
	(2.84)	(−1.30)	(0.64)	(1.28)	(1.87)
Control for three factors	**0.59**	−0.09	0.02	0.05	0.14
	(3.23)	(−1.22)	(0.74)	(1.09)	(1.70)
Control for four factors	**0.35**	−0.04	0.05	**0.07**	**0.11**
	(2.16)	(−1.16)	(1.51)	(2.03)	(2.24)
Panel F: All assets					
CAPM	**0.54**	**−0.13**	−0.07	−0.01	**0.14**
	(4.96)	(−2.64)	(−1.82)	(0.21)	(2.34)

estimation of ex ante betas (which determine the ex ante relative position sizes of the long and short sides of the portfolio). Hence, if the TED spread volatility was used in the ex ante beta estimation, then the BAB factor would be market-neutral conditional on this information. However, the BAB factor was constructed using historical betas that do not take into account the effect of the TED spread and, therefore, a high TED spread volatility means that the realized betas will be compressed relative to the ex ante estimated betas used in portfolio construction. Therefore, a high TED spread volatility should increase the conditional market sensitivity of the BAB factor (because the long side of the portfolio is leveraged too much and the short side is deleveraged too much). Indeed, Table 10 shows that when credit constraints are more volatile, the market beta of the BAB factor rises. The right-most column shows that the difference between low- and high-credit volatility environments is statistically significant (t-statistic of 3.01). Controlling for three or four factors yields similar results. The results for our sample of international equities (Panel E) and for the average BAB across all assets (Panel F) are similar, but they are weaker both in terms of magnitude and statistical significance.

Importantly, the alpha of the BAB factor remains large and statistically significant even when we control for the time-varying market exposure. This means that, if we hedge the BAB factor to be market-neutral conditional on the TED spread volatility environment, then this conditionally market-neutral BAB factor continues to earn positive excess returns.

7. Testing the Model's Portfolio Predictions

The theory's last prediction (Proposition 5) is that more-constrained investors hold higher-beta securities than less-constrained investors. Consistent with this prediction, Table 11 presents evidence that mutual funds and individual investors hold high-beta stocks while LBO firms and Berkshire Hathaway buy low-beta stocks.

Before we delve into the details, let us highlight a challenge in testing Proposition 5. Whether an investor's constraint is binding depends both on the investor's ability to apply leverage (m^i in the model) and on its unobservable risk aversion. For example, while a hedge fund could apply some leverage, its leverage constraint could nevertheless be binding if its desired volatility is high (especially if its portfolio is very diversified and hedged).

Given that binding constraints are difficult to observe directly, we seek to identify groups of investors that are plausibly constrained and unconstrained. One example of an investor that could be constrained is a mutual fund. The 1940 Investment Company Act places some restriction on mutual funds' use of leverage, and many mutual funds are prohibited by charter from using leverage. A mutual fund's need to hold cash to meet redemptions ($m^i > 1$ in the model) creates a further incentive to overweight high-beta securities. Overweighting high-beta stocks helps avoid lagging their benchmark in a bull market because of the cash holdings (some funds use futures contracts to "equitize" the cash, but other funds are not allowed to use derivative contracts).

A second class of investors that could face borrowing constraints is individual retail investors. Although we do not have direct evidence of their inability to employ leverage (and some individuals certainly do), we think that (at least in aggregate) it is plausible that they are likely to face borrowing restrictions.

The flipside of this portfolio test is identifying relatively unconstrained investors. Thus, one needs investors that could be allowed to use leverage and are operating below their leverage cap so that their leverage constraints are not binding. We look at the holdings of two groups of investors that could satisfy these criteria as they have access to leverage and focus on long equity investments (requiring less leverage than long/short strategies).

First, we look at the firms that are the target of bids by leveraged buyout (LBO) funds and other forms of private equity. These investors, as the name suggest, employ leverage to acquire a public company. Admittedly, we do not have direct evidence of the maximum leverage available to these LBO firms relative to the leverage they apply, but anecdotal evidence suggests that they achieve a substantial amount of leverage.

Second, we examine the holdings of Berkshire Hathaway, a publicly traded corporation run by Warren Buffett that holds a diversified portfolio of equities and employs leverage (by issuing debt, via insurance float, and other means). The advantage of using the holdings of a public corporation that holds equities such as Berkshire is that we can directly observe its leverage. Over the period from March 1980 to March 2012, its average book leverage, defined as (book equity + total debt) / book equity, was about 1.2, that is, 20% borrowing, and the market leverage including other liabilities such insurance float was about 1.6 (Frazzini, Kabiller, and Pedersen, 2012). It is therefore plausible to assume that Berkshire at the margin could issue more debt but choose not to, making it a likely candidate for an investor whose combination of risk aversion and borrowing constraints made it relatively unconstrained during our sample period.

Table 11 reports the results of our portfolio test. We estimate both the ex ante beta of the various investors' holdings and the realized beta of the time series of their returns.

Table 11 Testing the Model's Portfolio Predictions, 1963–2012

This table shows average ex ante and realized portfolio betas for different groups of investors. Panel A reports results for our sample of open-end actively-managed domestic equity mutual funds as well as results a sample of individual retail investors. Panel B reports results for a sample of leveraged buyouts (private equity) and for Berkshire Hathaway. We compute both the ex ante beta of their holdings and the realized beta of the time series of their returns. To compute the ex ante beta, we aggregate all quarterly (monthly) holdings in the mutual fund (individual investor) sample and compute their ex ante betas (equally weighted and value weighted based on the value of their holdings). We report the time series averages of the portfolio betas. To compute the realized betas, we compute monthly returns of an aggregate portfolio mimicking the holdings, under the assumption of constant weight between reporting dates (quarterly for mutual funds, monthly for individual investors). We compute equally weighted and value-weighted returns based on the value of their holdings. The realized betas are the regression coefficients in a time series regression of these excess returns on the excess returns of the Center for Research in Security Prices value-weighted index. In Panel B we compute ex ante betas as of the month-end prior to the initial takeover announcements date. *t*-Statistics are shown to right of the betas estimates and test the null hypothesis of beta = 1. All standard errors are adjusted for heteroskedasticity and autocorrelation using a Bartlett kernel (Newey and West, 1987) with a lag length of 60 months. A 5% statistical significance is indicated in bold.

		Ex ante beta of positions		**Realized beta of positions**	
Investor, method	**Sample period**	**Beta**	***t*-Statistics (H0: beta=1)**	**Beta**	***t*-Statistics (H0: beta=1)**
Panel A: Investors likely to be constrained					
Mutual funds, value weighted	1980–2012	**1.08**	2.16	**1.08**	6.44
Mutual funds, equal weighted	1980–2012	1.06	1.84	**1.12**	3.29
Individual investors, value weighted	1991–1996	**1.25**	8.16	**1.09**	3.70
Individual investors, equal weighted	1991–1996	**1.25**	7.22	**1.08**	2.13
Panel B: Investors who use leverage					
Private equity (all)	1963–2012	0.96	−1.50		
Private equity (all), equal weighted	1963–2012	**0.94**	−2.30		
Private equity (LBO, MBO), value weighted	1963–2012	**0.83**	−3.15		
Private equity (LBO, MBO), equal weighted	1963–2012	**0.82**	−3.47		
Berkshire Hathaway, value weighted	1980–2012	**0.91**	−2.42	**0.77**	− 3.65
Berkshire Hathaway, equal weighted	1980–2012	**0.90**	−3.81	**0.83**	− 2.44

We first aggregate all holdings for each investor group, compute their ex ante betas (equal and value weighted, respectively), and take the time series average. To compute the realized betas, we compute monthly returns of an aggregate portfolio mimicking the holdings, under the assumption of constant weight between reporting dates. The realized betas are the regression coefficients in a time series regression of these excess returns on the excess returns of the CRSP value-weighted index.

Panel A shows evidence consistent with the hypothesis that constrained investors stretch for return by increasing their betas. Mutual funds hold securities with betas above

one, and we are able to reject the null hypothesis of betas being equal to one. These findings are consistent with those of Karceski (2002), but our sample is much larger, including all funds over 30-year period. Similar evidence is presented for individual retail investors: Individual investors tend to hold securities with betas that are significantly above one.[20]

Panel B reports results for our sample of private equity. For each target stock in our database, we focus on its ex ante beta as of the month-end prior to the initial announcements date. This focus is to avoid confounding effects that result from changes in betas related to the actual delisting event. We consider both the sample of all private equity deals and the subsample that we are able to positively identify as LBO/MBO events. Since we only have partial information about whether each deal is a LBO/MBO, the broad sample includes all types of deals where a company is taken private. The results are consistent with Proposition 5 in that investors executing leverage buyouts tend to acquire (or attempt to acquire in case of a non-successful bid) firms with low betas, and we are able to reject the null hypothesis of a unit beta.

The results for Berkshire Hathaway show a similar pattern: Warren Buffett bets against beta by buying stocks with betas significantly below one and applying leverage.

8. Conclusion

All real-world investors face funding constraints such as leverage constraints and margin requirements, and these constraints influence investors' required returns across securities and over time. We find empirically that portfolios of high-beta assets have lower alphas and Sharpe ratios than portfolios of low-beta assets. The security market line is not only flatter than predicted by the standard CAPM for US equities (as reported by Black, Jensen, and Scholes (1972)), but we also find this relative flatness in 18 of 19 international equity markets, in Treasury markets, for corporate bonds sorted by maturity and by rating, and in futures markets. We show how this deviation from the standard CAPM can be captured using betting against beta factors, which could also be useful as control variables in future research (Proposition 2). The return of the BAB factor rivals those of all the standard asset pricing factors (e.g., value, momentum, and size) in terms of economic magnitude, statistical significance, and robustness across time periods, subsamples of stocks, and global asset classes.

Extending the Black (1972) model, we consider the implications of funding constraints for cross-sectional and time series asset returns. We show that worsening funding liquidity should lead to losses for the BAB factor in the time series (Proposition 3) and that increased funding liquidity risk compresses betas in the cross section of securities toward one (Proposition 4), and we find consistent evidence empirically.

Our model also has implications for agents' portfolio selection (Proposition 5). To test this, we identify investors that are likely to be relatively constrained and unconstrained. We discuss why mutual funds and individual investors could be leverage constrained, and, consistent with the model's prediction that constrained investors go for riskier assets, we find that these investor groups hold portfolios with betas above one on average.

Conversely, we show that leveraged buyout funds and Berkshire Hathaway, all of which have access to leverage, buy stocks with betas below one on average, another prediction of the model. Hence, these investors could be taking advantage of the BAB effect by applying

leverage to safe assets and being compensated by investors facing borrowing constraints who take the other side. Buffett bets against beta as Fisher Black believed one should.

Appendix A. Analysis and Proofs

Before we prove our propositions, we provide a basic analysis of portfolio selection with constraints. This analysis is based on Fig. A1. The top panel shows the mean-standard deviation frontier for an agent with $m < 1$, that is, an agent who can use leverage. We see that the agent can leverage the tangency portfolio T to arrive at the portfolio $\overline{T}$. To achieve a higher expected return, the agent needs to leverage riskier assets, which gives rise to the hyperbola segment to the right of $\overline{T}$. The agent in the graph is assumed to have risk preferences giving rise to the optimal portfolio $\overline{C}$. Hence, the agent is leverage constrained so he chooses to apply leverage to portfolio C instead of the tangency portfolio.

The bottom panel of Fig. A1 similarly shows the mean-standard deviation frontier for an agent with $m > 1$, that is, an agent who must hold some cash. If the agent keeps the minimum amount of money in cash and invests the rest in the tangency portfolio, then he arrives at portfolio T'. To achieve higher expected return, the agent must invest in riskier assets and, in the depicted case, he invests in cash and portfolio D, arriving at portfolio D'.

Unconstrained investors invest in the tangency portfolio and cash. Hence, the market portfolio is a weighted average of T and riskier portfolios such as C and D. Therefore, the market portfolio is riskier than the tangency portfolio.

A.1. Proof of Proposition 1

Rearranging the equilibrium-price Eq. (7) yields

$$\begin{aligned} E_t\left(r_{t+1}^s\right) &= r^f + \psi_t + \gamma \frac{1}{P_t^s} e_s' \Omega \chi^* \\ &= r^f + \psi_t + \gamma \frac{1}{P_t^s} \operatorname{cov}_t\left(P_{t+1}^s + \delta_{t+1}^s, \left[P_{t+1} + \delta_{t+1}\right]' \chi^*\right) \\ &= r^f + \psi_t + \gamma \operatorname{cov}_t\left(r_{t+1}^s, r_{t+1}^M\right) P_t' \chi^* \end{aligned} \tag{18}$$

where e_s is a vector with a one in row s and zeros elsewhere. Multiplying this equation by the market portfolio weights $w^s = x^{*s} P_t^s / \sum_j \chi^{*j} P_t^j$ and summing over s gives

$$E_t\left(r_{t+1}^M\right) = r^f + \psi_t + \gamma \operatorname{var}_t\left(r_{t+1}^M\right) P_t' \chi^* \tag{19}$$

that is,

$$\gamma P_t' \chi^* = \frac{\lambda_t}{\operatorname{var}_t\left(r_{t+1}^M\right)} \tag{20}$$

Inserting this into Eq. (18) gives the first result in the proposition. The second result follows from writing the expected return as

$$E_t\left(r_{t+1}^s\right) - r^f = \psi_t\left(1 - \beta_t^s\right) + \beta_t^s\left(E_t\left(r_{t+1}^M\right) - r^f\right) \tag{21}$$

Figure A1. Portfolio selection with constraints. The top panel shows the mean-standard deviation frontier for an agent with $m < 1$ who can use leverage, and the bottom panel shows that of an agent with $m > 1$ who needs to hold cash.

and noting that the first term is (Jensen's) alpha. Turning to the third result regarding efficient portfolios, the Sharpe ratio increases in beta until the tangency portfolio is reached and decreases thereafter. Hence, the last result follows from the fact that the tangency portfolio has a beta less than one. This is true because the market portfolio is

an average of the tangency portfolio (held by unconstrained agents) and riskier portfolios (held by constrained agents) so the market portfolio is riskier than the tangency portfolio. Hence, the tangency portfolio must have a lower expected return and beta (strictly lower if and only if some agents are constrained). □

A.2. Proof of Propositions 2–3

The expected return of the BAB factor is

$$\begin{aligned} E_t\left(r_{t+1}^{BAB}\right) &= \frac{1}{\beta_t^L}\left(E_t\left(r_{t+1}^L\right)-r^f\right)-\frac{1}{\beta_t^H}\left(E_t\left(r_{t+1}^H\right)-r^f\right) \\ &= \frac{1}{\beta_t^L}\left(\psi_t+\beta_t^L\lambda_t\right)=\frac{1}{\beta_t^H}\left(\psi_t+\beta_t^H\lambda_t\right) \\ &= \frac{\beta_t^H-\beta_t^L}{\beta_t^L\beta_L^H}\psi_t \end{aligned} \tag{22}$$

Consider next a change in m_t^k. Such a change in a time t margin requirement does not change the time t betas for two reasons. First, it does not affect the distribution of prices in the following period $t + 1$. Second, prices at time t are scaled (up or down) by the same proportion due to the change in Lagrange multipliers as seen in Eq. (7). Hence, all returns from t to $t + 1$ change by the same multiplier, leading to time t betas staying the same.

Given Eq. (22), Eq. (12) in the proposition now follows if we can show that ψ_t increases in m^k because this leads to

$$\frac{\partial E_t\left(r_{t+1}^{BAB}\right)}{\partial m_t^k}=\frac{\beta_t^H-\beta_t^L}{\beta_t^L\beta_t^H}\frac{\partial\psi_t}{\partial m_t^k}>0 \tag{23}$$

Further, because prices move opposite required returns, Eq. (11) then follows. To see that an increase in m_t^k increases ψ_t, note that the constrained agents' asset expenditure decreases with a higher m_t^k. Indeed, summing the portfolio constraint across constrained agents [where Eq. (2) holds with equality] gives

$$\sum_{i\ constrained}\sum_s \chi^{i,s}P_t^s=\sum_{i\ constrained}\frac{1}{m^i}W_t^i \tag{24}$$

Because increasing m^k decreases the right-hand side, the left-hand side must also decrease. That is, the total market value of shares owned by constrained agents decreases.

Next, we show that the constrained agents' expenditure is decreasing in ψ. Hence, because an increase in m_t^k decreases the constrained agents' expenditure, it must increase ψ_t as we wanted to show.

$$\frac{\partial}{\partial\psi}\sum_{i\ constrained}P_t'\chi^i=\sum_{i\ constrained}\left(\frac{\partial P_t'}{\partial\psi}\chi^i+P_t'\frac{\partial\chi^i}{\partial\psi}\right)<0 \tag{25}$$

to see the last inequality, note that clearly $(\partial P_t/\partial\psi)'\chi^i < 0$ since all the prices decrease by the same proportion [seen in Eq. (7)] and the initial expenditure is positive. The second term is also negative because

$$\begin{aligned}
\sum_{i\ constrained} P_t' \frac{\partial}{\partial\psi}\chi^i &= \sum_{i\ constrained} P_t' \frac{\partial}{\partial\psi}\frac{1}{\gamma^i}\Omega^{-1}\times\left(E_t\left(P_{t+1}+\delta_{t+1}\right)-\left(1+r^f+\psi_t^i\right)\frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi}\right)\\
&= -P_t'\frac{\partial}{\partial\psi}\Omega^{-1}\sum_{i\ constrained}\frac{1}{\gamma^i}\left(1+r^f+\psi_t^i\right)\frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi}\\
&= -P_t'\frac{\partial}{\partial\psi}\Omega^{-1}\frac{1}{\gamma}\left(q\left(1+r^f\right)+\psi\right)\frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi}\\
&= -\left(\frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi}\right)'\frac{\partial}{\partial\psi}\Omega^{-1}\frac{1}{\gamma}\left(q\left(1+r^f\right)+\psi\right)\\
&\quad \frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi}\\
&= -\frac{1}{1+r^f+\psi}\frac{1}{\gamma}\frac{\partial}{\partial\psi}\frac{q\left(1+r^f\right)+\psi}{1+r^f+\psi}\left(E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*\right)'\\
&\quad \Omega^{-1}\left(E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*\right)<0
\end{aligned} \tag{26}$$

where we have defined $q = \sum_{i\ constrained}(\gamma/\gamma^i) < 1$ and used that $\sum_{i\ constrained}(\gamma/\gamma^i)\psi^i = \sum_i(\gamma/\gamma^i)\psi^i = \psi$ since $\psi^i = 0$ for unconstrained agents. This completes the proof. □

A.3. Proof of Proposition 4

Using the Eq. (7), the sensitivity of prices with respect to funding shocks can be calculated as

$$\frac{\partial P_t^s}{P_t^s}/\partial\psi_t = -\frac{1}{1+r^f+\psi_t} \tag{27}$$

which is the same for all securities s. Intuitively, shocks that affect all securities the same way compress betas toward one. To see this more rigorously, we write prices as:

$$\begin{aligned}
P_t^i &= \frac{E_t\left(P_{t+1}^i+\delta_{t+1}^i\right)-\gamma e_i'\Omega\chi^*}{1+r^f+\psi_t}\\
&= a^i z_t + z_t E_t\left(P_{t+1}^i\right)\\
&= a^i\left(z_t + z_t E\left(z_{t+1}\right)+z_t E\left(z_{t+1}\right)E\left(z_{t+2}\right)+\ldots\right)\\
&= a^i\pi_t
\end{aligned} \tag{28}$$

where we use the following definitions and that random variables are i.i.d. over time:

$$
\begin{aligned}
a^i &= E\left(\delta_{t+1}^i\right) - \gamma e_i' \Omega \chi^* \\
z_t &= \frac{1}{1 + r^f + \psi_t} \\
\pi_t &= z_t + z_t E\left(z_{t+1}\right) + z_t E\left(z_{t+1}\right) E\left(z_{t+2}\right) + \ldots = \frac{z_t}{1 - E\left(z_{t+1}\right)}
\end{aligned}
\tag{29}
$$

with these definitions, we can write returns as $r_t^i = \left(P_t^i + \delta_t^i\right) / P_{t-1}^i = \left(a_{\pi_t}^i + \delta_t^i\right) / a_{\pi_{t-1}}^i$ and calculate conditional as follows and (using that new information about m_t *and* W_t affect only the conditional distribution of π_t):

$$
\begin{aligned}
\beta_{t-1}^i &= \frac{\operatorname{cov}_{t-1}\left(r_t^i, r_t^M\right)}{\operatorname{var}_{t-1}\left(r_t^M\right)} \\
&= \frac{\operatorname{cov}_{t-1}\left(\left(a^i \pi_t + \delta_t^i\right) / a^i \pi_{t-1}, \left(a^M \pi_t + \delta_t^M\right) / a^M \pi_{t-1}\right)}{\operatorname{var}_{t-1}\left(a^M \pi_t + \delta_t^M\right) / a^M \pi_{t-1}} \\
&= \frac{\operatorname{var}_{t-1}\left(\pi_t\right) + \left(1 / a^i a^M\right) \operatorname{cov}_{t-1}\left(\delta_t^i, \delta_t^M\right)}{\operatorname{var}_{t-1}\left(\pi_t\right) + \left(1 / \left(a^M\right)^2\right) \operatorname{var}_{t-1}\left(\delta_t^M\right)}
\end{aligned}
\tag{30}
$$

Here, we use that δ_t^s and π_t are independent since the dividend is paid to the old generation of investors while π_t depends on the margin requirements and wealth of the young generation of investors.

We see that the beta depends on the security-specific cash flow covariance, $\operatorname{cov}_{t-1}\left(\delta_t^i, \delta_t^M\right)$, and the market-wide discount rate variance, $\operatorname{var}_{t-1}(\pi_t)$. securities beta below (above) one, the beta is increasing (decreasing) in $\operatorname{var}_{t-1}(\pi_t)$. Hence, a higher $\operatorname{var}_{t-1}(\pi_t)$ compresses betas, and the reverse is true for a lower $\operatorname{var}_{t-1}(\pi_t)$.

Further, if betas are compressed toward one after the formation of the BAB portfolio, then BAB will realize a positive beta as its long side is more leveraged than its short side. Specifically, suppose that the BAB portfolio is constructed based on estimated betas $\left(\hat{\beta}_t^L, \hat{\beta}_t^H\right)$ using data from period a with less variance of so Þ ψ_t that $\hat{\beta}_t^L < \beta_t^L < \beta_t^H < \hat{\beta}_t^H$. Then the BAB portfolio will have a beta of

$$
\begin{aligned}
\beta_t^{BAB} &= \frac{1}{\operatorname{var}_t\left(r_{t+1}^M\right)} \operatorname{cov}_t\left(\frac{1}{\hat{\beta}_t^L}\left(r_{t+1}^L - r^f\right) - \frac{1}{\hat{\beta}_L^H}\left(r_{t+1}^H - r^f\right), r_{t+1}^M\right) \\
&= \frac{\beta_t^L}{\hat{\beta}_t^L} - \frac{\beta_t^H}{\hat{\beta}_t^H} > 0
\end{aligned}
\tag{31}
$$

A.4. Proof of Proposition 5

To see the first part of the proposition, note that an unconstrained investor holds the tangency portfolio, which has a beta less than one in the equilibrium with funding constraints, and the constrained investors hold riskier portfolios of risky assets, as discussed in the proof of Proposition 1.

To see the second part of the proposition, note that given the equilibrium prices, the optimal portfolio is

$$\begin{aligned}\chi^i &= \frac{1}{\gamma^i}\Omega^{-1}\left(E_t\left(P_{t+1}+\delta_{t+1}\right)-\left(1+r^f+\psi_t^i\right)\frac{E_t\left(P_{t+1}+\delta_{t+1}\right)-\gamma\Omega\chi^*}{1+r^f+\psi_t}\right)\\ &= \frac{\gamma}{\gamma^i}\frac{1+r^f+\psi_t^i}{1+r^f+\psi_t}\chi^* + \frac{\psi_t-\psi_t^i}{1+r^f+\psi_t}\frac{1}{\gamma^i}\Omega^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)\end{aligned} \tag{32}$$

The first term shows that each agent holds some (positive) weight in the market portfolio χ^* and the second term shows how he tilts his portfolio away from the market. The direction of the tilt depends on whether the agent's Lagrange multiplier ψ_t^i is smaller or larger than the weighted average of all the agents' Lagrange multipliers ψ_t. A less-constrained agent tilts toward the portfolio $\Omega^{-1}E_t(P_{t+1}+\delta_{t+1})$ (measured in shares), while a more-constrained agent tilts away from this portfolio. Given the expression (13), we can write the variance-covariance matrix as

$$\Omega = \sigma_M^2 bb' + \Sigma \tag{33}$$

where $\Sigma = \text{var}(e)$ and $\sigma_M^2 = \text{var}\left(P_{t+1}^M\right)$. Using the Matrix Inversion Lemma (the Sherman-Morrison-Woodbury (formula), the tilt portfolio can be written as

$$\begin{aligned}\Omega^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right) &= \left(\Sigma^{-1}-\Sigma^{-1}bb'\Sigma^{-1}\frac{1}{\sigma_M^2+b'\Sigma^{-1}b}\right)E_t\left(P_{t+1}+\delta_{t+1}\right)\\ &= \Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)-\Sigma^{-1}bb'\Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)\frac{1}{\sigma_M^2+b'\Sigma^{-1}b}\\ &= \Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)-y\Sigma^{-1}b\end{aligned} \tag{34}$$

where $y = b'\Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)/\left(\sigma_M^2+b'\Sigma^{-1}b\right)$ is scalar. It holds that $(\Sigma^{-1}b)_s > (\Sigma^{-1}b)_k$ because $b^s > b^k$ and because s and k have the same variances and covariances in Σ, implying that $(\Sigma^{-1})_{s,j} = (\Sigma^{-1})_{k,j}$ for $j \neq s,k$ and $(\Sigma^{-1})_{s,s} = (\Sigma^{-1})_{k,k} \geq (\Sigma^{-1})_{s,k} = (\Sigma^{-1})_{k,s}$. Similarly, it holds that $\left[\Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)\right]_s < \left[\Sigma^{-1}E_t\left(P_{t+1}+\delta_{t+1}\right)\right]_k$ since higher market exposure leads to a lower price (as seen below). So, everything else equal, a higher b leads to a lower weight in the tilt portfolio.

Finally, security s also has a higher return beta than k because

$$\beta_t^i = \frac{P_t^M \operatorname{cov}\left(P_{t+1}^i+\delta_{t+1}^i, P_{t+1}^M+\delta_{t+1}^M\right)}{P_t^i \operatorname{var}\left(P_{t+1}^M+\delta_{t+1}^M\right)} = \frac{P_t^M}{P_t^i}b^i \tag{35}$$

and a higher b^i means a lower price:

$$P_t^i = \frac{E_t\left(P_{t+1}^i + \delta_{t+1}^i\right) - \gamma\left(\Omega\chi^*\right)i}{1 + r^f + \psi_t}$$
$$= \frac{E_t\left(P_{t+1}^i + \delta_{t+1}^i\right) - \gamma\left(\Sigma\chi^*\right)_i - b^i b'\chi^*\gamma\sigma_M^2}{1 + r^f + \psi_t} \tag{36}$$

Appendix B and C

See the internet appendix at http://jfe.rochester.edu/appendix.htm

Endnotes

[1] While we consider a variety of BAB factors within a number of markets, one notable example is the zero-covariance portfolio introduced by Black (1972) and studied for US stocks by Black, Jensen, and Scholes (1972), Kandel (1984), Shanken (1985), Polk, Thompson, and Vuolteenaho (2006), and others.

[2] This effect disappears when controlling for the maximum daily return over the past month (Bali, Cakici, and Whitelaw, 2011) and when using other measures of idiosyncratic volatility (Fu, 2009).

[3] The dividends and shares outstanding are taken as exogenous. Our modified CAPM has implications for a corporation's optimal capital structure, which suggests an interesting avenue of future research beyond the scope of this paper.

[4] While the risk premium implied by our theory is lower than the one implied by the CAPM, it is still positive. It is difficult to empirically estimate a low risk premium and its positivity is not a focus of our empirical tests as it does not distinguish our theory from the standard CAPM. However, the data are generally not inconsistent with our prediction as the estimated risk premium is positive and insignificant for US stocks, negative and insignificant for international stocks, positive and insignificant for Treasuries, positive and significant for credits across maturities, and positive and significant across asset classes.

[5] A natural BAB factor is the zero-covariance portfolio of Black (1972) and Black, Jensen, and Scholes (1972). We consider a broader class of BAB portfolios because we empirically consider a variety of BAB portfolios within various asset classes that are subsets of all securities (e.g., stocks in a particular size group). Therefore, our construction achieves market neutrality by leveraging (and de-leveraging) the long and short sides instead of adding the market itself as Black, Jensen, and Scholes (1972) do.

[6] Garleanu and Pedersen (2011) find a complementary result, studying securities with identical fundamental risk but different margin requirements. They find theoretically and empirically that such assets have similar betas when liquidity is good, but when funding liquidity risk rises the high-margin securities have larger betas, as their high margins make them more funding sensitive. Here, we study securities with different fundamental risk, but the same margin requirements. In this case, higher funding liquidity risk means that betas are compressed toward one.

[7] SMB, HML, and UMD are from Ken French's data library, and the liquidity risk factor is from Wharton Research Data Service (WRDS).

[8] Our results are robust to the choice of benchmark (local versus global). We report these tests in Appendix B.

[9] These factors mimic their U.S counterparts and follow Fama and French (1992, 1993, 1996). See Asness and Frazzini (2013) for a detailed description of their construction. The data can be downloaded at http:// www.econ.yale.edu/~af227/data_library.htm.

[10] The data can be downloaded at https://live.barcap.com.

[11] The distress index was provided to us by Credit Suisse.

[12] Daily returns are not available for our sample of US Treasury bonds, US corporate bonds, and US credit indices.

[13] See, for example, De Santis and Gerard (1997).

[14] The Vasicek (1973) Bayesian shrinkage factor is given by $w_i = 1 - \sigma^2_{i,TS} / \left(\sigma^2_{i,TS} + \sigma^2_{XS}\right)$ where $\sigma^2_{i,TS}$ is the variance of the estimated beta for security i and $\sigma^2_{\chi S}$ is the cross-sectional variance of betas. This estimator places more weight on the historical times series estimate when the estimate has a lower variance or when there is large dispersion of betas in the cross section. Pooling across all stocks in our US equity data, the shrinkage factor w has a mean of 0.61.

[15] See Asness, Frazzini, and Pedersen (2012) for a detailed description of this market portfolio. The market series is monthly and ranges from 1973 to 2009.

[16] We would like to thank Mark Mitchell for providing us with these data.

[17] We keep the international portfolio country neutral because we report the result of betting against beta across equity indices BAB separately in Table 8.

[18] DataStream international pricing data start in 1969, and Xpress-feed Global coverage starts in 1984.

[19] We are viewing the TED spread simply as a measure of credit conditions, not as a return. Hence, the TED spread at the end of the return period is a measure of the credit conditions at that time (even if the TED spread is a difference in interest rates that would be earned over the following time period).

[20] As further consistent evidence, younger people and people with less financial wealth (who might be more constrained) tend to own portfolios with higher betas (Calvet, Campbell, and Sodini, 2007, Table 5). Further, consistent with the idea that leverage requires certain skills and sophistication, Grinblatt, Keloharju, and Linnainmaa (2011) report that individuals with low intelligence scores hold higher-beta portfolios than individuals with high intelligence scores.

References

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. J. Financial Econ. 77, 375–410.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. J. Finance 61, 259–299.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further US evidence. J. Financial Econ. 91, 1–23.

Ashcraft, A., Garleanu, N., Pedersen, L.H., 2010. Two monetary tools: interest rates and hair cuts. NBER Macroeconomics Annu. 25, 143–180.

Asness, C., Frazzini, A., 2013. The devil in HML's details. J. Portfol. Manage. 39, 49–68.

Asness, C., Frazzini, A., Pedersen, L.H., 2012. Leverage aversion and risk parity. Financial Anal. J. 68 (1), 47–59.

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as limits to arbitrage: understanding the low volatility anomaly. Financial Anal. J. 67 (1), 40–54.

Bali, T., Cakici, N., Whitelaw, R., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. J. Financial Econ. 99 (2), 427–446.

Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. J. Finance 55, 773–806.

Black, F., 1972. Capital market equilibrium with restricted borrowing. J. Bus. 45 (3), 444–455.

Black, F., 1993. Beta and return. J. Portfol. Manage. 20, 8–18.

Black, F., Jensen, M.C., Scholes, M., 1972. The capital asset pricing model: some empirical tests. In: Jensen, M.C. (Ed.), Studies in the Theory of Capital Markets, Praeger, New York, NY, pp. 79–121.
Brennan, M.J., 1971. Capital market equilibrium with divergent borrowing and lending rates. J. Financial Quant. Anal. 6, 1197–1205.
Brennan, M.J., 1993. Agency and asset pricing. Unpublished working paper. University of California, Los Angeles, CA.
Brunnermeier, M., Pedersen, L.H., 2009. Market liquidity and funding liquidity. Rev. Financial Stud. 22, 2201–2238.
Calvet, L.E., Campbell, J.Y., Sodini, P., 2007. Down or out: assessing the welfare costs of household investment mistakes. J. Political Econ. 115 (5), 707–747.
Carhart, M., 1997. On persistence in mutual fund performance. J. Finance 52, 57–82.
Cohen, R.B., Polk, C., Vuolteenaho, T., 2005. Money illusion in the stock market: the Modigliani-Cohn hypothesis. Q. J. Econ. 120 (2), 639–668.
Cuoco, D., 1997. Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. J. Econ. Theory. 72 (1), 33–73.
De Santis, G., Gerard, B., 1997. International asset pricing and portfolio diversification with time-varying risk. J. Finance 52, 1881–1912.
Duffee, G., 2010. Sharpe ratios in term structure models. Unpublished working paper. Johns Hopkins University, Baltimore, MD.
Elton, E.G., Gruber, M.J., Brown, S.J., Goetzmannn, W., 2003. Modern Portfolio Theory and Investment Analysis. Wiley, Hoboken, NJ. Falkenstein, E.G., 1994. Mutual Funds, Idiosyncratic Variance, and Asset Returns. Northwestern University, IL (dissertation).
Fama, E.F., 1984. The information in the term structure. J. Financial Econ. 13, 509–528.
Fama, E.F., 1986. Term premiums and default premiums in money markets. J. Financial Econ. 17, 175–196.
Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. J. Finance 47 (2), 427–465.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financial Econ. 33, 3–56.
Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. J. Finance 51, 55–84.
Frazzini, A., Kabiller, D., Pedersen, L.H., 2012. Buffett's Alpha. Unpublished working paper. AQR Capital Management and New York University, Greenwich, CT and New York, NY.
Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. J. Financial Econ. 91 (1), 24–37.
Garleanu, N., Pedersen, L.H., 2011. Margin-based asset pricing and deviations from the law of one price. Rev. Financial Stud. 24 (6), 1980–2022.
Gibbons, M., 1982. Multivariate tests of financial models: a new approach. J. Financial Econ. 10, 3–27.
Grinblatt, M., Keloharju, M., Linnainmaa, J., 2011. IQ and stock market participation. J. Finance 66 (6), 2121–2164.
Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. J. Financial Econ. 66, 361–407.
Hindy, A., 1995. Viable prices in financial markets with solvency constraints. J. Math. Econ. 24 (2), 105–135.
Kacperczyk, M., Sialm, C., Zheng, L., 2008. Unobserved actions of mutual funds. Rev. Financial Stud. 21, 2379–2416.
Kandel, S., 1984. The likelihood ratio test statistic of mean-variance efficiency without a riskless asset. J. Financial Econ. 13, 575–592.

Karceski, J., 2002. Returns-chasing behavior, mutual funds, and beta's death. J. Financial Quan. Anal. 37 (4), 559–594.

Lewellen, J., Nagel, S., 2006. The conditional CAPM does not explain asset-pricing anomalies. J. Financial Econ. 82 (2), 289–314.

Mehrling, P., 2005. Fischer Black and the Revolutionary Idea of Finance. Wiley, Hoboken, NJ.

Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. J. Financial Econ. 8, 323–361.

Moskowitz, T., Ooi, Y.H., Pedersen, L.H., 2012. Time series momentum. J. Financial Econ. 104 (2), 228–250.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. J. Political Econ. 111, 642–685.

Polk, C., Thompson, S., Vuolteenaho, T., 2006. Cross-sectional forecasts of the equity premium. J. Financial Econ. 81, 101–141.

Shanken, J., 1985. Multivariate tests of the zero-beta CAPM. J. Financial Econ. 14, 327–348.

Tobin, J., 1958. Liquidity preference as behavior toward risk. R.. Econ. Stud. 25, 65–86.

Vasicek, O.A., 1973. A note on using cross-sectional information in Bayesian estimation on security beta's. J. Finance 28 (5), 1233–1239.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48 (4), 817–838.

Common Factors in Corporate Bond Returns

Ronen Israel, Diogo Palhares, and Scott Richardson

We find that four well-known characteristics (carry, defensive, momentum, and value) explain a significant portion of the cross-sectional variation in corporate bond excess returns. These characteristics have positive risk-adjusted expected returns and are not subsumed by traditional market premia or respective equity anomalies. The returns are economically significant, not explained by macroeconomic exposures, and there is some evidence that mispricing plays a role, especially for momentum.

1 Introduction

Corporate bonds are an enormous—and growing—source of financing for companies around the world. As of the first quarter of 2016, there was $8.36 trillion of U.S. corporate debt outstanding, and from 1996 to 2015 corporate bond issuance grew from $343 billion to $1.49 trillion (Securities Industry and Financial Markets Association). Surprisingly little research, however, has investigated the cross-sectional determinants of corporate bond returns.

We study the drivers of the cross-section of corporate bond expected returns. To do so, we focus on a set of characteristics that has been shown to predict returns in other markets, yet researchers have not studied the viability of all these characteristics to predict returns in credit markets. The characteristics are carry, quality, momentum, and value (Koijen et al., 2014 for carry; Frazzini and Pedersen, 2014 for quality; Asness et al., 2013 for momentum and value). Our contribution includes (i) applying these concepts to credit markets; (ii) studying them together in a way that shines light on their joint relevance or lack thereof; (iii) evaluating their economic significance by examining both long-and-short, transaction-costs-oblivious portfolios and also long-only, transaction-costs aware portfolios; and (iv) exploring the source of the return premia by testing both risk- and mispricing explanations.

Reprinted from the *Journal of Investment Management*, 2nd Quarter 2018, Vol. 16 (2) 17–46, with permission.

Using traditional long-and-short portfolio analysis and cross-sectional regressions we find positive risk premiums that are highly significant (*t*-statistics of 3 or more) for all characteristics but carry. These premia are distinct from traditional long-only bond and equity market risk premia, as well as from the premium earned by long-and-short equity anomalies based on value, momentum and defensive. The strong relation among carry, defensive, value and momentum and future credit excess returns can be interpreted as out-of-sample evidence for the broader efficacy of these characteristics.

We also make a methodological contribution to long-and-short portfolio analysis for credit markets. The volatility and market beta of corporate bonds is tightly related to credit spreads and durations. Many return predictors are also correlated with credit spreads. As a consequence, if one just creates portfolios by sorting on these measures, the long-and-short portfolios will have very different risk profiles, making their expected returns hard to compare and the long-high-short-low portfolio far from market neutral. In the end, contrary to what happens in equity long-and-short portfolios that end up with small market exposures, simple long-and-short credit portfolios do not. Furthermore, credit spread itself is a return predictor so it is important to understand whether a candidate variable simply predicts returns because of its correlation with spreads or whether it has any extra forecasting power. To solve these general issues, we use a double sort on an ex-ante measure of beta (duration times spread) and the candidate characteristic.

Trading costs and liquidity are also very different in credit markets relative to equity markets. Corporate bonds are difficult to trade and the expected trading cost is high relative to the underlying volatility of the asset class (see, e.g., Harris, 2015). Thus, simple analyses of Sharpe ratios based on "academic" quintile long/short portfolios may substantially overstate the economic significance of any characteristic. Thus, when studying credit portfolios, we explicitly account for transaction costs and other potential trading restrictions.

To establish more realistic returns we also study long-only portfolios of relatively liquid corporate bonds with exposure to carry, defensive, momentum and value themes. We show that these portfolios generate high risk-adjusted returns, net of trading costs. Relative to a value-weighted benchmark of corporate bonds, the long-only portfolio yields a net (of transaction cost) active return of 2.20% annualized, which translates to an information ratio of 0.86. While the number is a point estimate out of a roughly 20-year sample, its exact magnitude is less important than the fact that it is well above zero.

We explore possible explanations for the observed return's patterns. We examine both risk explanations—characteristic portfolios expose the aggregate investor to losses at times in which those losses are particularly tough to bear—and mispricing theories—investors deviate from rationality because of mistakes or agency problems and limits-to-arbitrage stop arbitrageurs from fully correcting these mistakes and their impact on asset prices.

We examine the risk hypothesis with two tests. First we measure the exposure of each individual characteristic and a combination of all of them to traditional macroeconomic factors (e.g., Chen et al., 1986). While the coefficients are statistically significant, they suggest that the characteristics have a hedging profile. That is, the returns of the combined portfolio are higher when growth expectations are lower, volatility increases and inflation expectations increase. In the second test, we replace the traditional macroeconomic factors with changes in broker-dealer leverage. Adrian et al. (2014) found that exposures to broker-dealer balance sheets can explain equity anomalies as well as government bond returns. We

do not find evidence that broker-dealer leverage can explain credit characteristic returns. In particular, the combined portfolio, exposed to all individual characteristics, has a positive (hedging), but indistinguishable from zero, loading on leverage shocks.

We break down the drivers of mispricing into (i) factors that influence the likelihood that noise traders (Grossman and Stiglitz, 1980) are important for a given security; and (ii) factors that limit the activity of arbitrageurs (Shleifer and Vishny, 1997). We proxy for the likelihood of noise traders with (i) a measure of the investor base sophistication (institutional ownership of the bond); and (ii) a measure of firm transparency (analyst coverage of the issuing firm equity). For limits to arbitrage we measure liquidity (bond amount outstanding) and ease of shorting (as reflected by the shorting fee).

We run two tests using with those proxies. In the first test we examine whether the bonds most attractive to an arbitrageur—those with extreme values for the anomalies—are particularly hard to arbitrage or are more vulnerable to investors' errors. For the shorting fee the test is one sided: are bonds that represent the most attractive short from the point of view of arbitrageurs—those with unusually low anomaly scores—unusually hard to short? We do not find evidence of this pattern for any of the characteristics.

In the second test we look at long-and-short portfolios built on security universes that differ in expected mispricing and limits to arbitrage. The hypothesis here is that anomaly returns should be stronger among the hard-to-arbitrage-or-high-error bonds. Momentum returns are indeed larger among harder-to-arbitrage-or-high-error bonds and the result is statistically significant. The other anomalies perform similarly across the different universes.

Finally, the last test focuses on the investor mistake hypothesis. We test whether the returns can be explained by investors' errors in forecasting sales (e.g., Bradshaw et al., 2001). To proxy for investors' expectations, which are unobservable, we use equity analyst forecasts. If these forecasts were rational and unfettered by agency conflicts, analyst revisions should not be predictable by public information. On the other hand, if they underestimate the sales of firms with high scores and overestimate the sales of those with low scores, the correction of those expectations may explain the anomaly premium. The evidence is in the right direction, statistically significant and quantitatively important for momentum, but not for the remaining characteristics. Overall, returns to the momentum characteristic seem to be the most tightly linked to mispricing, with the evidence being less clear about the source of the returns of other characteristics.

The remainder of this paper proceeds as follows. Section 2 discusses a simple framework for corporate bond excess returns and links our analysis to earlier papers exploring determinants of cross-sectional variation in corporate bond expected excess returns. Section 3 explains our data sources, sample-selection criteria, characteristic measures and research design. Section 4 describes our empirical analyses and Section 5 concludes.

2 A Framework for Expected Corporate Bond Excess Returns

Unlike equity markets with variants of the dividend discount model to guide empiricists in their measurement of expected returns, there is not an agreed-upon framework for estimating excess credit returns. Ex post, researchers agree that credit excess returns can, and should, be measured as the difference between the returns to a corporate bond and

an appropriately cash flow-matched treasury bond (see, e.g., Hallerbach and Houweling, 2013; Asvanunt and Richardson, 2017). Ex ante, as credit and equity are related securities, one approach would be to simply explore whether characteristics known to explain cross-sectional variation in equity excess returns also explain credit excess returns. Indeed, some recent research has followed this approach (e.g., Chordia et al., 2016). This approach amounts to testing whether priced sources of risk span across markets (e.g., Fama and French, 1993). Whilst this approach is useful in commenting on whether characteristics share similar returns across equity and credit markets, this approach misses an important point that the relevant risk across credit and equity markets are not identical. After all, simply documenting that: (i) X is correlated with equity excess returns, (ii) equity excess returns and credit excess returns are correlated, and hence (iii) X is therefore correlated with credit excess returns is not that exciting (see e.g., Lok and Richardson, 2011).

Prices of corporate bonds are not independent from, equity prices, nor are they simply a mirror image. First, while the fundamental value of bonds and equities both depend on the underlying value of the assets of the firm (e.g., Merton, 1974), the way these two assets respond to changes in properties of asset values is not identical. Second, equity and bond values can change even when the underlying value of the firm business does not. Corporate events such as leveraged buyouts, for example, tend to benefit shareholders at the expense of debtholders. Third, bonds and equities are traded in two different markets and typically held by different investors. This can make stock and bond valuations diverge, as they are anchored to the risk aversion, liquidity demands and sentiment of different investor clienteles. As a consequence, knowledge about the cross-section of expected stock returns does not translate one-to-one to bond returns (see, e.g., Chordia et al., 2016; Choi and Kim, 2015).

Our approach is to directly measure characteristics that could inform about expected credit excess returns. A natural candidate is the spread of the corporate bond, which we call 'carry'. This is a suitable measure of expected returns if, and only if, there is no change in either default expectations or aggregate risk premium. To complement a measure of spread, we also look to multiple characteristics that could potentially inform about future changes in spreads. Such measures include dimensions of value, momentum and quality that have been examined in equity markets. But we need to tailor these measures to reflect the type of risk priced in the credit market (notably the risk of default). As such, our paper is related to prior research exploring cross-sectional determinants of corporate bond excess returns.

Correia et al. (2012) study value investing in corporate bond markets by comparing market spreads with model-implied spreads estimated using fundamental and market-based inputs. Kwan (1996) and Gebhardt et al. (2005b) document strong evidence for equity momentum in corporate bond markets by showing that past equity returns strongly predict future corporate bond returns of the same issuer, even after controlling for corporate bond momentum. Jostova et al. (2013) examine credit momentum and show that it is profitable when used to trade high-yield U.S. corporate bonds—even when controlling for equity momentum.

Koijen et al. (2014) evaluate carry factors across several markets: for credit markets, they test corporate bond indices of varying durations, maturities and rating categories. Carvalho et al. (2014) identify a low-risk anomaly across a broad universe of fixed income assets for various measures of risk. Similarly, Frazzini and Pedersen (2014) document

positive risk-adjusted returns for portfolios that take long positions for short duration and higher-rated corporate bonds and take short positions for long duration and lower-rated corporate bonds. In contrast, Ng and Phelps (2014) note that the low-risk anomaly in corporate bonds is sensitive to the selected measure of risk.

Our work extends this literature. First, we study the stand-alone performance of characteristics and investigate the relation between them and their combined efficacy. Second, we consider simple unconstrained long-short portfolios and also more realistically investable long-only portfolios, which account for transaction costs and shorting constraints typical for corporate bonds. The investable portfolios show that our results are economically meaningful. Third, we investigate the sources of return predictability. We explore risk-based and non-risk-based explanations and find that macroeconomic exposures are not consistent with a positive premia for the anomalies, whereas limits to arbitrage and investor errors seem to play a role in momentum strategies, though not the others.

3 Data and Methodology

3.1 *Corporate Bond Data*

Our analysis is based on a comprehensive panel of U.S. corporate bonds between January 1997 and April 2015 measured at a monthly frequency. This panel includes all constituents of the Bank of America Merrill Lynch ("BAML") investment-grade ("U.S. Corporate Master") and high-yield ("U.S. High Yield Master") corporate bond indices. The BAML dataset relies on the industry standard for valuations, aggregating data from TRACE as well as other sources. For an academic use of the data see Schaefer and Strebulaev (2008).

Following the criteria of Haesen et al. (2013), we select a representative bond for each issuer every month. The criteria used for identifying the representative bond are selected so as to create a sample of liquid and cross-sectionally comparable bonds. Specifically, we select representative bonds on the basis of (i) seniority, (ii) maturity, (iii) age and (iv) size.

First, we filter bonds on the basis of seniority, limiting ourselves to only senior debt. We then select only the bonds corresponding to the most prevalent rating of the issuer. To do this, we first compute the amount of bonds outstanding for each rating category for a given issuer. We keep only those bonds that belong to the rating category that contains the largest fraction of debt outstanding. This category of bonds tends to have the same rating as the issuer. Second, we filter bonds on the basis of maturity. If the issuer has bonds with time to maturity between 5 and 15 years, we remove all other bonds for that issuer from the sample. If not, we keep all bonds in the sample. Third, we filter bonds on the basis of time since issuance. If the issuer has any bonds that are at most 2 years old, we remove all other bonds for that issuer. If not, we keep all bonds from that issuer in the sample. Finally, we filter on the basis of size. Of the remaining bonds, we pick the one with the largest amount outstanding. A deliberate consequence of our bond selection criteria is that we will not be exploring a liquidity premium (such as issue size) for our primary empirical analyses.

Our resulting sample includes 274,665 unique bond-month observations, corresponding to 11,804 bonds issued by 4,296 unique firms. Table 1 reports annual statistics

Table 1 Universe Statistics (January 1997–April 2015)

Year	Count	Total notional	%IG	%HY	% Linked to Compustat
1997	1,096	239	60%	40%	54%
1998	1,188	278	61%	39%	53%
1999	1,104	306	63%	37%	52%
2000	1,026	335	65%	35%	50%
2001	1,026	375	70%	30%	49%
2002	1,099	443	70%	30%	49%
2003	1,263	511	63%	37%	49%
2004	1,398	562	60%	40%	47%
2005	1,291	569	59%	41%	45%
2006	1,268	560	58%	42%	43%
2007	1,256	578	56%	44%	43%
2008	1,046	553	64%	36%	47%
2009	967	540	66%	34%	49%
2010	1,269	689	56%	44%	46%
2011	1,380	768	53%	47%	46%
2012	1,406	812	53%	47%	46%
2013	1,521	893	51%	49%	45%
2014	1,564	936	50%	50%	45%
2015	1,533	948	51%	49%	46%
Average	1,247	573	59%	41%	48%

The table reports annual summary statistics of the Bank of America Merrill Lynch (BAML) bond sample. Each column statistic is computed monthly and averaged within the specified year. Investment grade (IG) and high yield (HY) classifications are based on S&P ratings. Bond issues are linked to Compustat based on CUSIPs and tickers as described in the text. Total notional is reported in billions of dollars.

describing the composition of our sample over time. The average month in the sample consists of 1,247 bonds representing $573 billion of total notional outstanding, of which 59% (41%) corresponds to investment grade (high yield) issues. To construct variables requiring financial statement information, we can link 48% of our universe to the Compustat database (using CUSIP and ticker identifiers contained in the BAML dataset).

Next, we describe a few key variables contained in the BAML dataset. Option-adjusted spread (OAS) is the fixed spread that needs to be added to the Treasury curve such that the corporate bond's discounted payments match to its traded market price (accounting for embedded options). Duration, which measures a bond's sensitivity to interest rates, is also adjusted for embedded optionality. BAML provides total returns as well as excess returns, which are equal to total returns minus the return of a duration-matched Treasury. Credit ratings are based on Standard & Poor's ratings classification system. To construct numerical ratings that can be used in our regressions, we map ratings of AAA, AA, A, BBB, BB, B, CCC, CC, C and D to scores of 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10, respectively. A rating less (greater) than or equal to 4 (5) therefore corresponds to investment grade (high yield).

Table 2 Issue and Issuer Characteristics (January 1997–April 2015)

	Mean	Std.	5%	10%	25%	50%	75%	90%	95%
OAS	386	308	85	107	161	302	512	783	1,002
Duration	5.1	2.2	1.6	2.4	3.8	5.0	6.3	7.3	8.2
Total Ret.	0.6%	3.1%	–2.9%	–1.6%	–0.4%	0.6%	1.7%	3.0%	4.2%
Excess Ret.	0.2%	3.0%	–3.2%	–1.9%	–0.7%	0.2%	1.2%	2.4%	3.6%
Amt. Out.	437	442	134	159	208	309	495	811	1,123
Time to Mat.	7.8	5.1	2.7	3.9	5.5	7.1	8.7	10.4	15.5
Age Percent	28%	19%	5%	7%	12%	24%	39%	54%	67%
Rating	4.7	1.4	2.5	3.0	3.8	4.7	6.0	6.6	6.9
Dist. to Def	6.0	3.5	1.4	2.0	3.4	5.5	8.0	10.6	12.2
Momentum	5%	16%	–16%	–10%	–3%	2%	11%	24%	36%
Leverage	0.31	0.41	–0.02	0.03	0.13	0.28	0.47	0.66	0.77

The table reports summary statistics of bond issue and issuer characteristics (as defined in Table A.I). For each characteristic, the column statistic is computed on a monthly basis and then averaged over the full sample period.

As newly issued bonds tend to be more liquid, we define a measure of bond illiquidity labelled "age percent," which is computed as time-since-issuance (in days) divided by original maturity (in days).

Table 2 provides a description of several issue and issuer characteristics. All of our variable definitions are presented in Table A.1. For each characteristic, we compute several statistics (e.g., mean, standard deviation and various percentiles) on a monthly basis and report the average of these monthly statistics in the table. The average issue in our sample has an OAS of 386 basis points, duration of 5.1 years, $437 million of notional outstanding, 7.8 years to maturity, and age percent of 28%. The average issuer in our sample has a six-month average credit and equity excess return of 5% and market leverage of 0.31.

3.2 Characteristic Measures

In this section, we define the four key characteristics that we use to explain cross-sectional variation in corporate bond excess returns. Our choices are driven by the desire to have intuitive and, to the extent possible, standard measures that span both public and private issuers of corporate bonds. When multiple measures satisfy those criteria, we combine them using equal-risk weights to obtain a more robust portfolio and make the results less susceptible to a specific variable selection.[1] We deliberately do not select size as a characteristic, as the corporate bond market is notoriously expensive to trade. Our interest is in the identification of characteristics that explain excess returns of large and liquid corporate bonds.

Carry is the return of a security if time passes but market conditions do not change and we measure it using the option-adjusted spread (OAS). We use OAS rather than bond yield because we are interested in credit returns in excess of key-rate-duration-matched treasuries. Bond yield reflects both the credit component and the Treasury component.

OAS also has its problems. It is a perfect measure of carry only if the credit curve is flat. If the curve has a positive or negative slope, OAS will underestimate and overestimate carry, respectively. Most issuers have upward sloping credit spread term structures, implying that the OAS will be an imperfect measure of carry. The alternative, however, is to estimate credit spread curves for each issuer. While potentially more precise, the curve interpolation exercise is model dependent and adds considerable complexity and opaqueness to the carry measure. In our view, OAS strikes a reasonable balance between precision on one hand and simplicity and transparency on the other.

Past research has identified a tendency for safer low-risk assets to deliver a higher risk-adjusted return (e.g., Frazzini and Pedersen, 2014; Carvalho et al., 2014). We apply this idea to corporate bonds by building a *defensive* (or low-risk) measure using multiple variables. Our first measure is market leverage, measured as the value of net debt (book debt + minority interest + preferred stocks – cash) divided by the sum of the value of net debt and market value of equity. Both intuitively and theoretically speaking, firms with higher levels of leverage (or greater use of debt) are more likely to default and are hence fundamentally riskier (e.g., Altman, 1968; Shumway, 2001).

Our second measure of safety is gross profitability as defined in Novy–Marx (2013). Unlike other profitability measures, such as net income over equity value, gross profitability speaks to the quality of the overall assets owned by the firm. As such, it reasonably proxies for the safety of the enterprise, covering both equity and debt claims.

Our third measure of safety is simply low duration. Binsbergen and Koijen (2015) document that short maturity securities across different asset classes tend to have higher risk-adjusted returns. Palhares (2013) has shown that this also holds among single-name credit default swaps. Here we apply the same concept to corporate cash bonds.

For financial instruments that trade in cash markets (i.e., government bonds and equities), there is reliable evidence of a negative relation between beta and future excess returns (e.g., Frazzini and Pedersen, 2014). One reason for this negative relation is the prevalence of leverage-averse investors in cash markets who seek higher returns by buying higher beta assets as opposed to levering up the mean–variant efficient portfolio. Indeed, evidence from holdings of equity mutual funds shows that the average stock held has a beta of about 1.08 (see Table 11 of Frazzini and Pedersen, 2014).

For credit markets, both systematic and idiosyncratic volatility can be captured by the product of duration and spread, or DTS (e.g., Ben Dor et al., 2007). The first component, duration, has been shown to be negatively associated with risk-adjusted returns in equities, bonds and several other asset classes (e.g., Palhares, 2013; Binsbergen and Koijen, 2015). The second component, credit spread, simply measures carry in credit markets. Beta and idiosyncratic volatility, therefore, implicitly combine two measures that have confounding effects on expected returns, leading to their inadequacy as suitable characteristics to explain corporate bond excess returns. As a consequence we have excluded beta and volatility as measures of the defensive theme.

For our *momentum* characteristic, we use two widely studied momentum measures. The first is credit momentum defined as the trailing six-month bond excess return. Jostova et al. (2013) shows that, in a broad sample of corporate bonds, including both high-yield and investment-grade securities, past winners tend to outperform past losers. The second momentum measure is the six-month equity momentum of the bond issuer. Kwan (1996) and Gebhardt et al. (2005b) show that stock returns tend to lead corporate bond returns.

To construct a value signal, we need a market value measure (price, yield, spread, etc.), a fundamental value measure and a way to compare the two. For example, Fama and French (2003) use the price of a stock for the market measure, the book value for the fundamental measure and the ratio to make a comparison. For credit markets we use the spread of the bond and credible measures of default risk as the fundamental anchor. A cheap bond has high spread relative to default risk.

We use two proxies for default risk. First, we follow Correia et al. (2012) and use the issuer default probability. We measure the default probability as in Bharath and Shumway (2008). One drawback of this approach is that it can only be computed for issuers with publicly traded equity. To increase coverage, we use a second value anchor that combines three broadly available fundamental measures: credit rating, bond duration and the volatility of bond excess return returns in the last 12 months.

3.3 Portfolio Construction

The traditional way of examining the relationship between expected returns and a candidate predictor in the equity literature consists of constructing portfolios based on the cross-sectional rank of the characteristic, averaging the returns within the portfolio and then averaging those over time (e.g., Fama and French, 1993). This approach does not guarantee that the different quantile portfolios will have similar ex-ante volatilities and beta, and, as consequence, that the long-top-minus-short-bottom portfolios will be market neutral. In spite of that, in the equity literature, the different anomaly portfolios do tend to have similar risk and the long-and-short risk factors tend to have moderate betas—though not zero, for example, the SMB and HML factors are notorious for their positive and negative betas, respectively (e.g., Fama and French, 1993) and the betting-against-beta factor (e.g., Frazzini and Pedersen, 2014) has negative beta.

This quirk of the traditional portfolio construction methodology is important for this paper because the cross-section of corporate bonds has a much larger dispersion in beta and risk than equity markets. Furthermore, many of the characteristic we examine correlate with beta. As a consequence, the long-and-short portfolios formed using those characteristics will not be beta neutral, complicating the interpretation of their expected returns and time-series properties as the reflection of something other than their embedded market exposure. To obtain long-and-short portfolios that are closer to market neutrality, we demean characteristics within five ex-ante beta quintiles, with beta being measured as duration times spread (DTS). We exclude duration and carry from that step because, mechanically, that would induce a portfolio that mixes high carry and low duration together.

We construct two types of characteristic portfolios. First, we follow the standard convention of computing a zero-cost portfolio, that is, long corporate bonds in the highest quintile of a given characteristic and short corporate bonds in the lowest quintile of a given characteristic. Within quintiles, we report excess returns based on value-weighted returns. Our inferences are unaffected if we instead use equal weighting. We also display a constant 5% volatility version of each long-and-short portfolio (Muir and Moreira, 2016). We use the 24-month realized volatility of the unsealed portfolio as the measure of ex-ante risk.[2]

We construct the quintiles and long-and-short portfolios for each characteristic individually and for a combination of them all. The combination sorting variable is an inverse-of-risk-weighted sum of the four characteristics. More precisely, for each characteristic we form a portfolio that is linear in ranks (Asness et al., 2014) and then multiply it by 5% and divide it by its 24-month realized volatility—the outcome is an alternative constant-volatility portfolio with linear weights instead of just having non-zero values for the most extreme quintiles. The combined characteristic is then just an equal-weighted average of those single-characteristic, linear-in-ranks portfolio weights.

A critical part of this paper is to examine the return-forecasting characteristics jointly. Each single characteristic informs us about properties of the stochastic discount factor that prices corporate credit securities, but the single portfolio that makes optimal use of the multiple characteristics goes beyond: it alone is sufficient to fully characterize that discount factor (e.g., Cochrane, 2009). From the point of view of an investor, that single portfolio is also interesting. For example, for a mean–variance investor allocating between these long-and-short credit strategies and cash, the allocation to that optimal portfolio would be sufficient to summarize its asset allocation policy.

The question is then how to build this optimal portfolio. Without observing expected returns and covariance matrices, one cannot observe the optimal portfolio weights. Using sample moments is problematic because of look-ahead bias and the relative shortness of a 20-year sample to estimate expected return. Our answer to the problem is an equal-weighted portfolio. It generalizes the robustness of the *1/N* portfolio (e.g., DeMiguel et al., 2009) by applying it to similarly risky characteristic portfolios rather than underlying assets.

For the combination of characteristics, we also analyze a second type of portfolio: a long-only portfolio that takes into consideration realistic implementation by solving a linear optimization problem. The analysis of a long-only portfolio is unusual when studying cross-sectional return predictability. But given the well-known challenges in shorting corporate bonds (e.g., Asquith et al., 2013) and the significant costs in trading corporate bonds relative to their underlying volatility (e.g., Bessembinder et al., 2006; Edwards et al., 2007), it is important to test whether the characteristic's premia survives difficult but realistic real-world constraints.

4 Results

4.1 Regression Analysis

Before reporting the performance of our portfolios, we first report Fama–Macbeth regressions of monthly corporate-bond excess returns regressed onto lagged characteristics along with control variables. Each month, we run cross-sectional regressions of the form:

$$\begin{aligned} R_{i,t+1} = \alpha &+ \beta_1 CARRY_{i,t} + \beta_2 DEF_{i,t} \\ &+ \beta_3 MOM_{i,t} + \beta_4 VALUE_{i,t} \\ &+ \gamma Z + \varepsilon_{i,t+1}, \end{aligned} \tag{1}$$

where $R_{i,t+1}$ denotes the duration-hedged excess return of bond i over month $t + 1$. Each of the four characteristics is converted to a normalized variable. Specifically, for each characteristic, for every month, we rank issues by their characteristic values, subtract the mean rank and then divide by the standard deviation of the ranks. We also fill missing values with zero, but the results are robust if we do not. As a result, estimated coefficients may be interpreted as the future one-month excess return difference for a one standard deviation difference in characteristic ranking. To rule out the hypothesis that the characteristics predict returns because they proxy for traditional measures of risk, we include control variables in the regression. The first variable is a market beta, where the market is defined as the credit return of the cap-weighted portfolio of all bonds in our database and the beta is computed using a 12-month rolling regression. For robustness, we also include two other traditional measures of risk in credit markets—rating and duration—as well as a proxy for illiquidity, age percent (e.g. Gebhardt et al., 2005a).

Table 3 reports our Fama–Macbeth regression estimates for the monthly sample period from January 1997 to April 2015. Regression (1) includes just an intercept and beta, and Regression (2) adds our control variables, which reduce the average number of bonds in the cross-section from 723 to 671. Regressions (3) through (6) evaluate the predictive ability of each of our characteristics on a stand-alone basis. Both individually and combined, the

Table 3 Fama–Macbeth Regressions (January 1997–April 2015)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.10	–0.02	0.04	–0.01	0.05	–0.10	–0.02
	[1.5]	–[0.2]	[0.5]	–[0.1]	[0.5]	–[1.2]	–[0.2]
Carry			0.00				0.14
			[1.0]				[2.3]
Defensive				0.15			0.03
				[5.0]			[0.9]
Momentum					0.15		0.22
					[3.3]		[7.1]
Value						0.26	0.30
						[5.8]	[10.7]
Mkt beta	0.05	0.04	0.10	0.04	0.06	0.08	0.14
	[0.7]	[0.6]	[1.6]	[0.7]	[0.9]	[1.2]	[2.3]
Rating		0.02	–0.03	0.02	0.00	0.03	0.00
		[0.8]	–[1.0]	[0.6]	[0.1]	[1.0]	–[0.1]
Duration		–0.01	–0.01	0.01	0.00	0.01	0.01
		–[0.5]	–[0.5]	[0.8]	–[0.4]	[1.1]	[1.1]
Age percent		0.25	0.23	0.22	0.23	0.14	0.09
		[2.2]	[2.0]	[1.9]	[2.0]	[1.2]	[0.9]
Avg. *R*-squared	0.07	0.10	0.14	0.10	0.11	0.11	0.15
Avg. Num. Obs.	723	671	671	671	671	671	671

The table reports Fama–Macbeth regressions of monthly bond excess returns regressed onto normalized carry, defensive, momentum and value style measures along with controls for market beta, rating, duration and age percent variables (as defined in Table A.1).

value and momentum characteristics have explanatory power for corporate bond excess returns. The carry characteristic does not exhibit a reliable association with future bond excess returns as a stand-alone variable but it is marginally significant when controlling for the remaining characteristics. The opposite is true with defensive: it is highly significant as a stand-alone variable but loses significance when controlling for value and momentum. This suggests that the defensive theme in credit may be spanned by the value and momentum themes. This is not surprising as the value factors we build for credit make explicit use of fundamental information. Our value measures identify a bond as cheap when its spread is wide relative to default probabilities. Our measures of default probabilities include distance to default and rating information. These fundamental anchors incorporate measures of leverage and expected profitability. As a consequence, it is not surprising that they help explain the defensive premium.

The average R-squared of the Fama–Macbeth cross-sectional regressions is 15%, suggesting that our characteristics collectively explain a nontrivial portion of the cross-sectional variation in bond excess returns. The interpretation of the 15% average explanatory power is not that we can predict 15% of the variation in corporate bond excess returns but rather that knowledge of the four characteristics combined with (unknown ex ante) time-varying loadings to our four characteristics can explain 15% of the variation in corporate bond excess returns. To put that number in context, Lewellen (2015) finds that 15 equity characteristics explain 7.6% of the cross-sectional variation of equity returns. The value and momentum characteristics have the strongest statistical relation with future excess returns, as indicated by the large positive Fama–Macbeth test statistics in the final column.

4.2 Long-short Quintile Portfolios

Table 4 reports performance statistics of our long-short quintile portfolios. Consistent with the Fama–Macbeth results, we see the strongest positive association between characteristics and returns for defensive, momentum and value. A portfolio that combines all of the factors at an equal-risk weight ("combined") performs even better, with an annualized Sharpe ratio of 2.19, indicating that the different characteristics are weakly correlated amongst themselves. Note also that the realized volatilities of the constant-volatility portfolios are close to the targeted value of 5%, confirming that our simple scalar methodology succeeds reasonably in estimating the volatility of the combined portfolio.

Across all characteristics, we can see that the long-short returns are driven by positive performance on the long-side and negative (or weaker) performance on the short-side. In fact, reading Sharpe ratios across each of the rows clearly illustrates that performance is generally monotonically increasing across quintiles for each of the characteristics.

Figure 1 plots cumulative excess characteristic returns over time. We can see that performance, especially for the combination of characteristics, is not driven by any particular sub-period and has not changed substantially over time. While different characteristics performed better and worse over different sub-periods, it is clear that the combined portfolio has been relatively stable in its outperformance. Not surprisingly the most visible drawdown is carry during the Global Financial Crisis, when investors sought safe assets and shunned riskier ones like high-yield bonds (e.g., Koijen et al., 2014). Whilst we are hesitant to draw too strong inferences from a relatively short time period, the relative smoothness

Table 4 Quintile Portfolio Tests (January 1997–April 2015)

		Q1	Q2	Q3	Q4	Q5	Q5–Q1	ConstVol
Carry	Ret.	–0.4%	1.1%	1.5%	3.7%	3.7%	4.1%	1.1%
	Vol.	2.9%	4.4%	6.6%	8.7%	13.9%	11.7%	5.8%
	S.R.	–0.12	0.26	0.22	0.43	0.27	0.35	0.19
Defensive	Ret.	0.0%	1.4%	2.0%	1.9%	2.7%	2.7%	8.3%
	Vol.	6.0%	5.8%	6.4%	6.2%	5.6%	2.4%	6.9%
	S.R.	0.00	0.24	0.31	0.32	0.49	1.11	1.21
Momentum	Ret.	–0.2%	1.3%	1.5%	1.4%	2.7%	2.9%	7.5%
	Vol.	7.2%	6.1%	5.2%	5.3%	6.5%	3.4%	6.7%
	S.R.	–0.03	0.21	0.28	0.27	0.41	0.85	1.12
Value	Ret.	–0.4%	0.7%	1.6%	2.4%	3.5%	3.9%	10.7%
	Vol.	5.5%	5.8%	6.3%	6.8%	5.6%	2.2%	6.0%
	S.R.	–0.07	0.13	0.25	0.35	0.62	1.75	1.80
Combined	Ret.	–0.5%	1.0%	1.5%	2.3%	4.9%	5.4%	14.0%
	Vol.	5.6%	5.6%	6.3%	6.8%	6.0%	2.5%	6.0%
	S.R.	–0.09	0.18	0.24	0.34	0.81	2.19	2.32

The table reports annualized performance statistics for value-weighted quintile portfolios formed on carry, defensive, momentum, value and combined style factors (as described in the text). "ConstVol" corresponds to quintile long-short portfolios targeting a constant volatility of 5% per annum (as described in the text).

Figure 1 Cumulative Style Factor Returns (January 1997–April 2015). The figure shows cumulative arithmetic returns for each of the carry, defensive, momentum, value and combined style factors (as defined in the text).

of the returns of the combined portfolio is initial evidence that risk-based explanations will be challenging to support.

To better understand the source of the characteristic's premia, we report return correlations for the various constant-volatility, long-and-short characteristic portfolios and well-known sources of risk premia. We report the various pairwise return correlations in Table 5 using the full time series of data for the period January 1997 through to April 2015, inclusive. We consider the following traditional risk premia: (i) credit risk premium ("CREDIT"), measured as the value-weighted corporate-bond excess returns; (ii) equity risk premium, measured as the difference between the total returns on the S&P500 index and one-month U.S. Treasury bills ("EQUITY"); and (iii) Treasury term premium ("TSY"), measured as the difference between total returns on 10-year U.S. Treasury bonds and one-month U.S. Treasury bills.

Several of the correlations in Table 5 are worth discussing. First, among the four characteristics, we see negative correlations between carry and the other three measures. This is not surprising as issuers with higher spreads will typically have considerable leverage and low profit margins (part of defensive), will have experienced poor recent performance (poor momentum) or both. The correlations reported here are still negative even after our attempt to mitigate the negative correlation with carry by first ranking bonds into duration-times spread groups and then ranking on characteristic measures within those groups. But they are considerably less negative than without this adjustment. While the carry characteristic is relatively less attractive on a stand-alone basis, it has low correlation with the other characteristics (see the correlations reported in Table 5). Conditional on each characteristic generating a positive risk-adjusted return on a stand-alone basis as was evident in Tables 3 and 4, the relatively low (and sometimes negative) correlations across characteristics suggest that characteristics do not span each other. Second, the correlations between the various characteristic measures and well-known sources of risk premia show that the

Table 5 Return Correlation Matrix (January 1997–April 2015)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Carry	1.00							
Defensive	–0.18	1.00						
Momentum	–0.30	0.40	1.00					
Value	–0.09	0.28	–0.16	1.00				
Combined	0.15	0.79	0.43	0.38	1.00			
CREDIT	0.80	–0.24	–0.17	–0.10	0.01	1.00		
EQUITY	0.55	–0.27	–0.05	–0.17	–0.05	0.59	1.00	
TSY	–0.49	0.02	0.03	0.10	–0.16	–0.50	–0.25	1.00

The table reports monthly excess return correlations for each of the constant-volatility, long-top-quintile-short-bottom-quintile portfolios for the carry, defensive, momentum, value and combined characteristic portfolios along with market indices corresponding to credit returns in excess of duration-matched Treasuries, equity returns in excess of the risk-free rate and Treasury returns in excess of the risk-free rate. All variables are defined in the Appendix.

characteristic premia are not simply a manifestation of these other well-known risk premia. With the exception of carry, the return correlations between the characteristic factors and risk premia are all less than 0.30 and are often *negative.*

We next test the hypothesis that the characteristic's long-and-short portfolio expected returns cannot be explained by loadings on traditional sources of market risk premia (CREDIT, EQUITY and TSY) as well as exposures to well-known equity anomalies (SMB, HML and UMD from Ken French's data library and QMJ from Asness et al. (2014)). The latter test examines whether stocks and bonds with a certain characteristic both earn their expected return due to a common exposure. For example, do cheap stocks (high book-to-market) and cheap bonds (high spread in relation to default risk) earn high average returns due to a common, shared exposure or are the two expected return sources distinct? To answer those questions, we run regressions of constant-volatility, long-and-short characteristic portfolio returns on market and equity anomaly returns as follows:

$$\begin{aligned} CONST_VOL_CHAR_t \\ &= \alpha + \beta_0 EQUITY_t + \beta_1 TSY_t + \beta_2 CREDIT_t \\ &\quad + \beta_3 SMB_t + \beta_4 HML_t + \beta_5 UMD_t \\ &\quad + \beta_6 QMJ_t + \varepsilon_{i,t}. \end{aligned} \tag{2}$$

Consistent with the simple correlations reported in Table 5, we see in Table 6 that the carry characteristic has a significant positive exposure to credit risk premium. After controlling for other well-known sources of return, the intercept is not significant for carry. The defensive characteristic is negatively correlated with market risk premia (e.g., credit risk premium), consistent with it reflecting a flight to quality or a risk-on/risk-off tendency of investors.

Momentum has a positive correlation with UMD and nothing else. Credit value exhibits a negative loading on SMB and QMJ, –2.0 and –2.3 *t*-statistics, respectively. Interestingly, the value characteristic in credit markets is mildly negatively associated with the HML factor, a result consistent with the evidence that characteristic portfolios in one asset class have limited correlations with those in other asset classes (Asness et al., 2015). In the fifth column of Table 6, we regress the combined characteristic long-short portfolio return onto the various market risk premia and equity factor returns. The combined portfolio does not have a statistically significant loading on any of the equity factors and a mildly negative relation with term premium. As a consequence, its intercept is a significant 123 basis points per month with a *t*-statistic of 9.6 and an information ratio of 2.25. The combination portfolio is superior to any individual characteristic portfolio, reassuring us that the equal-risk approach is a sensible way to combine the different characteristics. Furthermore, the fact that the combination portfolio does not load on traditional market risk premia and equity anomalies suggests that the source of return predictability is distinct from those.

Momentum has a positive correlation with UMD and nothing else. Credit value exhibits a negative loading on SMB and QMJ, –2.0 and –2.3 *t*-statistics, respectively. Interestingly, the value characteristic in credit markets is mildly negatively associated with

Table 6 Long-and-short Portfolio Alphas and Betas with Respect to Market and Equity Factors (January 1997–April 2015)

	Carry	Defensive	Momentum	Value	Combined
Intercept	0.05%	0.75%	0.55%	1.02%	1.23%
	[0.7]	[5.3]	[3.9]	[8.1]	[9.6]
CREDIT	0.58	–0.22	–0.19	–0.02	–0.07
	[10.5]	–[2.1]	–[1.8]	–[0.2]	–[0.7]
EQUITY	0.04	–0.07	0.09	–0.11	–0.02
	[0.0]	[0.0]	[0.0]	[0.0]	[0.0]
TSY	–0.12	–0.12	–0.10	0.05	–0.18
	–[2.9]	–[1.4]	–[1.2]	[0.7]	–[2.4]
SMB	0.01	0.06	0.01	–0.08	0.02
	[0.4]	[1.3]	[0.2]	–[2.0]	[0.5]
HML	–0.02	0.05	0.06	–0.02	0.06
	–[0.7]	[1.2]	[1.4]	–[0.6]	[1.4]
UMD	0.00	–0.01	0.06	–0.01	0.04
	[0.1]	–[0.2]	[2.3]	–[0.4]	[1.6]
QMJ	–0.04	0.03	0.11	–0.14	–0.06
	–[1.2]	[0.4]	[1.6]	–[2.3]	–[0.9]
R-squared	0.67	0.10	0.09	0.07	0.05

The table reports monthly excess return regressions of the carry, defensive, momentum, value and combined characteristic long-top-quintile-short-bottom-quintile, constant-volatility factors onto (i) market excess returns for Treasuries, credit and equity; (ii) equity anomaly factors SMB, HML and UMD from Ken French's website and QMJ from Asness et al. (2014). All variables are defined in the Appendix.

the HML factor, a result consistent with the evidence that characteristic portfolios in one asset class have limited correlations with those in other asset classes (Asness et al., 2015). In the fifth column of Table 6, we regress the combined characteristic long-short portfolio return onto the various market risk premia and equity factor returns. The combined portfolio does not have a statistically significant loading on any of the equity factors and a mildly negative relation with term premium. As a consequence, its intercept is a significant 123 basis points per month with a *t*-statistic of 9.6 and an information ratio of 2.25. The combination portfolio is superior to any individual characteristic portfolio, reassuring us that the equal-risk approach is a sensible way to combine the different characteristics. Furthermore, the fact that the combination portfolio does not load on traditional market risk premia and equity anomalies suggests that the source of return predictability is distinct from those.

The economic magnitude of the intercept requires further discussion. The literal interpretation would suggest that a 2.25 information ratio is available for investors. Such a statement needs to be interpreted very cautiously. Corporate bond and equity markets differ substantially in terms of their trading costs.

For example, Chen et al. (2007) show that the average bid-mid spread for BBB-rated and B-rated medium maturity bonds are 22 bps and 30 bps, respectively. Frazzini et al. (2012) report average value-weighted trading costs for global equities of 20 bps. These numbers, however, severely understate the impact of transaction costs, as stocks are much more volatile than bonds. Andersen et al. (2001) find that the median stock volatility is 22%, whereas the median bond in our sample has an excess return volatility close to 7%. More importantly, whereas our combined one-dollar-long-and-one-dollar-short portfolio from Table 4 has a 2.5% annualized volatility, Fama–French HML's factor—long 1 dollar of cheap stocks and short 1 dollar of expensive stocks—achieves 11.6% annual volatility over the same period.

Given the similarity in dollar transaction costs estimates across bonds and stocks, and similar turnover across bond and stock portfolios, the bond portfolio transaction cost per unit of risk is more than four times larger than that of equity. As a consequence, if a long-and-short portfolio of stocks and bonds are to have similar net-of-transaction costs Sharpe ratios, the bond portfolio must have a much larger gross-of-transaction cost Sharpe ratio.

To illustrate any time-varying performance across the various characteristics (in Figure 2), we use the full-sample regression coefficients from Table 6 to compute 36-month rolling average alphas for each respective long-and-short, constant-volatility characteristic portfolio. While outperformance has been marginally attenuated in recent

Figure 2 Rolling Regression Alphas. The figure shows 3-year rolling average regression alphas for each of the value, momentum, carry, defensive and combined style factors (as defined in the text). Regression alphas are computed monthly using the full-sample beta estimates (as reported in Table 6) and averaged over a trailing 36-month period.

years, it is clear that excess returns have been relatively stable and positive. Again the smoothness of the returns, albeit over a short time series, is difficult to reconcile with a risk-based explanation. We formally examine this issue in Section 4.4.

4.3 Long-only Optimized Portfolio

While our long-short characteristic portfolios suggest a robust relation between credit excess returns and each of the considered characteristics, they do not take into account actual portfolio implementation considerations. To more realistically address the hypothetical performance of our characteristic portfolios, we build and test optimized long-only portfolios with explicit portfolio implementation constraints. Hence our optimized portfolios are designed to be comparable with traditional actively managed corporate bond portfolios, which tend to be long-only (as individual bonds are difficult to short).

We build and rebalance long-only portfolios on a monthly frequency by solving a linear optimization problem. While mean–variance optimization is a commonly utilized objective function in portfolio construction, here we build our portfolios using a simpler objective function that does not require estimation of an asset-by-asset covariance matrix (i.e., an asset-level risk model). Our optimization problem is specified as follows:

$$
\begin{aligned}
&\textit{Maximize}: \quad \sum_{i=1}^{I} w_i \cdot COMBO_i \\
&\textit{subject to}: \quad w_i \geq 0, \ \forall \mathrm{i}\left(\textit{no shorting constraint}\right) \\
&\qquad \left|w_i - b_i\right| \leq 0.25\%, \quad \forall \mathrm{i}\left(\textit{deviation from benchmark constraint}\right) \\
&\qquad \sum_{i=1}^{I} w_i = 1 \left(\textit{fully invested constraint}\right) \\
&\qquad \sum_{i=1}^{I} \left|w_{i,t} - w_{i,t+1}\right| \leq 10\% \left(\textit{turnover constraint}\right) \\
&\qquad \sum_{i=1}^{I} \left|\left(w_{i,t} - w_{i,t-1}\right) \cdot PRICE_{i,t}\right| \geq \$100{,}000, \quad \forall \mathrm{i}\left(\textit{minimum trade size constraint}\right) \\
&\qquad \sum_{i=1}^{I} \left|\left(w_i - b_i\right) \cdot OAS_i\right| \leq 0.50\% \quad \left(\textit{deviation from benchmark spread constraint}\right) \\
&\qquad \sum_{i=1}^{I} \left|\left(w_i - b_i\right) \cdot Duration_i\right| \leq 0.50 \quad \left(\textit{deviation from benchmark duration constraint}\right),
\end{aligned}
$$

where ω_i is the portfolio weight for a given bond, and $COMBO_i$ is an equal-weighted combination of the carry, defensive, momentum and value long-short characteristic portfolios for a given bond. When computing the realized returns from our optimal portfolio holdings, we subtract an estimate of transaction costs based on each bond's rating and maturity in line with Table 1 of Chen et al. (2007). $PRICE_i$ is the bond price for a given

Table 7 Long-only Backtest Portfolio Performance (January 1997–April 2015)

	Optimized Portfolio	Benchmark	Active: Portfolio – Beta * Benchmark
Excess return (gross)	5.72	4.14	2.45
Excess return (net)	5.26	3.84	2.20
Volatility (net)	5.10	5.59	2.56
Sharpe ratio (net)	1.03	0.69	0.86

The table reports performance statistics for the long-only optimized backtest portfolio based on the optimization problem outlined below. The optimized portfolio refers to the stream of returns generated by the optimized long-only portfolio that maximizes the score of the bonds held as explained in the text. Benchmark is a cap-weighted portfolio of all the corporate bonds in our database; i.e., it includes both investment-grade and high-yield bonds. The active returns reported below are the returns from the optimized portfolio less the benchmark using a 24-month rolling beta. Gross returns are returns in excess of the risk-free rate only. Net returns subtract estimated transaction costs from gross returns.

bond, OAS_i is the option-adjusted spread for a given bond, $Duration_i$ is the effective duration for a given bond and b_i is the benchmark portfolio weight for a given bond based on a value-weighted benchmark of all corporate bonds in our one-bond-per-issuer dataset.

The solution to this optimization problem is a long-only corporate bond portfolio that has maximal exposure to the combined characteristic portfolio while taking into consideration the challenges of trading corporate bonds as well as the risk contribution of individual positions to the final portfolio. Importantly, we limit the portfolio's differences from (or tracking error to) the benchmark by limiting the portfolio's active weights relative to the benchmark (i.e., at most 25 bps), limit the portfolio's aggregate OAS exposure to be within 50 bps of the benchmark and limit the portfolio's aggregate duration exposure to be within 0.50 years of the benchmark. As discussed earlier, Ben Dor et al. (2007) document that spread and duration are the key determinants of volatility in credit markets. Hence constraining the aggregate active weights on these two dimensions is a simple and transparent way to control the active risk of the long-only portfolio. We also constrain turnover to at most 10% per month and force trades to be at least $100,000 (small trades are much more costly, e.g., Edwards et al., 2007). Despite our best efforts to incorporate constraints and transaction costs, the trading of corporate bonds is challenging. Thus we add the caveat to our empirical results that dynamic trading strategies in corporate bonds are not as implementable as those in more liquid assets.

Table 7 reports performance statistics for the optimized long-only portfolio as well as the benchmark. The portfolio earned an annual average excess return of 5.72% per year (and 5.26% after taking into account estimated transaction costs). Given its realized annualized volatility of 5.1%, the net Sharpe ratio over this period was 1.03. By comparison, the gross (net) benchmark earned a 4.14% (3.84%) annualized excess return with a Sharpe ratio of 0.69. The active portfolio (i.e., portfolio minus beta times the benchmark) realized an annualized net information ratio of 0.86 with a tracking error of 2.56%. Figure 3 shows the cumulative performance of the portfolio and the benchmark.

Figure 3 Cumulative Long-only Portfolio Returns (January 1997–April 2015). The figure shows cumulative returns for the optimized multi-style long-only portfolio (as described in the text) as well as a corporate bond market index constructed based on the value-weighted average of all corporate bonds in the BAML bond sample.

4.4 Investigating Risk and Behavioral Explanations

So far we have documented that value, momentum, carry and defensive measures can predict corporate bond excess returns. In other markets where these anomalies have been studied, both behavioral and risk-based explanations have been suggested. We run additional tests on credit characteristic portfolios aiming to distinguish between risk and behavioral explanations for their respective premiums.

4.4.1 Risk-based Explanations

In the first test, we ask whether exposures to traditional macroeconomic variables can explain the premiums that we uncover. We add three macroeconomic variables to the time-series regressions that we had previously run in Table 6, specifically we run:

$$\begin{aligned} CONST_VOL_CHAR_t \\ &= \alpha + \beta_1 X_t^{Market}, \beta_2 X_t^{Equity} \\ &\quad + \beta_7 \Delta LOGVIX_t + \beta_8 \Delta ALOGINDPRO_t \\ &\quad + \beta_9 \Delta LOGCPI_t + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where ΔLOGVIX, ΔLOGINDPRO and ΔLOG CPI are respectively the one-month change in the log of the VIX, seasonally-adjusted industrial production index and seasonally-adjusted consumer price index (CPI). While the intercept cannot be interpreted as a portfolio alpha because the macro variables are not tradable portfolios, we can still examine the regression slope coefficients, which are what we report in Panel A of Table 8. The combination portfolio tends to have higher returns when volatility and inflation rise and when growth falls. If anything, the combo portfolio behaves as a macroeconomic hedge and should have negative expected returns if that hedge is valuable, making its high and positive expected returns even more puzzling.

The single-characteristic portfolios behave much in the same way as the combo—coefficients suggest a macro hedge rather than macro risk profile. Carry is the exception. Its coefficient signs are consistent with the risk story, but only statistically significant for ALOGVIX. This suggests that macroeconomic risk may play a role in the carry anomaly but not for the remainder. The conclusions from the time-series regressions, however, have to be caveated by the relatively small sample (about 20 years) for this type of exercise.

In recent years, another class of rational models emerged that focus on financial intermediaries rather than on a single aggregate consumer (e.g., He and Krishnamurthy, 2013). In these models, the conditions of financial intermediaries (wealth, risk aversion, etc.) determine asset prices. Adrian et al. (2014) apply this idea to equity markets and find that exposures to increases in broker-dealers leverage can explain traditional stock factors: size, book-to-market and momentum. We test whether shocks to broker-dealer leverage can explain characteristic factor returns in credit by running time-series regressions of those quarterly returns on quarterly log changes in broker-dealers leverage, controlling for market and equity factor returns. Panel B of Table 8 shows that value has a statistically significant loading on broker-dealers leverage shocks of –0.18 (*t*-statistic of 2.2). This suggests that some of the value premium may be due to it being exposed to deterioration in dealer-brokers balance sheet. All other characteristics either have positive loadings (momentum) or loadings that are indistinguishable from zero. In particular, the combo portfolio is a hedge to broker-dealer leverage shocks, though that exposure is not statistically significant. As a consequence, although dealer's balance sheet may play a role in explaining the value characteristic, it cannot be a unified explanation for them all.

4.4.2 Mispricing Explanations

In the next set of empirical tests, we focus on deviations from market efficiency as explanations for the characteristic portfolios positive risk-adjusted returns. Investors deviate from rationality either because they make mistakes or because they are subject to portfolio management frictions (e.g., agency problems due to intermediation and regulations limiting their portfolio choices); and limits-to-arbitrage (Shleifer and Vishny, 1997) impede arbitrageurs from eliminating the ensuing price distortions. If these deviations from market efficiency are empirically descriptive, then we would expect to see bonds in the most extreme portfolios—those that end up in the top and bottom quintiles—to be illiquid, traded by more error-prone investors, issued by more obscure firms and on the short side would be costlier to borrow (for the bottom quintile). We test these hypotheses by measuring each of these attributes for each characteristic long and short portfolios separately.

Table 8 Long-and-short Portfolio Betas with Respect to Macroeconomic Variables (January 1997–April 2015)

	Carry	Defensive	Momentum	Value	Combined
Panel A: Monthly volatility, growth and inflation					
Intercept	0.00	0.01	0.01	0.01	0.01
	[1.1]	[4.4]	[3.2]	[5.6]	[7.9]
ΔLOGVIX	−0.01	0.03	0.02	0.02	0.02
	−[2.5]	[3.1]	[2.2]	[2.3]	[2.2]
ΔLOGINDPRO	0.09	−0.70	−0.41	−0.32	−0.68
	[1.0]	−[4.0]	−[2.2]	−[2.0]	−[4.2]
ΔLOGCPI	−0.02	0.91	0.23	1.08	1.12
	−[0.1]	[2.2]	[0.5]	[2.8]	[2.9]
Market controls	Yes	Yes	Yes	Yes	Yes
Equity factor controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.70	0.24	0.14	0.14	0.17
Panel B: Quarterly broker-dealer change in log leverage					
Intercept	0.00	0.03	0.02	0.03	0.04
	[1.4]	[5.0]	[4.1]	[6.0]	[8.4]
ΔLEV	−0.08	0.14	0.26	−0.18	0.07
	−[1.5]	[1.5]	[2.8]	−[2.2]	[0.8]
Market controls	Yes	Yes	Yes	Yes	Yes
Equity factor controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.77	0.17	0.33	0.26	0.20

The table reports monthly excess return regressions of the carry, defensive, momentum, value and combined characteristic long-top-quintile-short-bottom-quintile, constant-volatility factors onto (i) market excess returns for Treasuries, credit and equity; (ii) equity anomaly factors SMB, HML and UMD from Ken French's website and QMJ from Asness et al. (2014); (iii) macroeconomic variables: one-month change in log VIX, one-month change in log industrial production and one-month change in log CPI; (iv) broker-dealer log leverage change (deseasonalized as in Adrian et al. 2014).

We measure liquidity using bond issue size. We use this measure as it has the broadest coverage and has a high correlation with average daily trading volume (Hotchkiss and Jostova, 2007). Volume-based measures typically require coverage in TRACE, which is limited for the high-yield market because a large fraction of high-yield bonds are issued under 144A regulations and are not disseminated by TRACE before late 2014. We measure the investor population sophistication by the fraction of a bond which is owned by institutional investors, presumably better equipped than retail investors to avoid mistakes. We measure firm transparency by the number of analysts following the issuer equity (Hong et al., 2000) and, lastly, the cost of shorting is measured by the shorting fee score from MarkitDataExplorers.

Table 9 reports the results. If deviations from market efficiency and/or market frictions are empirically descriptive we would expect to see (i) lower analyst coverage, (ii) lower institutional ownership and (iii) smaller bonds in the extreme portfolios for each characteristic. We would also expect to see greatest short selling costs in the lowest portfolio for

each characteristic. Across all four characteristics we do not see any consistent evidence supporting these explanations. Bonds that score high in carry, momentum and value tend to be smaller, issued by more sparsely covered and owned by fewer institutions. Bonds, on the short side, however, display opposite rather than similar behavior. They are bigger, well covered and owned at a higher rate by institutional investors. Finally, the short side of each characteristic portfolio is in fact cheaper to borrow than its long side. Collectively, these patterns suggest that a simple mispricing hypothesis does not fit the data.

A separate implication of the mispricing hypothesis is that the relation between characteristics and future returns will be stronger among bonds in the segment of the corporate bond universe that are harder to arbitrage, less transparent or populated with a less sophisticated investor base. In Table 10, we test this hypothesis by comparing long-and-short portfolios formed on different universes of corporate bonds distinguished by the dimensions discussed above. We find that momentum long-and-short portfolios perform better in the less liquid, less transparent and less sophisticated segments of the corporate bond market. For carry, defensive and value the evidence is more mixed, sometimes performing better in the more mispricing prone arbitrage segments of the market, sometimes performing better in the less vulnerable one, but rarely statistically significant.

As a final test of behavioral explanations, we look at errors in expectations of equity analysts' sales forecasts. A systematic pattern between a characteristic and both future returns and revisions is consistent with mispricing (see, e.g., Bradshaw et al., 2001). We focus on sales instead of EPS because we are assessing senior claims. Increases in EPS can be detrimental for credit if the increase came about through releveraging. Our hypothesis is that analysts and investors have similar beliefs and, therefore, we can learn about the (unobservable) mistakes of the latter from the (observable) mistakes of the former. In other words, do the firms in the good carry, momentum, value and defensive portfolio experience more positive revisions than those in the bottom quintile of those characteristics?

To study analyst errors we focus on their forecast revisions for the next 12 months of sales. If analysts are fully rational and free from agency concerns, their revisions should not be predictable by any model relying on public information. If, on the other hand, their revisions are found to be predictable, it means they are ignoring certain information. To the extent that investors and analysts share the same beliefs, prices would not reflect that information as well and as a consequence would be predictable.

We build next-12-month revisions by averaging revisions over the next fiscal year sales number (FY1) and that of the subsequent fiscal year (FY2). We set the weights dynamically to assure that the weighted average horizon of the forecast is always 12 months.

The results are displayed in Figure 4. The dark lines are the cumulative revisions of the long portfolio and the dotted lines are the revisions for the short portfolio. If there are systematic predictable errors in sales forecasts, then we expect greater downward revisions for the short portfolio and greater upward revisions for the long portfolio. The mispricing hypothesis is consistent with what we see for the momentum portfolio and to a lesser extent for the defensive portfolio: after 12 months the long portfolio experiences sales revisions are larger than short by 5.7% (t-statistic = 7.28) and 1.2% (t-statistic = 1.86) for momentum and defensive, respectively. For value and carry the result goes in the opposite direction of that predicted by the mispricing hypothesis but the difference is only statistically

Table 9 Average Mispricing Susceptibility of Quintile Portfolios

	Carry	Defensive	Momentum	Value
Equity Analyst Coverage				
Low	19.36	14.87	15.19	16.48
40	15.13	13.90	15.20	15.61
60	12.46	14.50	15.04	14.47
80	9.42	14.95	14.59	12.98
High	7.64	15.41	12.88	11.37
High–low	−11.73	0.55	−2.31	−5.11
High–low t-statistic	−18.66	2.03	−4.87	−13.38
Issue market value in billions				
Low	1.28	1.34	1.06	1.10
40	1.04	0.87	0.98	1.05
60	0.79	0.85	0.93	0.93
80	0.66	0.72	0.85	0.74
High	0.54	0.66	0.80	0.71
High–low	−0.73	−0.69	−0.25	−0.39
High–low *t*-statistic	−17.24	−9.93	−3.81	−5.27
Institutional ownership of bond				
Low	0.50	0.48	0.49	0.51
40	0.54	0.52	0.50	0.49
60	0.50	0.50	0.50	0.48
80	0.46	0.47	0.49	0.49
High	0.40	0.47	0.46	0.49
High–low	−0.10	−0.02	−0.03	−0.02
High–low t-statistic	−5.18	−2.20	−3.50	−2.66
Average shorting cost score of bond				
Low	0.05	0.10	0.14	0.10
40	0.05	0.09	0.10	0.10
60	0.08	0.12	0.10	0.14
80	0.17	0.20	0.12	0.19
High	0.52	0.18	0.21	0.17
High–low	0.47	0.07	0.07	0.07
High–low *t*-statistic	9.79	7.73	4.90	3.70

For each characteristic and mispricing susceptibility proxy, the table reports the cap-weighted average value of the proxy for each one of the quintile portfolios formed on carry, momentum, value and defensive. The proxies for susceptibility are firm transparency measured by the number of equity analysts that follows the issuing company; liquidity as measures by bond issue size; investor base sophistication as measured by institutional ownership and easiness of shorting as measured by the shorting fee score (0 is lowest fee and 5 the highest). The table also displays the difference between the top and bottom quintile portfolios as well as its *t*-statistic computed using Newey–West standard errors and 18 lags.

Table 10 Average Returns of Characteristic Portfolios Across Bonds with Varying Levels of Mispricing Susceptibility

	Low	Medium	High	High-minus-low
Panel A: Characteristic returns across equity analyst coverage terciles				
Carry	0.35%	0.25%	0.15%	−0.20%
	[1.9]	[1.6]	[1.1]	−[1.8]
Defensive	0.14%	0.08%	0.17%	0.02%
	[2.0]	[1.2]	[2.6]	[0.4]
Momentum	0.40%	0.18%	0.16%	−0.23%
	[4.4]	[2.6]	[2.1]	−[2.8]
Value	0.42%	0.39%	0.22%	−0.20%
	[6.7]	[7.0]	[2.9]	−[2.0]
Panel B: Characteristic returns across institutional ownership terciles				
Carry	0.32%	0.30%	0.28%	−0.03%
	[1.6]	[1.5]	[1.8]	−[0.3]
Defensive	0.11%	0.10%	0.17%	0.05%
	[1.7]	[2.5]	[4.5]	[0.9]
Momentum	0.26%	0.15%	0.06%	−0.20%
	[3.1]	[2.1]	[1.2]	−[2.3]
Value	0.21%	0.29%	0.31%	0.10%
	[3.6]	[4.6]	[6.7]	[1.7]
Panel C: Characteristic returns across bond market value terciles				
Carry	0.28%	0.34%	0.23%	−0.06%
	[1.8]	[2.1]	[1.4]	−[0.6]
Defensive	0.14%	0.15%	0.11%	−0.02%
	[2.8]	[4.3]	[2.7]	−[0.4]
Momentum	0.37%	0.15%	0.12%	−0.25%
	[6.2]	[3.4]	[2.0]	−[3.4]
Value	0.28%	0.29%	0.19%	−0.08%
	[5.8]	[7.2]	[4.4]	−[1.4]
Panel D: Characteristic returns across of shorting costs				
	Fee = 0	Fee > 0	Fee 0 minus Fee > 0	
Carry	0.36%	0.35%	−0.01%	
	[1.3]	[1.0]	[0.0]	
Defensive	0.16%	0.52%	0.36%	
	[2.2]	[1.8]	[1.3]	
Momentum	0.06%	0.08%	0.02%	
	[0.6]	[0.2]	[0.1]	
Value	0.32%	0.53%	0.21%	
	[4.4]	[2.1]	[0.9]	

The table displays the average credit excess returns and *t*-statistics of long-and-short characteristic portfolios built from subsets of firms with different values for mispricing susceptibility measures. The proxies for susceptibility are firm transparency measured by the number of equity analysts that follows the issuing company; liquidity as measures by bond issue size; investor base sophistication as measured by institutional ownership and easiness of shorting as measured by the shorting fee score (0 is lowest fee and 5 the highest).

significant for carry (t-statistic = 6.17). As a consequence mispricing seems to play less of an obvious role for these two characteristics.

While the revision number for momentum is a sizable 5.7% sales drop, it is hard to evaluate its impact on the long-and-short portfolio return. To facilitate interpretation we compute the impact on credit spreads of a 5.7% sales drop for the median firm using a simple structural model. The median bond in our sample has an OAS of 302 bps and duration of 5, while its issuer has a leverage of 0.28. We feed those numbers through a structural model (Merton, 1974) to invert the asset volatility that is consistent with this quantity—the credit-implied volatility (Kelly et al., 2016). We then shock asset value by 5.7% assuming that it drops by the same value as sales and, keeping volatility constant, compute the new credit spread and the credit returns associated with this change. For momentum, a 5.7% sales increase translates into a roughly 112

Figure 4 Analysts revisions of top and bottom quintile portfolios formed on different characteristics (January 2001–April 2015). Average cumulative monthly revisions of analysts sales forecasts for the next 12 months since portfolio formation for issuers in the top and bottom quintiles of the four characteristics: carry, defensive, momentum and value as defined in the text. For every firm, analyst forecasts revisions for the next 12 months are built from an average of the revisions (log difference) of forecasts for the next fiscal year (FY1) and the following (FY2), with weights set to make the average horizon be 12 months. For every portfolio, the revision number is an equal weight of the revisions of all the firms that comprise that portfolio.

bps return: about one-third of the 290 bps momentum premium displayed in Table 4. Errors in expectations offer a partial explanation to the momentum returns in corporate bond markets.

5 Conclusion

We undertake a comprehensive analysis of the cross-sectional determinants of corporate bond excess returns. We find strong evidence of positive risk-adjusted returns to measures of carry, defensive, momentum and value. These returns are diversifying with respect to both known sources of market risk (e.g., equity risk premium, credit risk premium and term premium) and characteristic returns that have been documented in equity markets (e.g., size, value and momentum). These conclusions hold whether one examines traditional long-and-short academic portfolios or a long-only, transactions-costs aware portfolio. The latter helps dismiss the hypothesis that the returns are not economically significant.

In our final analysis we examine the source of the value, momentum, carry and defensive premiums in credit. We investigate risk and mispricing explanations. We do not find evidence that the anomalies earn their premiums through traditional risk exposures or to shocks to financial intermediaries' balance sheets—characteristic returns tend to be a hedge to traditional macroeconomic factors and exhibit mostly insignificant loadings on shocks to broker-dealers' leverage. Mispricing evidence is strongest for momentum: the momentum strategy has better performance among less liquid bonds issued by less transparent firms and owned by less sophisticated investors; it is also long (short) bonds of firms where analyst forecasts of sales are relatively too pessimistic (optimistic). The evidence for mispricing is mixed for the other characteristics.

Appendix

Table A.1 Variable Definitions

Variable	Definition
Duration Total	Option-adjusted duration as reported by BAML.
return	Monthly total return on the corporate bond, inclusive of coupons and accrued interest.
Excess return	Monthly excess return on the corporate bond, computed as the difference between the monthly total return on the corporate bond and the monthly return of a duration-matched U.S. Treasury bond.
Amt. Out.	The face value of the corporate bond measured in USD millions.
Time to maturity	Number of years before bond matures.
Age percent	Fraction of bond life that has expired (time since issuance divided by original maturity).
Rating	Standard & Poor's issuer-level rating, coded from 1 (AAA) to 10 (D).
Market beta	Slope from l2-month rolling regression of credit excess returns on the credit market excess return (see CREDIT below).

(*Continued*)

Table A.1 (*Continued*)

	Variable	Definition
Carry	OAS	Option-adjusted spread as reported in the Bank of America Merrill Lynch (BAML) bond database.
Value	Empirical	The residual from a cross-sectional regression of the log of OAS onto the log of duration, rating and bond excess return volatility (12 month).
	Structural	The residual from a cross-sectional regression of the log of OAS onto the log of the default probability implied by a structural model (Shumway, 2001).
Momentum	Credit Equity	The most recent six-month cumulative corporate-bond excess return. Equity momentum, defined as the most recent six-month cumulative issuer equity return.
Defensive	Leverage	Market leverage, measured as the ratio of net debt (book debt + minority interest + preferred stocks − cash) to the sum of net debt and market capitalization. Measured using data available at the start of each month (assuming a six-month lag for the release of financial statement information).
	Duration	Effective duration as reported in the Bank of America Merrill Lynch (BAML) bond database.
	Profitability	Gross profits over assets.
	CONST_VOL_CHAR	Credit excess returns of a characteristic portfolio that goes long bonds in the top characteristic quintile and short those in the bottom. Every month the portfolio is scaled to have an ex-ante volatility of 5%, where the ex-ante volatility is the realized volatility over the last 2 years.
	TSY	Excess returns to long-term government bonds, measured as the difference between monthly total returns on the Bank of America Merrill Lynch U.S. Treasuries 7–10 year index and one-month U.S. Treasury bills.
	CREDIT	Excess returns to corporate bonds, measured as the difference between the value-weighted monthly total returns of corporate bonds included in the BAML dataset and a portfolio of duration-matched U.S. Treasury bonds.
	EQUITY	Excess returns to the S&P 500 Index, measured as the difference between monthly total returns to the S&P 500 and one-month U.S. Treasury bills.
	SMB	Monthly mimicking-factor portfolio return to the size factor, obtained from Ken French's website.
	HML	Monthly mimicking-factor portfolio return to the value factor, obtained from Ken French's website.
	UMD	Monthly mimicking-factor portfolio return to the momentum factor, obtained from Ken French's website.
	QMJ	Monthly mimicking-factor portfolio return to the quality factor, obtained from AQR's library website.
	ΔLOGVIX	One-month change in log VIX (VIXCLS) from FRED website.

(*Continued*)

Table A.1 (*Continued*)

Variable	Definition
ΔLOGINDPRO	One-month change in log seasonally adjusted industrial production index (INDPRO) from FRED website.
ΔLOGCPI	One-month change in log seasonally adjusted CPI (CPIAUCSL) from FRED website.
ΔLEV	Change in log broker-dealers $leverage = \frac{financial\ assets}{financial\ assets - total\ liabilities}$. From FED's flow of funds data and seasonally adjusted (Adrian et al., 2014)
Equity analyst coverage	The number of analysts covering the issuer equity, from I/B/E/S.
Credit institutional ownership	Fraction of bond amount outstanding owned by non-retail investors, from Lipper emaxx database.
Bond shorting score	Score between 0 and 5 from Markit: 0 represents lowest cost to borrow and 5 represents the highest.
Equity analyst revisions in sales expectations	Weighted average change in the next fiscal year (FYl) and the one after (FY2) sales forecasts. The weights are chosen such that the forecast refers to a number that is on average 12 months into the future.

Acknowledgement

A previous version of this paper was titled "Investing with style in corporate bonds." We thank Demir Bektic, Maria Correia, Wayne E. Ferson, Patrick Houweling, Antti Ilmanen, Sarah Jiang, Toby Moskowitz, Narayan Naik, Lasse Pedersen, Kari Sigurdsson and participants at the 4th Alliance Bernstein Quant Conference, 24th European Pensions Symposium, 7th Inquire UK Business School Meeting, Norwegian Ministry of Finance and University of Cambridge, 2016 SFS Finance Cavalcade, London Quant Group, University of Notre Dame and UBS Quantitative Investment Conference 2016 for their helpful discussion and comments. We acknowledge the outstanding research assistance of Peter Diep, Johnny Kang and Mason Liang. The views and opinions expressed herein are those of the authors and do not necessarily reflect the views of AQR Capital Management LLC ("AQR"), its affiliates, or its employees. This information does not constitute an offer or solicitation of an offer, or any advice or recommendation, by AQR, to purchase any securities or other financial instruments and may not be construed as such.

Endnotes

[1]If one of the standardized measures is missing, we assign a zero score such that the combination will have a non-missing score for the union of names, which have at least one non-missing score.

[2]Between January 1997 and December 1998, we set the scalar equal to its value as of January 1999.

References

Adrian, T., Etula, E., and Muir, T. (2014). "Financial Intermediaries and the Cross-Section of Asset Returns," *The Journal of Finance* 69(6), 2557–2596.

Altman, E. I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance* 23(4), 189–209.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). "The Distribution of Realized StockReturn Volatility," *Journal of Financial Economics* 61(1), 43–76.

Asness, C. S., Andrea, F., and Lasse, H. P. (2014). "Quality Minus Junk," Working Paper.

Asness, C., A., Ilmanen, R. I., and Moskowitz, T. (2015). "Investing with Style," *Journal of Investment Management* 13, 27–63.

Asness, C., Moskowitz, T., and Pedersen, L. (2013). "Value and Momentum Everywhere," *Journal of Finance* 68, 929–985.

Asquith, P., Au, A. S., Covert, T., and Pathak, P. A. (2013). "The Market for Borrowing Corporate Bonds," *Journal of Financial Economics* 107, 155–182.

Asvanunt, A. and Richardson, S. (2017). "The Credit Risk Premium," *Journal of Fixed Income,* Winter 2017.

Ben Dor, A., Dynkin, L., Hyman, J., Houweling, P., Van Leeuwen, E., and Penniga, O. (2007). "DTS (Duration Times Spread)," *The Journal of Portfolio Management* 33, 77–100.

Bessembinder, H., Maxwell, W., and Venkataraman, K. (2006). "Market Transparency, Liquidity Externalities, and Institutional Trading Costs in Corporate Bonds," *Journal of Financial Economics* 82, 251–288.

Bharath, S. T. and Shumway, T. (2008). "Forecasting Default with the Merton Distance to Default Model," *Review of Financial Studies* 21(3), 1339–1369.

Binsbergen, J. H. and Koijen, R. S. J. (2015). "The Term Structure of Returns: Facts and Theory," Working Paper.

Bradshaw, M. T., Richardson, S. A., and Sloan, R. G. (2001). "Do Analysts and Auditors Use Information in Accruals?," *Journal of Accounting Research* 39(1), 45–74.

Carvalho, R., Dugnolle, P., Xiao, L., and Moulin, P. (2014). "Low-Risk Anomalies in Global Fixed Income: Evidence from Major Broad Markets," *The Journal of Fixed Income* 23, 51–70.

Chen, L., Lesmond, D. A., and Wei, J. (2007). "Corporate Yield Spreads and Bond Liquidity," *Journal of Finance* 62, 119–149.

Chen, N. F., Roll, R., and Ross, S. A. (1986). "Economic Forces and the Stock Market," *Journal of Business* 383–403.

Choi, J. and Kim, Y. (2015). "Anomalies and Market (Dis)integration," Working Paper, UIUC.

Chordia, T., Goyal, A., Nozawa, Y., Subrahmanyam, A., and Tong, Q. (2016). "Are Capital Market Anomalies Common to Equity and Corporate Bond Markets?," Working Paper, Emory University.

Cochrane, J. H. (2009). *Asset Pricing: Revised Edition*. Princeton University Press.

Correia, M., Richardson, S., and Tuna, I. (2012). "Value Investing in Credit Markets," *Review of Accounting Studies* 17(3), 572–609.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). "Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?," *Review of Financial Studies* 22(5), 1915–1953.

Edwards, A. E., Harris, L. E., and Piwowar, M. S. (2007). "Corporate Bond Market Transaction Costs and Transparency," *Journal of Finance* 62, 1421–1451.

Fama, E. F. and Kenneth, R. F. (1993). "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* 33(1), 3–56.

Frazzini, A., Israel, R., and Moskowitz, T. J. (2012). "Trading Costs of Asset Pricing Anomalies," Working Paper, AQR Capital Management.

Frazzini, A. and Pedersen, L. H. (2014). "Betting Against Beta," *Journal of Financial Economics* 111, 1–25.

Gebhardt, W. R., Hvidkjaer, S., and Swaminathan, B. (2005a). "The Cross-Section of Expected Corporate Bond Returns: Betas or Characteristics?," *Journal of Financial Economics* 75, 85–114.

Gebhardt, W. R., Hvidkjaer, S., and Swaminathan, B. (2005b). "Stock and Bond Market Interaction: Does Momentum Spill Over?," *Journal of Financial Economics* 75, 651–690.

Grossman, S. J. and Stiglitz, J. E. (1980). "On the Impossibility of Informationally Efficient Markets," *The American Economic Review* 70(3), 393–408.

Haesen, D., Houweling, P., and van Zundert, J. (2013). "Residual Equity Momentum for Corporate Bonds," Working Paper, Robeco Quantitative Strategies.

Hallerbach, W. G. and Houweling, P. (2013). "Ibbotson's Default Premium: Risky Data," *The Journal of Investing* 22, 95–105.

Harris, L. (2015). "Transaction Costs, Trade Throughs, and Riskless Principal Trading in Corporate Bond Markets," Working Paper, USC.

He, Z. and Krishnamurthy, A. (2013). "Intermediary Asset Pricing," *The American Economic Review* 103(2), 732–770.

Hong, H., Lim, T., and Stein, J. C. (2000). "Bad News Travels Slowly: Size, Analyst Coverage and the Profitability of the Momentum Strategies," *The Journal of Finance* 55(1), 265–295.

Hotchkiss, E. S. and Jostova, G. (July 9, 2007). "Determinants of Corporate Bond Trading: A Comprehensive Analysis," Working Paper.

Houweling, P. and van Zundert, J. (2014). "Factor Investing in the Corporate Bond Market," Working Paper, Robeco Quantitative Strategies.

Jostova, G., Nikolova, S., Philipov, A., and Stahel, C. W. (2013). "Momentum in Corporate Bond Returns," *Review of Financial Studies* 26(7), 1649–1693.

Kelly, B. T., Manzo, G., and Palhares, D. (2016). "Credit-Implied Volatility," Manuscript.

Koijen, R., Moskowitz, T., Pedersen, L., and Vrugt, E. (2014). "Carry." Working Paper, University of Chicago Booth School of Business, New York University, University of Amsterdam.

Kwan, S. H. (1996). "Firm-Specific Information and the Correlation Between Individual Stocks and Bonds," *Journal of Financial Economics* 40, 63–80.

Lewellen, J. (2015). "The Cross-Section of Expected Stock Returns," *Critical Finance Review* 4, 1–44.

Lok, S. and Richardson, A. (2011). "Credit Markets and Financial Information," *Review of Accounting Studies* 16, 487–500.

Merton, R. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance* 29, 449–470.

Moreira, A. and Muir, T. (2016). "Volatility Managed Portfolios (No. w22208)," *National Bureau of Economic Research.*

Ng, K. Y. and Phelps, B. (2014). "Structure of U.S. Corporate Excess Returns: The Hunt for a 'Low-Risk' Anomaly," Working Paper, Barclays.

Novy-Marx, R. (2013). "The Other Side of Value: The Gross Profitability Premium," *Journal of Financial Economics* 108(1), 1–28.

Palhares, D. (2013). "Cash-Flow Maturity and Risk Premia in CDS Markets," Working Paper, AQR Capital Management.

Schaefer, S. M. and Strebulaev, I. A. (2008). "Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds," *Journal of Financial Economics* 90(1), 1–19.

Shleifer, A. and Vishny, R. W. (1997). "The Limits to Arbitrage," *The Journal of Finance* 52(1), 35–55.

Shumway, T. (2001). "Forecasting Bankruptcy More Accurately: A Simple Hazard Model," *Journal of Business* 740, 101–124.

This page intentionally left blank

Size Matters, If You Control Your Junk*

Clifford Asness, Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen

The size premium has been accused of having a weak historical record, being meager relative to other factors, varying significantly over time, weakening after its discovery, being concentrated among microcap stocks, residing predominantly in January, relying on price-based measures, and being weak internationally. We find, however, that these challenges disappear when controlling for the quality, or its inverse, junk, of a firm. A significant size premium emerges, which is stable through time, robust to specification, not concentrated in microcaps, more consistent across seasons, and evident for non-price-based measures of size, and these results hold in 30 different industries and 24 international equity markets. The resurrected size effect is on par with anomalies such as value and momentum in terms of economic significance and gives rise to new tests of, and challenges for, existing asset pricing theories.

* We thank David Blitz, Mark Carhart, Kent Daniel, Ken French, April Frieda, Jacques Friedman, Nicolae Garleanu, Will Goetzmann, Antti Ilmanen, Sarah Jiang, Don Keim, Bryan Kelly, Christine Kemper, John Liew, Robert Novy-Marx, Scott Richardson, Richard Roll, Laura Serban, Rodney Sullivan, Eric Wu, and seminar participants at Dartmouth College, the FTSE Russell World Investment Forum, the Jacobs Levy Center Conference, the Chicago Booth Asset Pricing Conference, the Q Group 2016 Spring Seminar, London Business School, London School of Economics, and the 2017 American Finance Association meetings (Chicago, Illinois) for helpful comments. We also thank Alex Bennett, Brandon Lee, Xiao Qiao, Kaushik Vasudevan, and Ryan Wei for outstanding research assistance. Moskowitz thanks the Center for Research in Security Prices for financial support when this project began. AQR Capital Management is a global investment manager who may or may not use the ideas in this paper. The views and opinions expressed are those of the authors and do not necessarily reflect the views of AQR Capital Management.

https://doi.org/10.1016/j.jfineco.2018.05.006 0304-405X/© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Reprinted from the *Journal of Financial Economics*, 129 (2018) 479–509, with authors' permission.

1. Introduction

Does size matter? With respect to capital markets, the answer to this question is unclear. Academic research on the relation between firm size and expected returns dates back to at least Banz (1981), who finds that small stocks in the US (those with lower market capitalizations) have higher average returns than large stocks, an effect not accounted for by the higher market beta of small stocks. The relation between firm size and expected returns is important for several reasons. First, the size effect has become a focal point for discussions about market efficiency. Second, a size factor has become one of the main building blocks of current asset pricing models used in the literature and in practice (e.g., Fama and French, 1993, 2016). Third, the size premium implies that small firms face larger costs of capital than large firms, having important implications for corporate finance, incentives to merge and form conglomerates, and broader industry dynamics. Fourth, the size effect has had a large impact on investment practice (Reinganum, 1983a), including spawning an entire category of investment funds, giving rise to small cap indices, and serving as a cornerstone for money management classification.

We provide new evidence on the size effect and test several competing theories for its existence:

1. *Risk-based theories of frictionless capital markets.*
   (a) Standard asset pricing models such as the capital asset pricing model (CAPM): If size per se is not a risk, standard models predict that size does not matter when controlling for risk exposures.
   (b) Size captures time-varying risk premia: Size can be correlated to expected returns only because size is measured by market value, which is influenced by risk premia (Ball, 1978; Berk, 1995a). Riskier firms have higher required returns, leading to lower market value, everything else equal. Hence, any misspecification of risk premia, due perhaps to time-varying risk or risk premia, will be picked up by market prices. According to this theory, size factors based on market prices will mechanically pick up these movements, but size measures other than market value should not predict returns.
   (c) Theories of growth options: If small firms have more growth options and growth options are risky, then small firms are riskier and have higher required returns (Carlson et al., 2004; Garleanu et al., 2012). Hence, the size effect should be smaller when controlling for measures of risk and growth options.
2. *Behavioral finance theories*. Small firms are more difficult to arbitrage, making these firms more mis-priced (Shleifer and Vishny, 1997). If this effect drives the size effect, then small stocks need to be underpriced on average relative to large stocks, and the size effect should be smaller when controlling for measures of mispricing such as investor sentiment, disagreement, and limited arbitrage activity (e.g., trading costs or short-sale costs). Investors could overreact to growth and other quality measures; so according to these theories, the size effect should be smaller when we control for such effects (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999).
3. *Theories of liquidity and liquidity risk*. Size matters because small firms are less liquid (Amihud and Mendelson, 1986) and face more liquidity risk (Acharya and Pedersen, 2005), both requiring higher expected returns. Hence, the size premium should be related to liquidity level and risk measures, and the size and liquidity premia should be more evident when controlling for other risk exposures, especially those negatively correlated with size.

To test these theories, we consider the size effect controlling for other factors, which can proxy for other sources of risk, growth, mispricing, and liquidity. We focus on the interaction between size and firm quality (or its inverse, junk). The interaction between size and quality is especially interesting for three reasons. First, quality can be defined as a characteristic of an asset that, all else equal, commands a higher price. As such, size, which is based on market values, should have a strong connection to quality. Second, Novy-Marx (2013) shows that quality, as measured by profitability, has a strong connection to the value effect and helps clean up the relation between value and expected returns. A similar argument can be made for the size effect, where size's relation to average returns can be clearer once we account for quality. Because size also interacts with value (Fama and French, 1993; 2012), sorting out the interactions between these three predictors of returns could help better characterize the cross section of expected returns. Third, quality has a direct theoretical link to some of the theories we aim to test, namely, quality can be related to growth options underlying rational theories for size or sources of overreaction underlying prominent behavioral theories, or both. Measures of quality such as profitability, stability, and growth are intuitively empirical proxies for growth options and have been used as variables related to overreaction by investors.

Supporting these motivations for looking at the size-quality interaction, we find empirically that the key control variable for the size effect is firm quality (measured by profitability, stability, growth, and safety) or its inverse, junk. The interaction between size and quality or junk is far stronger than size's interaction with other factors (beta, value, momentum) and accounting for it produces a more significant size premium that helps test some of the competing theories for size.

Controlling for quality or junk (along with the standard asset pricing factors such as value and momentum) resurrects a strong size effect and helps distinguish among the competing theories. The resurrected strong size effect controlling for other factors can be viewed as a rejection of theory 1(a). We find that non-price-based size measures perform just as well as price-based ones, rejecting theory 1(b). The fact that controlling for quality, including growth, makes the size effect stronger, not weaker, seems inconsistent with the growth theory 1(c), if growth options are more prevalent among the high growth firms as theory predicts. We find that small stocks have higher shorting costs and more disagreement, which, according to behavioral theories, makes them more likely to be overvalued not undervalued, which is inconsistent with behavioral theories 2. Finally, the fact that size matters more when controlling for other factors is consistent with liquidity-based theories 3, in which controlling for these other factors, particularly quality, helps clean up the relation between size, liquidity, and average returns. However, this result does not offer any additional direct evidence in favor of a liquidity story.

Given the importance and prominence of the size effect as the first major challenge to the CAPM, it has naturally come under heavy and often critical scrutiny. Considering almost a century-long sample of US stocks and a broad sample of global stocks in 24 different markets, we confirm the main criticisms of the standard size effect. Consistent with risk-based theories 1, size has a weak historical record in the US, especially after adjusting for market risk (e.g., the CAPM). It suffers from long periods of poor performance, particularly after its discovery in the 1980s.[1] It has an even weaker record internationally (Crain, 2011 and Bryan, 2014). What little size effect is present seems to be concentrated in difficult to invest in microcap stocks, whose returns occur only in January,[2] making it

arguably fairly insignificant. Based on this evidence, a simple risk-based model such as the CAPM, in which size per se is not a risk factor, seems broadly consistent with the data.

However, we find that measures of size studied by the literature load strongly and consistently negatively on a large variety of what have come to be termed "quality factors." At a broad level, quality is a characteristic, or set of characteristics, of a security that investors should in theory be willing to pay a high price for, all else equal. A high-quality firm can be one that is well-managed and has strong economic and accounting performance, such as high profitability and stability of earnings, good growth prospects, and low risk.[3]

Regardless of the quality metric used, for metrics that vary substantially both qualitatively and in terms of measured correlation, we find a much stronger and more stable size effect when controlling for a firm's quality or its inverse, junk. Controlling for quality, using any of a broad set of measures, reconciles many of the empirical irregularities associated with the size premium that have been shown in the literature. In short, firm size is highly confounded with firm quality, which distorts the relation between size and expected returns. Large firms tend to be high-quality firms on any of the above dimensions or measures, while small firms tend to be "junky" (i.e., have the opposite characteristics). Given that the literature shows that high-quality stocks, however defined, tend to outperform junk stocks, including when comparing stocks of similar size (Asness et al., 2014; Fama and French, 2016), this means that the basic size effect is fighting a strong quality effect. By going long small stocks and short large stocks, a size-based strategy is long a potential size premium but also short a quality premium, which both understates the actual size effect and introduces additional variation from another factor.

We show that controlling for quality, a significant and more robust size effect emerges. Small quality stocks significantly outperform large quality stocks, and small junk stocks outperform large junk stocks, but the basic size effect suffers from a size-quality composition effect. We show that controlling for quality does even more than simply raise the size premium. Accounting for the influence of quality also explains all of the many empirical irregularities and challenges researchers have identified with the size effect. Controlling for any number of a variety of quality measures with no notable failures (some used in the literature, others novel) not only resuscitates the overall size effect by more than doubling its Sharpe ratio, but also resurrects it in the 1980s and 1990s when it is otherwise conspicuously absent, restores a more linear relation between size and average returns (i.e., no longer concentrated among the tiniest firms), revives the returns to size outside of January while simultaneously diminishing the returns to size in January (making it more uniform across months), and uncovers a larger size effect in almost two dozen other international equity markets (where size has been notably weak). This stronger size effect controlling for other factors, including the market and other risk-based factors, is inconsistent with standard theories of asset pricing. We next consider theory 1(b) that size matters only because it is measured by market capitalization, which contains market prices. The argument is that any misspecification in the asset pricing model, such as time-varying risk premia, is likely to show up in a cross-sectional relation between any market-based measure containing price and returns. Consistent with this argument, Berk (1995b, 1997) shows that using non-price-based measures of size does not yield a significant relation between size and average returns. Hence, no size effect could exist per se, but rather model misspecification may be showing up in price-based measures. To test this argument, we construct size factors in which stocks are sorted based on measures that do not include the stock price. We construct five new size factors based on book assets, sales, book equity, number of employees, and

fixed assets (property, plant, and equipment, PP&E). For each of these size factors, we find a significant alpha when controlling for quality and other factors, rejecting theory 1(b).

Our evidence also does not support the theories of growth options, 1(c). Controlling for growth, which is included among our quality measures, and other risk factors makes the size effect stronger, not weaker. Although we cannot perfectly measure growth options and their risks, any mismeasurement of growth options likely has the opposite theoretic effect on the size premium to be consistent with our results. Hence, our overall evidence on size appears to challenge many risk-based theories of frictionless capital markets.

We next turn to the behavioral theories, 2. For these theories to explain the size effect, small stocks would not simply have to be mispriced, but more importantly, small stocks would have to be *underpriced* in the sense of being too cheap relative to large stocks. A key behavioral finance prediction related to certain stocks being cheap versus expensive is limits to arbitrage, in particular, theories of short-selling constraints following Miller (1977). According to these theories, certain stocks are difficult to sell short, implying that these stock prices mostly reflect the opinions of optimists, leading such stocks to become overvalued. Hence, these theories suggest a negative instead of a positive size premium, where small stocks are more likely to be overvalued rather than undervalued, relative to large stocks. Further, this effect should be stronger when differences of opinion among investors are larger, which also tends to go the opposite way of the size effect.

To test the behavioral theories, we examine the levels of short selling, the costs of short selling, the degree of dispersion among analysts' forecasts of earnings, and the degree of mispricing of small cap stocks relative to large cap stocks. To explain the size effect with a limit of arbitrage argument, the costs of short selling should be lower for small cap stocks relative to large cap stocks. We find evidence to the contrary. That is, small stocks have larger shorting costs in each quality group. This finding, while not surprising, implies that a negative small cap expected return should be evident, the opposite of a small cap premium, according to behavioral theories. In addition, we find evidence to support greater differences of opinion among small cap stocks, which, again, should lead to lower returns, not higher returns, for small cap stocks. Finally, when looking at past five-year returns, we find the returns to small cap stocks to be low, not high, providing further evidence that small stocks do not seem to be overvalued as predicted by behavioral theories. When we control for variables that could be related to mispricing or risk premia, such as value, quality, and momentum, the size effect gets stronger, not weaker, which is also inconsistent with a behavioral story for the size premium.

Lastly, we consider the liquidity-based theories, 3. The idea is that size can just be a proxy for illiquidity and liquidity risk, and investors generally require compensation for holding illiquid securities facing the risk of worsening liquidity (here we refer to attempts to measure liquidity itself and a premium for illiquidity, which is more than simply noting that the size effect is largely in microcaps and more difficult to arbitrage). Consistent with these theories, the returns to size seem to be captured by measures of illiquidity suggested by Brennan and Subrahmanyam (1996), Amihud (2002), Hou and Moskowitz (2005), Sadka (2006), and Ibbotson et al. (2013), and measures of liquidity risk (the covariance with changes in liquidity) such as those of Pastor and Stambaugh (2003) and Acharya and Pedersen (2005). Crain (2011) summarizes this evidence.

We study the link between liquidity and size when controlling for quality. We find a difference in bid-ask spreads between small and large stocks (as in the literature), but, more importantly, we find that this difference is similar across each quality quintile.

In other words, small quality stocks are less liquid than large quality stocks and like-wise for junk stocks. We also estimate that small stocks have greater market impact costs than large stocks and this difference is similar across quality quintiles. Hence, the return spread between small versus large stocks lines up with the corresponding liquidity spreads across small versus large stocks, after controlling for quality. However, controlling for size, little relation seems to exist between liquidity and quality measures. High-quality small stocks face similar liquidity to junky small stocks (among large stocks it is the same). This is consistent with liquidity-based theories for the size premium, in which size is also correlated with a quality factor that is unrelated to liquidity and so the size-liquidity relation could be partly obscured by quality. We find another sign of illiquidity, that the size factor loads on lagged market returns, consistent with non-synchronous trading for small, illiquid stocks, and that this lagged market exposure is the same whether we control for quality or not. Hence, size seems to be related to both illiquidity (positively) and quality (negatively), but liquidity and quality are not strongly related. Therefore, our results are consistent with the size premium being an illiquidity premium, though they do not offer any additional evidence for their connection.

The size factor is also highly correlated to factors that attempt to capture liquidity premia more directly. When we regress size on the liquidity risk factor IML (illiquid minus liquid) from Amihud (2014), we find a highly significant loading (t-statistic above 30), consistent with the idea that the size premium is (at least partly) explained by compensation for liquidity risk. The size factor also loads on an alternative liquidity risk factor based on bid-ask spreads and turnover. Controlling for these liquidity risk factors naturally lowers the alpha of the size factor, even after controlling for quality, to a marginally positive effect (t-statistic of 2.03). Size loading on liquidity is similar whether or not we control for quality, indicating that liquidity risk helps explain the size premium, but liquidity risk is relatively unrelated to quality. A theory consistent with these facts is that stocks with higher liquidity risk have higher required returns and, separately, quality stocks have higher expected returns, but these return premia can be driven by different mechanisms. Our results show that quality helps resurrect a premium associated with small stocks, whether that premium is a size effect or a liquidity effect. Because small stocks tend to be junky, the standard size factor underestimates this premium, which is more clearly seen once we control for quality. If purely driven by liquidity risk, then size should have zero alpha when we control for both quality and liquidity risk. Hence, the marginally significant alpha suggests that there is more going on, our attempt to measure liquidity risk contains error, the result itself is just noise (the t-statistic barely passes conventional significance), or the size effect is compensation for the level of both liquidity and liquidity risk, when we capture only the latter. The model of Acharya and Pedersen (2005) suggests that small stocks have higher required return, everything else equal (e.g., for equal quality), because small stocks are both less liquid on average and face more liquidity risk in the sense that their liquidity deteriorates more when investors value liquidity the most.

Our results resurrect the size premium, putting it on a more equal footing with other anomalies such as value and momentum in terms of its efficacy. We motivate and evaluate our analysis in the context of theory, interpreting our results through the lens of rational, behavioral, and liquidity-based theories. We motivate why quality in particular is an interesting characteristic to interact with size in addressing these theories, but we also conclude that the interaction between size and quality presents a set of new stylized facts that seem to challenge standard frictionless asset pricing theories, are not easily explained

by existing behavioral theories, and seem most consistent with theories of liquidity and liquidity risk.

The paper proceeds as follows. Section 2 describes the data and reviews the evidence on the size effect, highlighting the challenges to the size premium identified in the literature. As a first test of the competing theories, Section 3 shows that the size effect is resurrected when controlling for a firm's quality or junk. Section 4 tests further distinguishing predictions of the risk-based, behavioral, and liquidity-based theories. Section 5 concludes.

2. Data and Preliminaries: Reexamining the Size Effect

We detail the data used in this study and reexamine the evidence of the basic size effect by replicating some of the challenges identified in the literature using an updated sample.

2.1. Data

We examine long–short equity portfolios commonly used in the literature based on firm size. For US equities, we obtain stock returns and accounting data from the union of the Center for Research in Security Prices (CRSP) tapes and the Compustat North America database. Our US equity data include all available common stocks on the merged CRSP and Compustat data (sharecode 10 or 11) between July 1926 and December 2012, our longest historical sample. We include delisting returns when available in CRSP.

For size portfolios, we primarily use the Fama and French SMB (small minus big) factor and a set of value-weighted decile portfolios based on market capitalization sorts, obtained from Ken French's webpage (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ data _ library.html). The decile portfolios are formed by ranking stocks every June by their market capitalization (price times shares outstanding) and forming deciles based on NYSE breakpoints, in which the value-weighted average return of each decile is computed monthly from July to June of the following year. The size factor, SMB, is the average return on three small portfolios minus the average return on three big portfolios formed by ranking stocks independently by their market cap and their book-to-market equity ratio (BE/ME) every June and forming two size portfolios using the NYSE median size and three book-to-market portfolios using 30%, 40%, and 30% breakpoints, value, middle, and growth respectively. The intersection of these groups forms six size and BE/ME portfolios split by small and large (e.g., small value, small middle, small growth and large value, large middle, and large growth), in which SMB is the equal-weighted average of the three small portfolios minus the equal-weighted average of the three large portfolios.

In addition to SMB, the value factor, HML (high minus low), is formed from the equal-weighted average return of the two value portfolios minus the two growth portfolios, HML = ½ (small value + big value) – ½ (small growth + big growth). Fama and French (1993) also add the market factor, RMRF, which is the value-weighted index of all CRSP-listed securities minus the one-month Treasury bill rate.

Ken French's website also provides a momentum factor, which is a long–short portfolio constructed in a similar manner, in which six value-weighted portfolios formed on size and prior returns (the cumulative return from months t-12 to t-2) are used. The portfolios are the intersections of two portfolios formed on size and three portfolios formed on prior

returns. The momentum factor, UMD (up minus down) is constructed as UMD = ½ (small up + big up) – ½ (small down + big down).

Ken French's website also provides a short-term reversal factor, STREV, which is formed in the same way as the momentum factor except using past returns from just the most recent month *t*-1 instead of *t*-12 to *t*-2, and with stocks sorted by the negative of their past one-month return.

We also form SMB and value-weighted size decile portfolios using non-price-based measures of size, as suggested by Berk (1995b, 1997), in lieu of a firm's market capitalization to rank stocks. Using the same methodology, we form five sets of non-price size portfolios based on book value of assets, book value of equity, sales, PP&E, and number of employees.

We also form SMB portfolios within each of 30 industries used by Fama and French (1997) and available on Ken French's website. We construct SMB in a similar fashion within each industry so that we obtain 30 SMB industry-neutral portfolios.

For shorting costs and levels of shorting, we use data from Markit Data Explorers, which provides data on more than $15 trillion of global securities in the lending programs of more than 20 thousand institutional funds. We focus on the supply of lendable shares divided by the market cap for each stock, expressed as a percentage, and the daily cost of borrow score (DCBS), a number from one to ten indicating the fee for each stock. Our sample uses data from January 2010 to October 2016, when coverage is highest for our universe of stocks.

We use a variety of quality or junk measures proposed in the literature that pertain to different notions of quality: profitability, growth, safety, payout, credit, and investment. We use some alone and some in composites. We use as different determinants of quality the profitability and investment measures of Fama and French (2016); the profitability, growth, safety, and payout measures of Asness et al. (2014) as well as their composite index of quality, which is an average of these measures; a revised composite index of Asness et al. (2017) that excludes payout and drops accruals from the growth composite; a safe-minus-risky (SMR) factor based off of the betting-against-beta (BAB) factor from Frazzini and Pedersen (2014) — for simplicity we will refer to this factor throughout the paper as simply BAB;[4] and credit ratings of corporate debt.[5] For each of these measures, we form a portfolio that is long high quality and short low quality (or junk), formed in a manner similar to the methodology used by Fama and French (1993), in which stocks are ranked by size and quality measures independently into two size and three quality groups and the intersection of the groups forms six portfolios, with the resulting portfolio equally long the two quality portfolios and short the two junk portfolios. The only exception to these factors being constructed as long the top third and short the bottom third (neutral to a size split) is the credit portfolio, CRED, which goes long the equity of firms with A-rated or better debt and short the equity of firms with C-rated or lower debt. The results are robust across all of these measures.

With international data, we form all of the described portfolios and factors in each of 23 other developed equity markets following a similar methodology. Our international equity data include all available common stocks on the Compustat Global database for 23 developed markets from January 1983 to December 2012.[6] The 23 markets correspond to the union of all countries belonging to the MSCI World Developed Index as of December 2012. We assign individual stocks to the corresponding market based on the location of the primary exchange. For companies traded in multiple markets we use the primary trading vehicle identified by Compustat.

Global portfolio construction closely follows Fama and French (2012), Asness and Frazzini (2012), and Asness et al. (2013). The portfolios are country neutral in the sense that we form long–short portfolios within each country and then compute a global factor by weighting each country's long–short portfolio by the country's total (lagged) market capitalization. The global market factor, RMRF, is the value-weighted return of all available stocks across all markets minus the one-month US Treasury bill rate. The size and value factors are constructed using six value-weighted portfolios formed on size and book-to-market sorts just like in the US. However, while for the US the size break-point is the median NYSE market equity, for the international sample the size breakpoint is the 80th percentile by country to roughly match the US size portfolios. Because some countries have a small cross section of stocks in the early years of our sample, we use conditional sorts that first sort on size and then on book-to-market to ensure we have enough securities in each.

Portfolios are value weighted and rebalanced every calendar month. We require a firm, to be included in any of our tests, to have a non-negative book value and non-missing price at fiscal year-end as well as in June of calendar year *t*.[7] All portfolio returns are in US dollars and excess returns are relative to the one-month US Treasury bill rate.

2.2. Reexamining the Evidence on the Size Effect

Table 1 replicates the evidence on the size effect from the literature, including data outside of the sample periods from the original studies. Columns 1–3 report results for SMB; columns 4–6 report the difference in returns between deciles 1 and 10 (a more extreme difference in size than SMB and also unadjusted through bivariate sorts for book-to-price). Table 1 reports the mean, standard deviation, and *t*-statistic of the size premium over the longest historical sample period from July 1926 to December 2012. SMB yields a premium of 23 basis points (bps) per month that is statistically significant at the 5% level (*t*-statistic = 2.27). The decile spread returns also yield a positive return of 55 bps per month, with a *t*-statistic of 2.32. This result highlights that the size effect is significant but relatively weak compared with other anomalies such as value and momentum, whose *t*-statistics over the same period are 3.7 and 4.6, respectively (using the Fama and French factors HML and UMD), indicating stronger and more reliable return premia.[8] The Sharpe ratio generated for size-based portfolios is roughly 60% of the Sharpe ratio for value portfolios and half the Sharpe ratio that is generated from momentum (all without an attempt to adjust for other risk factors, including market beta).

The table also separates the returns to size into the month of January versus the months of February through December. The returns to SMB are enormous in January at 2.3% for the month and the 1–10 spread in size decile returns is even larger at 6.8% in January. However, from February through December, SMB delivers a negligible 4 bps and the 1–10 portfolio spread is −1 bp, both of which are statistically and economically zero. Hence, what reliable positive premium exists for size appears to solely reside in January and is completely absent the rest of the year. This result illustrates one of the biggest challenges to the interpretation of the size effect, i.e., all of the returns to size come from small stocks in January.

Table 1 reports results over the original sample period studied by Banz (1981) from 1936 to 1975 as well as the out-of-sample period from the original Banz (1981) study: 1926–1935 and 1976–2012. As the table indicates, SMB is insignificant over Banz's original sample period and the 1–10 decile spread is marginally significant (*t*-statistic of 1.82),

though the mean returns are similar to the full period results. The results from Banz (1981) over the same time period for similar decile portfolios are stronger than what we find here, which is likely due to data errors being fixed by CRSP after publication of Banz (1981).[9] The out-of-sample evidence from Banz (1981) is a bit stronger for SMB but weaker for the decile spread returns. Overall, the original size effect is weaker than originally found, consistent with Israel and Moskowitz (2013).

The basic size effect has also experienced significant variation over time, including over relatively long periods (a possible consequence of having a low Sharpe ratio). Table 1 reports summary statistics over the periods for which quality measures are available. The Frazzini and Pedersen (2014) BAB measure is available beginning in January 1931, the Fama and French (2016) profitability and investment factors are available from July 1963, credit from July 1987, and the quality measures used by Asness et al. (2014) from July 1957. We refer to the period from July 1957 to December 2012 as the "quality sample," and we break this period up into three subperiods: (1) from July 1957 to December 1979, shortly before the discovery and publication of the size effect, which we term the "golden age" because the late 1970s was when most researchers were looking at the size effect, which happened to coincide with when its performance was highest,[10] (2) from January 1980 to December 1999, which we call the "embarrassment" because this is when the size effect appears to have vanished, promptly after being discovered and published, and (3) from January 2000 to December 2012, which we term the "resurrection" as the size effect appears to be revitalized during this period. Table 1 highlights these patterns as the size effect seems to have disappeared in the 1980s and 1990s following its discovery but also appears to have made a comeback in the last 13 years of the sample.

Finally, Table 1 reports results for the size premium outside of the US. We report results for an SMB portfolio, constructed in the same manner as for the US, in each of 23 other equity markets, with the data availability ranging from 1986 to 2012. We compute only SMB returns because forming decile portfolios in some of these markets results in too few stocks in some of the deciles. We then average the SMB returns across countries globally excluding the US, and in Europe, North America, and the Pacific, separately, with countries weighted by their lagged total market capitalization. The average returns to SMB outside of the US are weaker, averaging 13 bps with an insignificant *t*-statistic of 0.91. Over the same time period from 1986 to 2012 SMB in the US also averaged only 12 bps with a *t*-statistic of 0.70. Hence, the weak international results could be due to the sample period.

To further show the size premium, the first row of Panel A of Table 2 reports time series regression results of SMB on the market portfolio, RMRF, over the July 1957 to December 2012 time period. The intercept or alpha from the regression is 12 bps per month with a *t*-statistic of 1.12, which is insignificantly different from zero, suggesting that the CAPM explains much of the modest returns to size that existed in Table 1. Next, the lagged return on the market from the previous month is added to capture delayed price responses of stocks, particularly small stocks, to marketwide news [following the results and implications of Lo and MacKinlay (1988) and Hou and Moskowitz (2005) and in the spirit of Asness et al. (2001) to account for non-synchronous price responses due to liquidity differences and lead-lag effects]. SMB has a significantly positive coefficient on the lagged market return, which further pushes its alpha down to an insignificant 7 bps. Next reported are results that add HML and UMD to capture value and momentum exposure. The alpha is 14 bps with a *t*-statistic of 1.23. In the presence of the market and other factors (value and momentum), no reliable size premium is evident.

Table 1 The Size Effect Across Sample Periods and Markets.

The table reports summary statistics on the size premium over time. Two zero-cost portfolios are used to capture the returns to size: the small minus big (SMB) stock factor of Fama and French (1993), obtained from Ken French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data _ library. html), and the return spread between size-sorted value-weighted decile portfolios. The annualized mean and standard deviation of the returns are reported on these spread portfolios, as well as the *t*-statistic of the mean, over the longest historical sample period (from July 1926 to December 2012), for January and February–December separately over the longest sample period, for the same sample period as the Banz (1981) study (January 1936 to December 1975), over the period before and after the Banz (1981) study treated as one period ignoring the discontinuity, over the period when the betting-against-beta (BAB) strategy of Frazzini and Pedersen (2014) is available (January 1931 to December 2012), when the Fama and French (2016) new five-factor model is available (July 1963 to December 2012), when returns to credit portfolios are available (July 1987 to December 2012), when the quality variables of Asness et al. (2014) are available (July 1957 to December 2012), and over three ex post selected subperiods when the size effect is strongest (July 1957 to December 1979, golden age), weakest (January 1980 to December 1999, embarrassment), and recently resurged (January 2000 to December 2012, resurrection). Also reported are SMB returns internationally across 23 other markets, including globally (excluding the US), Europe, North America, and the Pacific. The measure of a stock's size is its market capitalization (share price times shares outstanding) from June of the previous year.

		SMB			1–10 decile spread		
Sample Period	**Years**	**Mean**	**Standard deviation**	***t*-statistic**	**Mean**	**Standard deviation**	***t*-statistic**
Longest sample	1926–2012	0.23%	3.26%	2.27	0.55%	7.69%	2.32
January		2.30%	3.26%	6.50	6.83%	8.41%	7.49
February–December		0.04%	3.19%	0.41	−0.01%	7.37%	−0.06
Banz (1981)	1936–1975	0.16%	2.83%	1.22	0.61%	7.34%	1.82
Pre-and Post-Banz (1981)	1926–1935; 1976–2012	0.29%	3.59%	1.92	0.50%	7.99%	1.49
BAB sample	1931–2012	0.29%	3.28%	2.78	0.67%	7.74%	2.73
Fama and French five-factor sample	1963–2012	0.25%	3.13%	1.95	0.33%	4.89%	1.66
Credit sample	1987–2012	0.14%	3.31%	0.74	0.16%	4.89%	0.56
Quality sample	1957–2012	0.22%	3.01%	1.93	0.33%	4.72%	1.80
Golden age	1957–1979	0.35%	2.87%	2.00	0.68%	4.80%	2.35
Embarrassment	1980–1999	−0.04%	2.66%	−0.23	−0.40%	4.12%	−1.49
Resurrection	2000–2012	0.42%	3.67%	1.41	0.82%	5.31%	1.92
Global ex US	1986–2012	0.13%	2.50%	0.91			
Europe	1991–2012	−0.08%	2.50%	−0.60			
North America	1986–2012	0.09%	2.02%	0.77			
Pacific	1992–2012	−0.31%	3.38%	−1.63			

Table 2 Size Premium Controlling for Junk.

The table reports regression results for the size premium (small minus big, SMB) on the Fama and French factors that include the market (RMRF), its lagged return, high minus low (HML), and up minus down (UMD) and controlling for various measures of quality or junk. Panel A reports results from regressions that add various measures of quality or junk: the quality composite factor (QMJ) from Asness et al. (2014, 2017) and their four dimensions of quality related to profitability, growth, safety, and payout; the Fama and French (2016) five-factor model that includes the factors RMW and CMA, representing profitability and investment, respectively, the Frazzini and Pedersen (2014) betting-against-beta (BAB) factor, which is long low beta stocks and short high beta stocks, and the equity return difference between firms with A-rated debt and higher and firms with C-rated debt and lower (CRED). The sample period for QMJ and its components is July 1957 to December 2012; for the Fama and French (2016) factors July 1963 to December 2012; for the Frazzini and Pedersen (2014) BAB factors January 1931 to December 2012; and for the credit sample July 1987 to December 2012. Panel B reports multivariate regressions of quality or junk, and Panel C reports results using QMJ over the subperiods of the golden age (July 1957 to December 1979), embarrassment (January 1980 to December 1999), and resurrection (January 2000 to December 2012).

Panel A: Controlling for quality/junk

$$SMB_t = \alpha + \beta RM_t + \beta_{-1} RMRF_{t-1} + \mathrm{h}HML_t + \mathrm{m}UMD_t + \mathrm{q}Q_t^* + \varepsilon_t$$

Sample	α	$t(\alpha)$	β	$t(\beta)$	β_{-1}	$t(\beta_{-1})$	h	$t(h)$	m	$t(m)$	q	$t(q)$	RMW	$t(r)$	CMA	$t(c)$	R^2
Quality sample	0.0012	1.12	0.21	8.30													0.09
Quality sample	0.0007	0.63	0.20	7.96	0.13	5.09											0.13
Quality sample	0.0014	1.23	0.17	6.36	0.13	5.42	−0.16	−3.96	0.00	0.13							0.15
Q^* = QMJ (2014)	0.0049	4.89	−0.04	−1.42	0.10	4.82	−0.24	−6.75	0.06	2.70	−0.74	−15.09					0.37
Q^* = Profit	0.0042	3.95	0.06	2.36	0.11	5.07	−0.33	−8.04	0.03	1.24	−0.67	−10.98					0.28
Q^* = Growth	0.0020	1.80	0.17	6.57	0.13	5.50	−0.27	−5.39	0.01	0.27	−0.26	−3.68					0.17
Q^* = Safety	0.0035	3.53	−0.03	−1.12	0.10	4.82	0.20	4.61	0.05	1.98	−0.87	−14.94					0.36
Q^* = Payout	0.0044	4.60	−0.12	−4.28	0.09	4.35	−0.28	−7.93	0.08	3.63	−0.70	−16.86					0.40
Q^* = QMJ (2017)	0.0051	4.88	−0.03	−0.97	0.11	5.19	−0.39	−9.82	0.05	2.25	−0.74	−13.76					0.34
Fama and French sample	0.0016	1.31	0.17	6.13	0.14	5.33	−0.17	−3.87	0.01	0.52							0.16
Q* = RMW, CMA	0.0033	2.82	0.11	4.04	0.14	5.63	−0.09	−1.52	0.04	1.57			−0.54	−9.74	−0.15	−1.81	0.28

(Continued)

Table 2 (*Continued*)

	$SMB_t = \alpha + \beta RM_t + \beta_{-1}RMRF_{t-1} + hHML_t + mUMD_t + qQ_t^* + \varepsilon_t$																
Sample	α	$t(\alpha)$	β	$t(\beta)$	β_{-1}	$t(\beta_{-1})$	h	$t(h)$	m	$t(m)$	q	$t(q)$	RMW	$t(r)$	CMA	$t(c)$	R^2
BAB sample	0.0007	0.72	0.19	10.09	0.13	7.54	0.03	1.09	−0.01	−0.28							0.17
Q^* = BAB	0.0023	2.50	−0.13	−4.77	0.14	8.85	0.01	0.24	0.07	3.39	−0.42	−14.85					0.33
Credit sample	0.0005	0.27	0.11	2.77	0.13	3.39	−0.31	−5.23	0.04	1.15							0.17
Q^* = Cred	0.0035	2.12	0.04	1.13	0.08	2.10	−0.28	−5.02	0.07	2.15	−0.12	−7.82					0.31

Panel B: Multiple measures of quality/junk

	$SMB_t = \alpha + \beta RM_t + \beta_{-1}RMRF_{t-1} + hHML_t + mUMD_t + rRMW_t + cCMA_t + q\,QMJ_t + bBAB_t + dCred_t + \varepsilon_t$																				
Sample	α	$t(\alpha)$	B	$t(\beta)$	β_{-1}	$t(\beta_{-1})$	h	$t(h)$	m	$t(m)$	r	$t(r)$	c	$t(c)$	q	$t(q)$	b	$t(b)$	d	$t(d)$	R^2
Fama and French sample	0.0047	4.36	−0.16	−4.69	0.10	4.62	−0.18	−3.06	0.11	4.29	0.08	0.96	0.09	1.15	−0.64	−6.64	−0.24	−5.61			0.41
Credit sample	0.0047	3.12	−0.28	−5.39	0.04	1.25	−0.17	−2.09	0.18	5.46	0.00	0.02	0.12	1.14	−0.43	−3.00	−0.30	−5.36	−0.06	−3.81	0.50

Panel C: Controlling for quality/junk over time

	$SMB_t = \alpha + \beta RM_t + \beta_{-1}RMRF_{t-1} + hHML_t + mUMD_t + qQ_t^* + \varepsilon_t$												
Subperiod	α	$t(\alpha)$	β	$t(\beta)$	β_{-1}	$t(\beta_{-1})$	h	$t(h)$	m	$t(m)$	q	$t(q)$	R^2
Q* = QMJ													
Golden age	0.0025	1.52	0.27	7.19	0.15	4.10	0.07	0.95	−0.09	−1.83			0.24
	0.0057	4.00	0.07	1.96	0.14	4.70	−0.24	−3.73	−0.06	−1.39	−0.97	−10.73	0.48
Embarrassment	−0.0011	−0.64	0.04	0.97	0.18	5.05	−0.24	−3.56	−0.08	−1.63			0.18
	0.0050	3.06	−0.14	−3.43	0.15	4.85	−0.42	−6.84	−0.06	−1.34	−0.83	−9.08	0.40
Resurrection	0.0054	2.06	0.25	4.25	0.10	1.75	−0.34	−4.46	0.14	3.00			0.25
	0.0089	4.04	−0.17	−2.43	−0.03	−0.59	−0.18	−2.68	0.17	4.43	−0.84	−8.40	0.49

Overall, a weak size effect exists, with substantial variation over time and across season, and meager evidence outside of the US.

3. Resurrecting Size by Controlling for Junk

In this section, we show that accounting for the quality or junk of the stock helps to resurrect the size effect. The strong returns to size when controlling for junk present a challenge to the asset pricing theories.

3.1. The Size Effect Controlling for Quality: Regression Analysis

We first consider the magnitude and significance of the size effect in a regression setting. Table 2 shows the results from adding a quality factor to the regression of size on standard factors in the literature. In this case, we add the Asness et al. (2014) quality minus junk (QMJ) factor, which is a long-short portfolio created from a composite measure of quality, that is long quality stocks and short junk stocks. SMB loads significantly negatively on quality, driving the SMB alpha from 14 to 49 bps per month, which is almost five standard errors from zero (t-statistic = 4.89). The addition of a quality factor to the regression not only raises significantly the average return to size but also increases the precision of the SMB premium, as quality explains a substantial fraction of the variation in SMB's returns, evidenced by the R-squared rising from 15 to 37%.

We show that various other factors or portfolios formed from other measures of quality or junk give similar results on resurrecting the size effect. The QMJ factor constructed by Asness et al. (2014) combines many measures designed to capture quality or junk by looking at variables that proxy for a variety of attributes. We take each component separately (profitability, growth, safety, and payout) as a quality factor and repeat the regression for SMB using each subcomponent. Despite the different measures, in each case the loading on quality, no matter how defined, is significantly negative and the SMB alpha is significantly positive (and more reliable). Using profitability to define quality, the SMB alpha jumps to 42 bps, which is almost four standard errors from zero. Controlling for safety or payout as measures of quality yields very similar numbers (35 and 44 bps alphas). The weakest, based on realized average return, quality measure is growth, yet even here there is a marginally significant 20 bps size premium after adjusting for growth, and SMB loads significantly negatively on this measure of quality, too. Table 2 then uses the latest version of the Asness et al. (2017) quality composite, which excludes payout and drops accruals from the growth composite. As the table shows, the results are unchanged to various perturbations of the quality factor.

Panel A of Table 2 switches to using the two additional Fama and French (2016) factors from their five-factor model as quality proxies: RMW (robust minus weak) profitability factor and CMA (conservative minus aggressive) investment factor, which have been suggested as measures of firm quality (Novy-Marx, 2013; Fama and French, 2016). These two factors are similar to subcomponents of QMJ, but with different specific formations and different creators. Thus, they are not an independent test but a robustness check. Intuitively, both profitability and investment are characteristics that should differ among high versus low quality firms. Fama and French (2016) offer three separate

versions of their factors from sorting on combinations of size and profitability and investment. We show the results for the 2 × 3 versions of their factors from Kenneth French's website, which are nearly identical to using their 2 × 2 and 2 × 2 × 2 × 2 factor specifications. The factor returns are available from July 1963, so we first report the regression of SMB on the market, its lag, HML, and UMD over this period for reference. The SMB alpha is an insignificant 16 bps (*t*-statistic = 1.31). Adding the Fama and French (2016) profitability and investment factors, SMB loads significantly negatively on both, which doubles the SMB alpha to a significant 33 bps per month (*t*-statistic of 2.82). Hence, using Fama and French's (2016) two new factors as measures of quality also resurrects the size effect.

We use, as another measure of quality, a further robustness test, and an additional out-of-sample test, the stand-alone BAB factor from Frazzini and Pedersen (2014), which is available from January 1931. While BAB is part of the QMJ composite, it is the only variable available earlier than our quality period start in 1957. That, and the motivation for BAB, coming from Black (1972, 1992), makes it an interesting factor to highlight alone. The BAB factor is long low beta or safe stocks and short high beta risky stocks and hence can be viewed as a quality measure. Panel A of Table 2 further reports regressions of SMB on the market, its lag, HML, and UMD with and without the BAB factor over the 1931–2012 sample period. Absent the BAB factor, the SMB alpha is only 7 bps with a *t*-statistic of 0.72 from 1931 to 2012, indicating no size premium. However, adding just BAB as a single quality factor to the regression bumps up the SMB alpha to a significant 23 bps (*t*-statistic = 2.50). SMB loads significantly negatively on BAB (coefficient of −0.42 with a *t*-statistic of −14.85), indicating that even this very simple measure of quality is strongly and reliably negatively related to size and resurrects the otherwise absent size premium.[11]

The last two rows of Panel A use another measure of quality that is novel to the literature: the equity returns between firms with A-rated or higher debt minus the equity returns of firms with C-rated or below debt, in which the market capitalization-weighted average of returns is computed for each group. This factor, CRED, captures the equity return difference between firms with high credit-worthy debt minus low-rated debt. Because credit ratings are available for enough firms only beginning in July 1987, the sample period is limited. Consistent with other measures of quality, CRED has positive average returns over the sample period, exhibiting a 0.63 annual Sharpe ratio, which is consistent with the performance of other quality portfolios over this period. The correlation between CRED and QMJ is 0.53. As the table shows, even over this very short time period, this novel measure of quality resurrects the size effect, too. The SMB alpha over this period without controlling for quality is an insignificant 5 bps (*t*-statistic of 0.27). But, controlling for quality using CRED raises the SMB alpha to a significant 35 bps (*t*-statistic of 2.12), and there is a strong negative loading of SMB on this unique quality factor (−0.12 coefficient with a *t*-statistic of −7.82).

Panel B of Table 2 examines multiple measures of quality simultaneously by running a regression of SMB on all of the quality factors. It reports results for the Fama and French (2016) RMW and CMA, Asness et al. (2014) QMJ, and Frazzini and Pedersen (2014) BAB factors simultaneously over the common period July 1963 to December 2012. The negative coefficients on RMW and CMA are soaked up by the very strong negative loadings on the QMJ composite measure and the BAB factor, indicating that they pick up the information

in the Fama and French (2016) profitability and investment factors. This regression is then repeated, adding the credit factor, CRED, to the list of quality factors over the common, shorter sample period July 1987 to December 2012. Even over this shorter time, the Fama and French (2016) profitability and investment factors are still subsumed by the other factors, but QMJ, BAB, and CRED all have significant negative loadings with respect to SMB, suggesting that each captures different aspects of quality that are consistently inversely related to size.

Table OA1 in the Online Appendix reports the correlation matrix of the Fama and French factors, UMD, and the various quality portfolios. SMB is consistently negatively correlated with every quality factor, ranging from –0.18 (CMA) to –0.54 (QMJ). In addition, the quality factors themselves are generally positively correlated with each other, though they seem to capture distinct aspects of quality. Given that quality and size are strongly negatively correlated, we examine the correlation among the different quality factors after hedging out size, reported at the bottom of Table OA1. The correlations are still largely positive, albeit smaller after removing the common size component from returns.

Overall, the results indicate that all of the measures of quality are negatively related to size and are helpful in resurrecting the size premium, even over different sample periods, with the results not particularly sensitive to any particular measure of quality.

Another way to see the size and quality interaction is to look at optimal portfolio weights from a Sharpe ratio maximizing portfolio. Table OA2 in the Online Appendix reports in-sample optimal portfolio weights of SMB with and without a quality factor (using QMJ) in the investment opportunity set and does the same for quality (using QMJ) with and without SMB in the opportunity set. The ex post optimal portfolio weights are a reflection of the regression alphas from before. The optimal portfolios' in-sample Sharpe ratio and correlation to SMB are also reported. Both SMB and QMJ by themselves improve the efficient frontier in the presence of the market, but as the table shows, an optimizer wants to put positive weight on both to maximize Sharpe ratio. The analysis is then repeated, adding the value and momentum factors HML and UMD. Here, both SMB and QMJ again receive positive weight, but are given even larger weight each in the presence of the other. An optimal portfolio wants more SMB when QMJ is also present and vice versa. With all factors present, simply adding SMB to the Fama and French factors only slightly increases the Sharpe ratio from an ex post 1.08 to 1.11. But, adding SMB to the Fama and French factors plus QMJ increases the Sharpe ratio from 1.36 to 1.57. Thus, the marginal impact on the optimal portfolio from adding SMB is small without QMJ present, but it is substantially larger in the presence of QMJ. Essentially, choosing the weight on SMB is not very helpful unless also being exposed to quality. Absent quality, tilting toward SMB means also tilting strongly against quality, with the two effects almost canceling each other out. In all, this evidence suggests that both size and quality are valuable factors not spanned by each other or a rotation of the other factors.

Fig. 1 shows the impact of controlling for quality or junk on the size effect by examining SMB hedged with respect to the market, its lagged value, and HML and UMD factors and using the composite QMJ factor. The figure plots the cumulative sum of returns over time of SMB un-hedged, SMB hedged with the market, its lagged value, HML, and UMD, and SMB hedged with all of those factors plus QMJ. The plot uses the full sample estimates of the betas from July 1957 to December 2012 to estimate the hedged returns to SMB. As Fig. 1 shows, hedging SMB for exposure to the market, value, and momentum

Fig. 1. Cumulative abnormal returns of small minus big (SMB) with and without controlling for junk. The figure plots the cumulative sum of returns over time of SMB; SMB hedged with the market, its lagged value, high minus low (HML), and up minus down (UMD); and SMB hedged with the market, its lagged value, HML, UMD, and the quality factor QMJ. The cumulative returns use the sample estimates of the betas on all factors from the entire period.

factors reduces its returns, but hedging SMB with respect to quality or junk significantly improves returns.

Fig. 2 reports results across 30 different industries. We form SMB portfolios (long the smallest half of firms and short the largest half of firms) within each of 30 industries available from Ken French's data library. We then examine whether the improvement in SMB after controlling for quality or junk is similar within each industry. Though not 30 completely independent tests, this provides 30 different samples of firms from which we can test the robustness of the results.

We compute the alpha of SMB within each industry relative to the market, its lagged value, HML and UMD. We then repeat this computation using the same factors plus QMJ and compare the difference in alphas within each industry. Panel A of Fig. 2 shows the improvement in SMB alpha after controlling for QMJ for each of the 30 industries. The results are remarkably consistent. Every single industry shows positive improvement in SMB's returns after controlling for quality or junk, and for most industries the improvement is significant (with significance harder to achieve in a much smaller sample of firms within a single industry).

Panel B plots the betas of each SMB portfolio on QMJ, which are all negative and are the mirror image of the improvement in alphas in Panel A. These results indicate that the relation between size and quality or junk is robust. Not a single industry fails to find a strong negative relation between size and quality, and as a result, the size premium, controlling for quality, is consistently stronger within every single industry. Small stocks are, on average, junky stocks and this holds for every industry. Controlling for quality improves the size effect within every industry.[12]

Quality, however it is measured and across a variety of specifications, simultaneously resurrects the return premium to size as well as explains much of its variation, transforming it from a small and insignificant effect to an economically and statistically large one and doing so consistently.

Fig. 2. Intra-industry evidence of small minus big (SMB) premia controlling for quality. Panel A plots the improvement in SMB alphas [relative to the Fama and French market return (RMRF), its one-month lag, high minus low (HML), and up minus down (UMD) factors] after controlling for quality using the Asness et al. (2014) quality factor, QMJ, within the 30 industries defined in Ken French's webpage (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ data_library.html). The difference in SMB alphas between the Fama and French factors and the Fama and French factors augmented with the quality factor is plotted by industry. Panel B plots the betas of each SMB portfolio on quality.

3.2. Variation in the Size Premium Over Time

Fig. 1 anticipates the main result in this subsection as casual perusal of the plots in the figure shows a far more consistent size premium when hedged for quality exposure. More formally, we test the stability of the size premium with and without controlling for quality in Panel C of Table 2 by rerunning the regressions of SMB on the market, its lag, HML, and UMD with and without a quality factor over the three subsample periods (golden age, embarrassment, and resurrection) corresponding to the periods over which the unconditional basic size premium varies substantially (and chosen ex post precisely for this purpose). The golden age from July 1957 to December 1979 has a more positive size premium of about 25 bps when adjusting for the market, its lag, HML and UMD, though the *t*-statistic is only 1.52. Adding a quality factor, however, makes the age more golden, as it more than doubles the alpha to 57 bps with a *t*-statistic of 4.00.

During the embarrassment period, from 1980 to 1999, when we know SMB did not do well, the size premium has a negative alpha of −11 bps. However, controlling for quality using QMJ restores SMB's positive alpha over this period to a sizable 50 bps (*t*-statistic of 3.06), which is in-distinguishable from SMB's alpha over the golden age period. Hence, the embarrassing period is no longer embarrassing and the golden age is not relatively golden once we control for quality. Controlling for quality or junk fully explains the seemingly very different performance of the basic size premium over these two different periods. Finally, the resurrection period has positive SMB alphas but the alpha is larger once we control for quality. Like the other two subperiods, the alpha of SMB controlling for quality is of similar magnitude and highly significant. Hence, accounting for quality, the premium for size is positive and more stable through time.

Because quality seems to explain SMB's performance variation through time, determining whether the variation in basic size's performance is driven by variation in the quality premium or variation in size's exposure to quality would be of interest. Panel A of Fig. 3 plots the average quality premium and SMB's beta to quality over the sample and subsample periods (golden age, embarrassment, and resurrection). Panel A shows that SMB's beta to quality is stable over time and consistently negative, but the returns to quality vary, being lower during the golden age (hence, SMB without a quality adjustment looks relatively stronger in this period) and higher in the embarrassment period (hence, why SMB looks worse during this time).

Panel B of Fig. 3 plots the five-year moving averages of SMB's returns, SMB's alpha with respect to the market, its lag, HML, UMD, and quality and the product of SMB's beta on QMJ times the average return on QMJ. As the plot shows, time variation in SMB's alpha is largely determined by its exposure to, and the returns of, quality. Panel C of Fig. 3 plots SMB's quality beta and the quality premium separately over time and shows that most of the variation comes from the quality premium. Although some variation exists in SMB's quality exposure, it is consistently negative. Hence, the realized quality premium drives most of the variation in the size premium through time.[13] Size has a very stable negative exposure to quality, but the returns to quality, and not size per se, vary over time in a manner that has confounded previous interpretations of time variation in the basic size premium.

Fig. 3. Time variation in the quality premium and quality beta on small minus big (SMB). Panel A plots the average realized return to long–short portfolios based on quality as well as the quality portfolio's SMB beta over the sample period July 1957 to December 2012 (the quality sample), as well as the golden age, embarrassment, and resurrection subperiods. Panel B plots the five-year moving averages of SMB returns, alpha with respect to the Fama and French factors augmented with up minus down (UMD) and the quality factor QMJ from Asness et al. (2014), and the product of SMB's beta on quality times the moving average return on quality over time. Panel C plots the five-year moving average of SMB's quality beta and the quality premium separately and their product.

Fig. 4. Distribution of size among junk and quality stocks. Panel A plots the fraction of the number of stocks over time across five size categories based on size quintiles [S1 (small), S2, S3, S4, and S5 (big)] that make up the 20% of stocks with the lowest quality and highest junk ranking (junk). Panel B plots the fraction of the number of stocks over time across the five size groups that make up the 20% of stocks with the highest quality and lowest junk ranking (quality).

3.3. The Size Effect Controlling for Quality: 25 Size-Quality Portfolios

We form size and quality or junk portfolios to look more closely at the size-quality interaction. We form 25 portfolios based on five independent sorts on size (market cap) and five sorts on quality or junk (using QMJ) and group stocks into five quintiles using size and independently five quintiles using quality or junk. The intersection of each of the five categories for each variable determines the composition of the 25 portfolios. Because these are independent sorts with strongly (negatively) correlated sorting variables, the number of firms in each of the 25 portfolios is quite different. The value-weighted average monthly returns in excess of the monthly T-bill rate and their *t*-statistics are then computed over the sample period from July 1957 to December 2012 for each portfolio.

To get a sense of the intersection between size and quality or junk, Fig. 4 examines the size distribution of stocks within the lowest and highest 20% of quality or junk stocks. Panel A plots the fraction of the number of stocks over time within each of the five independent size quintiles, for the 20% of stocks with the lowest quality ranking (junk). Panel B does the same from the universe of the 20% highest quality stocks. As Panel A shows, junk stocks are mostly small stocks. As Panel B shows, among the highest quality stocks the average size is larger, but plenty of small stocks are represented among the quality group (and the distribution among the various sizes is considerably more even among high-quality stocks than among junk stocks). While junk is more correlated with small stocks and quality is more associated with big stocks, there are plenty of large, junky stocks and plenty of small, quality stocks that enable us to examine the interactions between size and quality, controlling for the other.[14]

Fig. 5 plots the reverse exercise of looking at the distribution of quality or junk among the 20% smallest and largest stocks, separately. Panel A shows the distribution of quality among the smallest quintile of stocks, with quality or junk characteristics fairly evenly distributed among the smallest stocks, though a slight tilt toward more junk and less quality. Panel B reports the quality or junk distribution among the largest quintile of stocks and shows that there is a stronger tilt toward high quality and away from junk stocks.

Table 3 reports summary statistics of the 25 size-junk portfolios. The average monthly returns in excess of the Treasury bill rate are reported for each portfolio (along with their *t*-statistics).[15] Moving across the columns of Table 3, a significant size effect is revealed, as the smallest stocks outperform the largest stocks consistently across the size quintiles. The 25 size-quality sorts represent another way to control for quality or junk when considering the size effect, although the relation between size and quality is so strong that these double-sorts are only a partial control. Each row represents the relation between size and average returns within (e.g., controlling for) a quality quintile. The only exception to the significant negative relation between size and average returns is among the junkiest, lowest-quality stocks, although the difference between the smallest and largest quintiles is still a healthy 23 bps per month (the *t*-statistic is insignificant). The rest of the quality or junk quintiles exhibit a very strong size effect and a clear relation between size and average returns.

The reverse is also true, that is, controlling for size, there is a clear quality premium. In every size quintile, quality outperforms junk and the relation is fairly stable. Hence, quality or junk and the size effect are not the same thing, though they are (negatively) related.

Fig. 5. Distribution of quality or junk among large and small stocks. Panel A plots the fraction of the number of stocks over time across five quality groups based on quality quintiles [Q1 (junk), Q2, Q3, Q4, and Q5 (quality)] that make up the 20% of smallest stocks. Panel B plots the fraction of the number of stocks over time across the five quality groups that make up the 20% of largest stocks.

Table 3 Size and Junk Double Sorts.

Panel A reports results from time series regression tests of 25 portfolios sorted on size (market cap) and quality or junk as defined by Asness et al. (2014). The 25 portfolios are formed from independent sorts of stocks into five quintiles using size and quality or junk. The average returns in excess of the monthly T-bill rate and their *t*-statistics are reported over the sample period from July 1957 to December 2012. Panel B reports summary statistics on Fama and French's SMB factor as well as a small minus big (SMB) factor adjusted for quality (SMBQ), which is an average of the "small" minus "big" returns within each quality or junk quintile, averaged equally across the five quality or junk groups. Reported are the annualized means and Sharpe ratios of SMB, SMBQ, market return (RMRF), high minus low (HML), up minus down (UMD), and quality composite portfolio from Asness et al. (2014) (QMJ), along with their correlations with all of the other factors.

Panel A: Time series regression tests

Portfolio	Small	2	3	4	Big	Small–Big
Excess returns (%)						
Junk	0.35	0.42	0.44	0.40	0.12	0.23
2	0.84	0.73	0.68	0.56	0.37	0.46
3	0.87	0.77	0.74	0.59	0.34	0.54
4	0.89	0.86	0.76	0.77	0.47	0.42
Quality	0.97	0.89	0.83	0.78	0.53	0.44
Quality – Junk	0.62	0.47	0.39	0.38	0.42	
t-statistics						
Junk	1.18	1.46	1.65	1.59	0.50	1.21
2	3.40	3.12	3.18	2.80	1.99	3.02
3	3.80	3.56	3.74	3.19	1.93	3.44
4	4.15	4.09	3.95	4.10	2.72	2.82
Quality	4.55	4.23	4.11	4.08	3.23	2.87
Quality – Junk	4.78	3.82	3.37	3.19	2.78	

Panel B: Summary statistics

			Correlation with:				
	Annual Mean (%)	Sharpe	SMB	RMRF	HML	UMD	QMJ
SMBQ	5.0	0.39	0.85	0.15	0.02	−0.18	−0.31
SMB	2.7	0.26	1.00	0.31	−0.22	−0.02	−0.48
RMRF	5.8	0.38		1.00	−0.29	−0.14	−0.54
HML	4.5	0.47			1.00	−0.17	0.06
UMD	8.7	0.60				1.00	0.22
QMJ	5.5	0.59					1.00

The results in Table 3 provide further insight into our earlier findings. Controlling for quality resurrects size in many places where it was previously and seemingly absent. As Table 3 highlights, the junk stocks could be the most interesting set of firms, where among them the relation between returns and size breaks down. These junk stocks have on average poor and very volatile returns and, hence, help explain many of the empirical challenges to the size effect.

Another effective way to look at the size premium independent of quality is to create a quality-hedged size portfolio. Taking an equally-weighted average of the five small minus big portfolios within each quality quintile, we find a return spread of 42 bps per month with a *t*-statistic of 3.18. This average size portfolio, which we refer to as small minus big adjusted for quality or SMBQ, represents the average size premium controlling for quality, although the control is imperfect, as SMBQ still has a negative correlation with quality but a far less strong one than SMB without the quality control. In Panel B of Table 3, we report the annualized mean and Sharpe ratio of SMBQ as well as the Fama and French factor SMB, which does not control for quality in its construction. SMBQ has nearly twice the mean return as SMB and has a 50% higher Sharpe ratio, 0.39 for SMBQ versus 0.26 for SMB. For comparison, we also report the means and Sharpe ratios of the market (RMRF), HML, UMD, and QMJ. SMB has the weakest performance among all factors, but this is because SMB is short a very significant quality premium. Reducing SMB's exposure to quality as SMBQ does (though not eliminating it) puts its performance on a more equal footing with the other factors.

Looking at the correlations among the factors reveals that SMBQ is highly correlated with SMB (0.85) and less negatively correlated with QMJ (−0.31) than SMB is (−0.48). The −0.31 correlation shows, however, that even SMBQ is still negatively correlated to quality, indicating that making the size portfolio completely quality-neutral is difficult, unless specifically hedged from a regression. The other correlations are interesting, too, as SMB has a strong negative value tilt (−0.22) and SMBQ is value-neutral (correlation of 0.02). For these reasons, to examine the independent variation of size, considering alphas from multivariate regressions on all these factors, as in Table 2, is more intuitive.

3.4. Is the Size Premium Concentrated in Extreme Stocks?

Fig. 6 examines the returns to size more finely by looking across size-sorted decile portfolios. From this analysis, we can address whether the size premium is concentrated in the extremes and if the relation between size and average returns is monotonic.

Fig. 6 plots the alphas of each size decile with respect to the Fama and French factors RMRF, RMRF lagged a month, HML, and UMD and these same factors augmented with a quality factor. All regressions are run over the sample period from July 1957 to December 2012. As the figure shows, the alphas adjusted for the Fama and French factors are barely higher for the smallest decile of stocks (Decile 1) compared with the largest (Decile 10) and are essentially flat across Deciles 2 through 9 and exhibit no reliable pattern. In short, no consistent relation exists between size and average returns across the deciles in terms of market or Fama and French–adjusted alphas. This finding is consistent with claims in the literature that the size-return relation is not linear. However, when adding quality as a factor, not only is the return difference between the smallest and largest size deciles magnified,

Fig. 6. Size decile alphas. The alphas of each size decile are plotted with respect to the Fama and French market return (RMRF), RMRF lagged a month, high minus low (HML), and up minus down (UMD) factors and those same factors augmented with the quality factor, QMJ, from Asness et al. (2014).

but perhaps more interesting, an almost linear monotonic relation between the size deciles and their alphas also emerges. Moving from small to big stocks, the alphas steadily decline and become negative for the largest stocks. Hence, controlling for quality or junk restores a more consistent relation between size and average returns.

The fact that quality resurrects the size premium can in part be related to it restoring more decile-monotonicity as well, if a larger absolute premium can reduce the influence of noise on each portfolio. However, it did not have to work out this way. A quality factor could have just as easily raised the returns on all size deciles equally without improving an increasing relation, or it could have added more to the larger deciles or to random deciles and reduced monotonicity. The fact that when size increases proportionately less alpha is evident when controlling for quality suggests that quality or junk exposure is related to size in a linear way and that controlling for quality or junk therefore restores a tighter linear relation between size and average returns, which current asset pricing models (e.g., Fama and French, 2015, 2016) assume.

The relation between size and quality is also stable through time. Panel A of Fig. 7 plots ten-year rolling beta estimates of SMB on QMJ over the sample period (July 1967 to December 2012) and shows that the betas are always negative and range from −0.40 to −1.25. Panel B of Fig. 7 plots the rolling ten-year betas of each size decile. Again, the time series variation in the betas is relatively small, but more interesting, the ordered relation between size and quality or junk is extremely stable though time, as smaller size deciles consistently have more negative quality betas and the effect is very stable throughout the sample. Few periods exist in which betas with respect to quality are not ordered almost perfectly by size, a remarkable feat considering the estimation error inherent in beta estimates. Repeating the same exercise for other measures of quality or junk using the various measures of Asness et al. (2014), Fama and French (2016), or Frazzini and Pedersen (2014) yields similar results.

Fig. 7. Rolling beta estimates of size portfolios on quality. Panel A plots the ten-year rolling beta estimates of small minus big (SMB) on quality, using the Asness et al. (2014) composite quality factor QMJ, estimated from a model that also includes RMRF, its lag, high minus low (HML), and up minus down (UMD), using the monthly returns over the preceding 120 months. Panel B plots the rolling quality betas of each size decile portfolio.

3.5. Seasonality in Size: The January Effect

Prior research has shown that the size effect mostly resides in January (Keim, 1983; Roll, 1983; Reinganum, 1983b). Table 1 showed that all of the returns to size are concentrated in January, with no evidence of any size effect (economically or statistically) outside of January.

Table 4 reexamines the seasonality in the size premium after controlling for quality. Reported are results from regressions of the returns to SMB on a January dummy, a non-January dummy (February through December), and the Fama and French factors RMRF, RMRF lagged, HML, and UMD over the full sample period. Confirming earlier results and those in the literature, there is a large January size premium (2.09% with a *t*-statistic of 5.59), no evidence of any size effect outside of January (−0.04% with a *t*-statistic of −0.32), and the Fama and French factors do not capture much of this seasonality. The last column of the table reports a statistical test for the difference in January versus non-January alphas, which easily rejects the null that January has the same returns as the rest of the year.

Table 4 also reports results when adding a quality factor (QMJ) to the regression. Controlling for quality delivers a positive and significant size premium outside of January of 38 bps (*t*-statistic = 3.62). Controlling for quality mitigates the very large size premium in January, dropping it from 2.09 to 1.57%. While a large January premium still remains, the premium for size is now present throughout the year and the difference between the January and non-January alpha, which is still significant, is now halved.

Table 4 then repeats this exercise over the subsample periods golden age, embarrassment, and resurrection. In every subperiod, quality resurrects the size effect outside of January, delivering a consistent premium of at least 33 bps (golden age) and as much as 91 bps (resurrection). Outside of January, the returns to size controlling for quality are larger (almost twice as large) during the embarrassment period than they are during the supposed golden age for size. Hence, regarding February to December, the notion of a golden age period for size and an embarrassment period for size is backward. As Table 4 shows, this is due to the quality premium confounding the performance of size over these periods. A major reason the golden age for size exists is because a size portfolio is long junk and short quality, with junk greatly outperforming quality in January months over this period. Hence, failure to control for quality or junk obscures the size effect. Consistent with this notion, quality also diminishes the size premium in January, when it is insignificant in the last two subperiods and insignificantly different from the returns in February to December over these subperiods.

These results suggest that quality or junk also helps explain the strong seasonality associated with size-based strategies. The strong performance of junk stocks in January drives a significant fraction of the apparently high returns to size in January, while depressing the returns to size outside of January. Controlling for quality or junk reduces this seasonal component substantially and shows a strong size premium throughout the year.[16]

3.6. International Evidence

Finally, we examine the size effect in 23 other countries to perform out of sample tests on the role of quality in reviving the size effect. We form SMB portfolios within each

Table 4 Seasonal Patterns and the Size Premium.

The table reports regression results for the size premium (SMB) on the factors RMRF, its lagged value, HML, and UMD and the composite quality factor from Asness et al. (2014), where the alphas are estimated for the months of January and non-January separately using dummy variables for those months. Also reported is the difference between January and other months, along with a *t*-statistic on that difference in the last column. Results are reported over four sample periods: the full quality sample period (July 1957 to December 2012), and the golden age (July 1957 to December 1979), embarrassment (January 1980 to December 1999), and resurrection (January 2000 to December 2012) subperiods for the size premium.

Period				$SMB_t = \alpha_{Non\text{-}Jan.} + \alpha_{Jan.} + \beta RMRF_t + \beta_{-1} RMRF_{t-1} + hHML_t + mUMD_t + qQMJ_t + \varepsilon_t$													
	αNon-Jan.	$t(\alpha)$	αJan.	$t(\alpha)$	β	$t(\beta)$	β_{-1}	$t(\beta_{-1})$	h	$t(h)$	m	$t(m)$	q	$t(q)$	R^2	January difference	t(difference)
Quality sample	–0.0004	–0.32	0.0209	5.59	0.16	6.21	0.13	5.29	–0.19	–4.68	0.02	0.90			0.18	0.0213	5.46
	0.0038	3.62	0.0157	4.74	–0.03	–1.28	0.10	4.77	–0.26	–7.10	0.07	3.08	–0.71	–14.37	0.38	0.0119	3.42
Golden age	–0.0001	–0.08	0.0354	6.34	0.25	6.95	0.14	4.02	–0.10	–1.41	–0.03	–0.67			0.34	0.0355	6.13
	0.0033	2.42	0.0359	7.61	0.05	1.55	0.14	4.75	–0.38	–6.02	–0.01	–0.21	–0.94	–11.27	0.55	0.0326	6.67
Embarrassment	–0.0016	–0.89	0.0045	0.79	0.03	0.79	0.18	5.01	–0.25	–3.67	–0.07	–1.46			0.19	0.0061	1.04
	0.0058	3.35	–0.0013	–0.27	–0.14	–3.42	0.15	4.87	–0.42	–6.81	–0.06	–1.51	–0.86	–9.12	0.40	–0.0071	–1.37
Resurrection	0.0041	1.50	0.0180	1.98	0.27	4.44	0.09	1.55	–0.33	–4.22	0.15	3.22			0.26	0.0139	1.45
	0.0091	3.86	0.0069	0.90	–0.18	–2.40	0.03	–0.58	–0.18	–2.68	0.17	4.33	–0.84	–8.19	0.49	–0.0022	–0.27

Fig. 8. International evidence of SMB premia controlling for quality. Panel A plots the improvement in SMB alphas [relative to the Fama and French factors market return (RMRF), RMRF lagged a month, high minus low (HML), and up minus down (UMD)] after controlling for quality (using the Asness et al. (2014) composite quality factor (QMJ)) across 24 countries, as well as five regions: global, global excluding US, Europe, North America, and Pacific. Plotted is the difference in SMB alphas between the Fama and French factors versus the Fama and French factors augmented with the quality factor, by country and region. Panel B plots the betas of each SMB portfolio on quality. The regressions are estimated using rolling five years of data for each country.

international equity market following the same procedure as above. Similarly, we form a quality factor based on the Asness et al. (2014) QMJ factor construction in each of these markets following their procedure. Fig. 8 reports the change in SMB's alpha for each country after controlling for quality by regressing SMB in each country on the local stock market index, the market lagged, HML, and UMD factors constructed within that market. The same regression is repeated including the quality factor in that market as a regressor, and the difference in SMB alphas between the two regressions, with and without controlling for quality, are plotted country by country in Fig. 8. Panel A shows a positive increase in SMB alpha for 23 out of 24 countries once we control for quality (the exception being Ireland, where the point estimate is very close to zero and statistically no different from zero). Panel B of Fig. 8 shows that the betas of SMB on quality are also uniformly negative (excluding Ireland) and where we find that 14 out of 24 are statistically significant. The *t*-statistics range from 0.5 to −17, and the median *t*-statistic is −2.2.

Aggregating across all countries outside of the US, the *t*-statistic of the beta on quality is −3.32. These results are remarkably consistent across countries, providing evidence of a negative relation between size and quality across countries and, therefore, a healthy size premium internationally once we control for quality and, hence, a wealth of out-of-sample evidence for our earlier findings.

As a further test of the international results, we replace the international version of QMJ with the international versions of the Fama and French (2016) RMW and CMA factors (available on Ken French's website) to control for quality. Using SMB formed across countries, excluding the US, and regressing it on the international versions of the market, HML, and UMD, we obtain an insignificant alpha of 11 bps with a *t*-statistic of 0.99. Adding RMW and CMA to the regression to control for quality, SMB's alpha goes up to a significant 29 bps with a *t*-statistic of 2.38. Regressing SMB adjusted for quality on SMB unadjusted for the Fama and French quality factors, the difference in alphas (of 18 bps) has a *t*-statistic of 6.36, more than six standard errors from zero.

These results are further testament to the strength of the size-quality interaction, in which the same patterns emerge in 23 other international markets using a variety of quality metrics, including those created by other researchers, such as Fama and French (2016). The size premium appears to be alive and well across all of these markets once we control for quality in any number of ways.

4. Further Tests of the Theories of the Size Effect

We turn now to further tests of the theories that have been offered for the size effect, and in particular how the interaction between size and quality/junk may help inform us about various theories.

4.1. Risk-based Theories: Characteristics and Asset Pricing Tests

The strong returns to size when controlling for quality present a challenge to standard asset pricing theories to the extent that size is not a risk in and of itself. In other words, for size to be priced in a rational model, size would have to proxy for a rational risk factor

such as the risk associated with growth options. Hence, the more we control for actual risk factors, the smaller should be the residual return of size. Our results presented in Section 3 test this hypothesis using a variety of risk factors proposed in the literature and what we find is the exact opposite: Controlling for growth and other factors makes the size effect stronger, not weaker.

Hence, our results in Section 3 present a rejection of this interpretation of standard asset pricing models, although we acknowledge that size could still proxy for other risk factors (unknown to us, but presumably known to the investors who are setting prices) or measurement errors, e.g., in the measurement of growth options (for which the theoretical literature does not offer clear guidance on the best empirical measures). In this section, we conduct further standard asset pricing tests of size-quality or junk portfolios.

We conduct asset pricing tests using the 25 size-quality/junk portfolios described in Section 3.3 as well as the 25 size-BE/ME portfolios of Fama and French (1996). We examine two different factor models: (1) the Fama and French (1993) factors RMRF, SMB, and HML, augmented with the UMD momentum factor and (2) a factor model that includes those same factors plus the composite quality factor, QMJ. The goal is to see how these models explain the portfolio returns. In other words, here we follow Fama and French (1996) and include size in the risk model even though standard theory does not predict that size is a priced risk. We show that, even when size is included as a risk factor, we can reject the Fama and French model, but including quality significantly improves the fit. The success of the model that includes both size and quality (and the other standard factors) can either be viewed as supportive of a rational model where we are yet to understand why size is a rational risk factor (as well as why some of the other factors are risk-based) or an expression of the fact that we need to include size and quality to explain the cross section of returns, rejecting the standard theories that do not predict such risk factors.

Fig. 9 plots the actual average returns of portfolios versus the predicted expected returns of those portfolios from each asset pricing model. A 45-degree line forced through the origin is plotted to highlight the pricing errors from each asset pricing model for each portfolio, with the distance between the points and the 45-degree line representing the pricing error under each model, assuming that the intercept from the cross-sectional regression between average and model-predicted returns is zero, which is the same as forcing each model to price the equity risk premium in addition to the cross section of returns. Also reported are the average absolute pricing errors or alphas from each model, the Gibbons et al (1989) (GRS) F-statistic on whether the alphas are jointly zero, its p-value, and the cross-sectional R-squared from a regression of average returns on predicted model expected returns. Panel A plots the results for the 25 size-quality portfolios.

As Fig. 9 shows, the Fama and French plus momentum factors do not explain the returns to size-junk portfolios, leaving an average absolute alpha of almost 20 bps per month with a Gibbons et al (1989) (GRS) F-statistic of 4.37 that easily rejects the null that the alphas are zero. The cross-sectional R-squared is 0.006, indicating no reliable relation between the predicted returns from the Fama and French factors and the average returns of the size-quality portfolios. The Fama and French size, value, and momentum factors fail to capture the dimension of quality in returns and the size-quality interaction. Adding QMJ to the Fama and French factors, however, explains the portfolio returns nicely, as the average absolute alpha drops to 7.8 bps with an F-statistic of only 1.54 that fails to reject the null. The cross-sectional R-squared here is 0.914, suggesting a tight relation between the model's predicted returns and actual average returns. It is also worth noting that both

Fig. 9. Asset pricing tests and pricing errors on size, book-to-market equity ratio (BE/ME), and quality portfolios. Plotted are the actual average returns of portfolios versus the predicted expected returns of those portfolios from an asset pricing model. The data points plotted in normal font correspond to the predicted expected returns from the Fama and French (1993) model augmented with a momentum factor (FF4) and the data points plotted in bold correspond to the predicted expected returns from the FF4 model plus the quality factor, QMJ. A 45-degree line forced through the origin is plotted to highlight the pricing errors from each asset pricing model for each portfolio, with the distance between the points and the 45-degree line representing the pricing error under each model assuming the intercept from the cross-sectional regression between average and model-predicted returns is zero, which is the same as forcing each model to price the equity risk premium in addition to the cross section of returns. Also reported are the average absolute pricing errors or alphas from each model, the Gibbons et al (1989) F-statistic on whether the alphas are jointly zero, its p-value, and the cross-sectional R-squared from a regression of average returns on predicted model expected returns. Panel A plots the results for the 25 size-quality portfolios; Panel B, for the 25 size and BE/ME portfolios from Fama and French (1996).

models contain SMB, yet the first model could not explain the variation in average returns across the size-quality spectrum without the quality factor. This indicates that quality is not subsumed by size and is necessary in explaining these portfolio returns.

Panel B of Fig. 9 shows the results for the 25 size and BE/ME portfolios of Fama and French (1996). This exercise is interesting because the test portfolios are not sorted by quality, hence we can see how well the quality factor prices the classic Fama and French 25 size and BE/ME portfolios. Here, the Fama and French factors do better, but not as well as adding quality. Under the Fama and French factors, an average absolute alpha of 8.6 bps with a GRS *F*-statistic of 2.63 easily rejects the null. The cross-sectional *R*-squared is 0.682. As is well known, Fama and French (1993, 1996) statistically reject their own model due largely to the poor performance of extremely small growth stocks (and to a lesser extent, large growth stocks). These are highlighted on the graph as S1V1 (smallest size, lowest BE/ME) and S5V1 (largest size, lowest BE/ME). Adding quality to the regression reduces the pricing errors to an absolute average 7.9 bps with a GRS *F*-statistic of only 1.95 and a cross-sectional *R*-squared of 0.811. The quality factor helps improve the fit of the 25 size and BE/ME portfolios. The plot shows precisely how the quality factor is helping, as the two most troublesome portfolios, smallest and largest growth, are captured by their covariance with the quality factor that Fama and French's SMB and HML portfolios fail to capture. The arrows on the graph highlight how these two portfolios are priced under the two different models (one excluding and one including the quality factor). Hence, quality also helps capture the extreme low returns to small growth stocks which lead to rejection of the Fama and French model.

4.2. Risk-based Theories: Non-price-based Size Measures

Berk (1995a), using an argument from Ball (1978), shows that because size is typically measured by market capitalization, which contains market prices, any misspecification of the asset pricing model leads to a negative relation between size and average returns. In other words, according to this theory, size is a measure of time-varying risk premiums because size is measured by market value, which is influenced by risk premia. Berk (1995b, 1997) suggests that using non-price-based measures of size is therefore a better way to test the true relation between size and average returns and finds that using non-price-based size measures results in no reliable size premium.

Table 5 reexamines the relation between non-price-based measures of size and average returns. In Panel A, we rank stocks based on the non-price size measures suggested by Berk (1997) plus two others, i.e., book assets, book equity, sales, PP&E, and number of employees. For each non-price size measure, stocks are ranked into deciles every June and the value-weighted returns of each decile are computed over the following year (the exact same procedure we use to form the market cap size deciles). Panel A reports the alphas of the return difference between the smallest and largest decile portfolios using the non-price-based size measures from regressions on the factors RMRF, RMRF lagged, HML, and UMD. No reliable size premium exists for any of the non-price-based size measures over the sample period. The last column of the table reports results for the portfolios sorted on market capitalization for comparison. Here, too, the results are insignificant. Thus, consistent with Berk (1995a, 1995b, and 1997), we find no reliable relation between non-price-based size measures and average returns, but inconsistent with Berk's theoretical argument we also fail to find any relation between price-based size measures and returns.

Panel A then repeats the regressions but adds a quality factor (QMJ) to the regression. Controlling for quality systematically resurrects a size premium among every non-price-based size measure (and the market cap measure, too). The contrast in results across all size measures is striking: Every estimated alpha from the first row is insignificant (ranging from 0 to 17 bps), and every alpha in the second row, which controls for quality, is large, positive, and significant (ranging from 58 to 83 bps with *t*-statistics of 4.5 to almost 6). Comparing the magnitude of these alphas with those based on market capitalization (in the last column), we reject the Berk (1995a) conjecture that the non-price-based size deciles deliver a smaller or insignificant size premium. Book assets, sales, book equity, PP&E, and number of employees produce size decile premia of 83, 67, 66, 58, and 68 bps per month, respectively, and the market cap size decile premium is 64 bps over the same period, after controlling for quality.

Table 5 also regresses the non-price-based size portfolios on the same factors plus the market cap–based size portfolio SMB. These regressions test whether the non-price-based size portfolio returns are any different from the price-based size factor, in the presence of the other factors that include quality. Without quality, no reliable alpha is associated with the non-price-based size portfolios. However, controlling for quality, the non-price-based size portfolios deliver positive alphas even after adjusting for market cap–based size via SMB. All five non-price-based size measures produce positive alphas, ranging from 9 to 29 bps, with respect to SMB after controlling for quality, and three out of five are statistically significant. Hence, non-price-based size measures seem to deliver at least as high return premia as market cap–based measures, which is the opposite of the conjecture in Berk (1995a).[17]

Panel B of Table 5 examines the relation between non-price-based size measures and returns using double sorted portfolios based on book assets and size/junk. The analysis is identical to that used for Table 3, where we independently sort stocks based on size and quality or junk into quintiles and calculate returns to the intersection of those sorts to form 25 size-junk portfolios, except here we sort by book assets instead of market cap for our size measure. Panel B reports the average returns in excess of the one-month T-bill rate for each of these 25 portfolios, along with their *t*-statistics, as well as returns adjusted for the market, HML, and UMD (and their *t*-statistics). The results show a healthy size premium within each quality category. Controlling for quality, small stocks significantly outperform large stocks, even when size is measured without using market prices.

The challenge that the size premium shows up only for market price-based measures of size is met by controlling for quality or junk. Doing so, we find a healthy and at least as large (if not larger) size premium associated with portfolios sorted on non-price-based measures of size that is evident in different sample periods, within 30 different industries, and across 23 other international equity markets. This evidence rejects the notion that the size premium is driven by the effect of time-varying risk premia on market-based measures of size.

4.3. Behavioral Versus Risk-based Theories: Characteristics of Size-quality Portfolios

Our results challenge standard risk-based theories of asset pricing, but what about the behavioral theories? Behavioral finance suggests that stocks can be mispriced and more so for small stocks with greater limits of arbitrage (e.g., greater implementation costs). However,

Table 5 Size Premium for Portfolios Sorted on Non-Price-Based Measures of Size.

Panel A reports regression results for the return difference between the smallest size decile portfolio (P1) and the largest size decile portfolio (P10), P1–P10, where portfolios are value-weighted based on sorts using non-priced-based measures of size, book assets, sales, book equity, property, plant, and equipment (PP&E), and number of employees. We also include portfolios sorted on market cap for comparison in the last column. The P1–P10 spread portfolio for each of the non-price-based size measures is constructed in the same manner used for market capitalization-sorted portfolios. We form decile portfolios by sorting stocks each July, based on their June measure of size using each of the non-price-based size measures, and then compute returns to each decile portfolio, in which securities are weighted by their market values, over the following year. Results are reported for regressions of the non-price-based size premia on the market, the lagged market, high minus low (HML), and up minus down (UMD), with and without controlling for quality using the composite factor, QMJ, from Asness et al. (2014) as a regressor. We also report results from regressions that also include small minus big (SMB) as a regressor (market cap-based size portfolio) to examine if the non-market-based size portfolios exhibit any different returns from the market cap-based size portfolios, with and without controlling for junk. For brevity, we report only the estimated alphas and their *t*-statistics. Panel B reports returns and *t*-statistics of double sorted portfolios based on the non-price-based measure of size, book assets, and quality or junk, with stocks sorted independently on book assets and junk into five quintiles and the intersection of the groups forming the 25 portfolios. Raw returns in excess of the one-month Treasury bill rate, as well as risk-adjusted returns or alphas net of the Fama and French market return (RMRF), HML, and UMD factors, are reported. Results are reported over the full sample period over which quality variables are available (July 1957–December 2012).

Panel A: Regressions of P1-P10 using non-price-based measures of size

	Book assets		Sales		Book equity		PP&E		Employees		Market Cap	
Size measure	α	$t(\alpha)$	α	$t(\alpha)$	α	$t(\alpha)$	α	$t(\alpha)$	α	$t(\alpha)$	α	$t(\alpha)$
$P1 - P10_t = \alpha + \beta RMRF_t + \beta_{-1} RMRF_{t-1} + hHML_t + mUMD_t + qQMJ_t + \varepsilon_t$												
No control for quality ($q = 0$)	0.0017	0.96	0.0002	0.10	0.0004	0.22	0.0008	0.00	0.0000	0.01	0.0004	0.20
Control for Quality ($q \neq 0$)	0.0083	5.98	0.0067	5.52	0.0066	4.98	0.0058	4.57	0.0068	5.78	0.0064	4.21
$P1 - P10_t = \alpha + \beta RMRF_t + \beta_{-1} RMRF_{t-1} + sSMB_t + hHML_t + mUMD_t + qQMJ_t + \varepsilon_t$												
Control for market cap size												
No control for quality ($q = 0$)	−0.0002	−0.31	−0.0015	−1.71	−0.0015	−2.14	−0.0070	−0.88	−0.0015	−1.48		
Control for Quality ($q \neq 0$)	0.0025	3.42	0.0020	2.66	0.0009	1.43	0.0010	1.19	0.0029	3.30		

(*Continued*)

Table 5 (*Continued*)

Panel B: Book size and junk double sorted portfolios						
Portfolio	Small	2	3	4	Big	Small – Big
Excess Returns (%)						
Junk	0.71	0.45	0.30	0.33	0.23	0.49
2	1.11	0.85	0.67	0.65	0.55	0.56
3	1.22	0.93	0.84	0.77	0.60	0.62
4	1.14	0.94	0.82	0.79	0.67	0.48
Quality	1.07	0.95	0.91	0.83	0.73	0.34
Quality–Junk	0.35	0.50	0.60	0.49	0.50	
t-statistics for excess returns						
Junk	1.97	1.31	0.93	1.08	0.74	1.96
2	3.44	2.96	2.59	2.73	2.47	2.31
3	4.05	3.59	3.61	3.70	3.21	2.63
4	4.18	3.83	3.83	3.98	3.69	2.30
Quality	4.15	4.11	4.33	4.34	4.27	1.84
Quality–Junk	2.06	2.75	3.38	2.74	2.44	
Returns adjusted for RMRF, HML, UMD (%)						
Junk	0.37	0.13	−0.04	−0.08	−0.26	0.63
2	0.75	0.50	0.22	0.13	−0.03	0.78
3	0.82	0.49	0.35	0.25	0.04	0.79
4	0.75	0.53	0.34	0.28	0.13	0.62
Quality	0.68	0.55	0.48	0.38	0.30	0.38
Quality–Junk	0.31	0.42	0.52	0.46	0.56	
t-statistics for returns adjusted for RMRF, HML, UMD						
Junk	1.31	0.54	−0.18	−0.45	−1.70	2.65
2	3.05	2.72	1.50	1.12	−0.41	3.24
3	3.67	3.15	2.87	2.69	0.61	3.40
4	3.86	3.56	3.16	3.29	2.37	3.05
Quality	3.95	4.38	4.92	4.94	5.60	2.03
Quality – Junk	1.82	2.39	3.12	2.89	3.26	

controlling for the level of mispricing, the behavioral theories do not predict any size effect. Hence, the fact that we find a strong size effect when controlling for value and quality factors is puzzling in light of this reading of the behavioral theories, although we should acknowledge that future research could come up with a new behavioral theory to match these findings.

The size effect is about a particular kind of mispricing in a particular direction. That is, small stocks deliver higher returns than large stocks, or small stocks are cheap relative to large stocks, ceteris paribus. The parts of behavioral finance that speak most directly to such a directional mispricing is the literature on short-selling frictions and price-optimism models caused by dispersion in opinion. These theories hypothesize that stocks that are more difficult to sell short and have greater dispersion of opinion about their value are more likely to become overvalued (Miller, 1977).

Table 6, Panel A, shows that small stocks have higher short-selling costs than large stocks over the sample period from January 2010 to October 2016. We compare the daily cost of borrow score (DCBS) from Markit Data Explorers, a number from 1 to 10 indicating the relative fee to short the stock, for our 25 size-quality portfolios, and find the average score for small cap stocks to be higher than the average score for large cap stocks.[18] This is especially true for the smallest and junkiest stocks. When controlling for the quality of a stock, the same pattern emerges in which the smallest stocks have the highest shorting costs.

Panel B shows the amount of shorting for small versus large cap stocks, measured by the lendable shares percentage, defined as the supply of lendable shares divided by the market cap of each stock, over the sample period from January 2010 to October 2016. Small stocks on average are shorted less than large stocks, especially the smallest and junkiest

Table 6 Level and Costs of Shorting for Size and Junk.

Reported are pooled averages of shorting activity and shorting costs for the 25 portfolios sorted on size (market cap) and quality or junk from Table 3. We report the daily cost of borrow score (DCBS), a number from 1 to 10 indicating the fee for each stock (Panel A) and the lendable shares defined as the supply of lendable shares divided by the market cap for each stock, expressed as a percentage (Panel B). The data cover the period January 2010 to October 2016.

Portfolio	Small	2	3	4	Big
		Panel A: DCBS score			
Junk	4.4	3.0	2.5	2.3	2.3
2	3.0	1.9	1.7	1.6	1.7
3	2.1	1.5	1.3	1.2	1.4
4	1.6	1.3	1.2	1.2	1.3
Quality	1.3	1.1	1.1	1.1	1.0
		Panel B: Lendable shares (%)			
Junk	3.1	3.7	3.6	3.1	3.3
2	9.5	12.9	12.8	12.0	11.6
3	15.6	18.9	21.5	22.6	21.0
4	18.0	22.0	23.7	25.0	24.9
Quality	17.3	20.9	22.0	22.6	23.6

stocks. Given this evidence, according to the behavioral theories, smaller stocks should be more overvalued, which is counter to the evidence on the size effect. The smallest, junkiest stocks have the worst expected returns, consistent with them being the most overvalued. For both Panels A and B of Table 6, the junkiest stocks are the most expensive to short and the least shorted, possibly consistent with their expected returns being poor according to the behavioral theory. However, the costs of shorting and amount of shorting seem to go the opposite way for explaining why small stocks have higher expected returns than large stocks.

To further test whether larger shorting constraints lead to smaller stocks becoming more overvalued as in Miller (1977), we examine the past five-year returns for our 25 size-quality portfolios. DeBondt and Thaler (1985) argue that markets exhibit an overreaction behavioral bias as evidenced by the return reversals that occur at five-year horizons. They find that past five-year winners are overvalued and tend to subsequently revert to lower valuations, while past five-year losers are undervalued and tend to subsequently revert to higher valuations. Thus, if small stocks are overvalued, we would expect to see high past five-year returns. Table 7, Panel A, presents contradictory evidence showing that the smaller stocks have experienced the weakest past five-year returns on average. This is especially true for the smallest and junkiest stocks. Here, we find little support that small stocks are overvalued. Moreover, we find that the worst past five-year performance is in the junkiest stocks, particularly the small junky stocks. So, despite facing large shorting costs and few lendable shares, these stocks do not appear to be overvalued, as Miller (1977) predicts.

The price-optimism models starting with Miller (1977) also suggest that when greater disagreement exists about the value of a stock, the stock becomes more overvalued, leading to lower subsequent returns. Following Diether et al. (2002), we use the dispersion in analysts' forecasts of earnings as our measure of disagreement. Panels B and C of Table 7 present two different versions of the measure of dispersion. Panel B is normalized by the absolute value of the mean forecast, as in Diether et al. (2002). Panel C normalizes by the price of the stock to avoid issues of having zero in the denominator which can affect the measure in Panel B. Based on the evidence in Panels B and C, smaller and junkier stocks have greater levels of dispersion which should lead to overvaluation of those stocks. While junkier stocks have lower subsequent returns than quality stocks, consistent with them being overvalued according to this theory, small stocks outperform large stocks. Once again, the behavioral theories fail to predict the size effect, in which the overvaluation theories go the wrong way.

4.4. Liquidity-based Theories: Transaction Costs and the Level of Liquidity

We next consider theories for the pricing of liquidity (Amihud and Mendelson, 1986) and liquidity risk (Acharya and Pedersen, 2005). Per these theories, investors prefer stocks that are more liquid and face less liquidity risk, all else equal (i.e., assuming similar risk exposures). Therefore, more illiquid stocks and stocks with more liquidity risk must offer higher expected returns. Because small stocks are more illiquid and face more liquidity risk, these theories can explain the size effect. Investors who buy small stocks must be compensated for their relatively high transaction costs and for the risk that these transaction costs rise

Table 7 Past Five-Year Returns and Analysts' Earnings Forecasts Dispersion for Size and Junk.

Reported are average statistics on past five-year returns and dispersion of earnings for the 25 portfolios sorted on size (market cap) and quality or junk from Table 3. We report the past five-year returns as an indication of market overreaction (Panel A) and two measures of dispersion in analysts' earnings forecasts, one normalized by average forecast (Panel B) and one normalized by stock price (Panel C). The data cover the period January 1980 to December 2012.

Portfolio	Small	2	3	4	Big
Panel A: Past five-year returns (%)					
Junk	−7.2	−2.0	4.0	9.3	15.8
2	2.8	9.8	15.4	21.6	28.7
3	11.1	17.5	22.1	27.9	36.4
4	19.3	23.0	27.4	31.5	37.7
Quality	25.4	24.5	28.1	32.2	38.2
Panel B: Earnings dispersion/mean earnings (%)					
Junk	67	71	61	74	80
2	69	55	42	35	26
3	58	41	29	21	14
4	49	31	18	11	8
Quality	46	25	12	7	5
Panel C: Earnings dispersion/price (%)					
Junk	40	13	8	6	4
2	13	4	3	2	1
3	5	2	1	1	1
4	3	1	1	1	0
Quality	2	1	1	0	0

unexpectedly when liquidity is most desired. This compensation is in fact the size premium, according to these theories.

We first consider the level of liquidity across our size and quality portfolios. Table 8 reports two measures of liquidity across the 25 size-junk portfolios, namely the average percentage half bid-ask spread (Panel A) and the market impact cost per dollar traded (Panel B). The market impact costs are estimated from Frazzini et al. (2013), who use live proprietary trading data from a large institutional trader to calibrate a trading cost model. The model computes price impact for a given fund size (net asset value, NAV). For the present calculations, we assume a constant fund NAV of $1 billion in capital for use in the trading cost model.[19]

As Table 8 shows, small firms face larger percentage bid-ask spreads than large firms consistently across the quality spectrum. However, the bid-ask spreads are similar across the quality groups, controlling for size. Hence, liquidity seems to vary strongly with size, but not with quality. Market impact cost per dollar traded portrays a similar picture.

From this evidence, we draw two conclusions. Liquidity can help explain the size effect and is tightly connected to size, and liquidity appears unrelated to quality and not

Table 8 Liquidity Level: Size, Junk, and Trading Costs.

Reported are pooled averages of liquidity and trading cost measures for the 25 portfolios sorted on size (market cap) and quality or junk from Table 3.We report the average half bid-ask spread as a percentage of share price (Panel A), which is half of the bid minus ask price divided by the mid-price, and the market impact cost per dollar traded estimated from Frazzini et al. (2013) assuming a constant fund net asset value (NAV) of $1 billion plus one half of the effective bid-ask spread, all expressed in basis points (Panel B). The trading cost data cover the period January 2000 to December 2012.

	Small	2	3	4	Big
Panel A: Bid/ask spread (%)					
Junk	3.1	0.9	0.3	0.2	0.1
2	3.4	1.0	0.3	0.2	0.1
3	3.8	1.1	0.3	0.1	0.1
4	3.4	1.2	0.3	0.1	0.1
Quality	2.5	1.2	0.3	0.1	0.1
Panel B: Market impact cost per dollar traded (bps)					
Junk	33.98	20.46	15.50	12.47	6.61
2	35.76	21.10	15.51	12.09	5.70
3	38.15	21.74	15.49	12.08	4.88
4	36.43	22.34	15.56	12.02	4.50
Quality	33.04	22.14	15.58	11.89	4.42

likely to explain the quality premium. Thus, two unrelated effects could be correlated with size: a liquidity effect and a quality effect. Looking at size unconditionally conflates liquidity and quality, just as it does size and quality. As such, any liquidity-related size effect should become stronger when controlling for quality by cleaning up the relation between size and liquidity. Consistent with this notion, the spread in liquidity across small versus large stocks is similar across quality groups, highlighting a clear pattern between size and liquidity within each quality group. From our previous evidence, we know that the spread in returns across small versus large stocks is also cleaned up when controlling for quality and is similar across quality groups.

This evidence is consistent with a liquidity-driven story for the size effect, where quality is a separate factor related to size but unrelated to liquidity that confounds the relation between size and liquidity and average returns. While the evidence in Table 8 is consistent with separate liquidity and quality effects related to size, the results do not provide any further direct evidence of a liquidity story for size per se. In addition, the variation in liquidity across the quality dimension within each size group is negligible, consistent with quality not being driven by liquidity.[20]

4.5. Liquidity-based Theories: Liquidity Risk

Some researchers have found that liquidity risk can help explain the size effect (Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005). Absent controlling for quality, the size

Table 9 Can Liquidity Risk Explain Size Controlling for Quality?

Panel A reports regression results for the size premium (small minus big, SMB) on the market return (RMRF), high minus low (HML), up minus down (UMD), the quality factor (QMJ), and two proxies for liquidity risk. IML (illiquid minus liquid) is the return of a portfolio that is long illiquid stocks and short liquid stocks, with liquidity measured as in Amihud (2014). LIQ is the decile spread in returns from portfolios sorted on bid-ask spreads. The sample is 1957–2012, and *t* -statistics are shown in parentheses. Panel B reports returns from double sorted portfolios based on illiquidity and quality or junk. Stocks are sorted independently based on illiquidity and quality into quintiles, and the intersection of those groups form 25 illiquidity-junk portfolios. The value-weighted returns of the 25 portfolios are computed and their averages net of the market are reported along with their *t*-statistics. Two measures of illiquidity are used: Amihud's (2002) measure, which forms the basis of the IML factor used in Amihud (2014), and the market impact cost per dollar traded estimated from Frazzini et al. (2013) assuming a constant fund net asset value (NAV) of $1 billion. Amihud's measure covers the 1957–2012 sample period, and the trading cost data cover the period August 1998 to December 2012.

Panel A: Regressions of SMB on liquidity factors

alpha	RMRF	HML	UMD	QMJ	IML	LIQ
0.20%	0.17	−0.15	−0.00			
(1.70)	(6.49)	(−3.64)	(−0.10)			
−0.19%	0.26	−0.35	0.07		0.79	
(−2.70)	(16.01)	(−13.07)	(3.97)		(33.47)	
0.65%	−0.09	−0.24	0.06	−0.84		
(6.64)	(−3.58)	(−6.69)	(2.45)	(−18.04)		
0.16%	0.07	−0.38	0.10	−0.56	0.68	
(2.72)	(4.44)	(−17.65)	(7.14)	(−19.55)	(34.94)	
0.12%	0.06	−0.34	0.09	−0.51	0.70	0.10
(2.03)	(3.54)	(−15.53)	(6.82)	(−17.59)	(36.56)	(6.53)

Panel B: Illiquidity and junk double sorted portfolios

Portfolio	Illiquid	2	3	4	Liquid	Illiquid – Liquid
Adjusted returns to Amihud (2002, 2014) and junk-sorted portfolios (%)						
Junk	−0.62	−0.52	−0.38	−0.41	−0.52	−0.10
2	0.00	−0.05	−0.08	−0.05	−0.25	0.25
3	0.24	0.02	0.11	0.07	−0.11	0.34
4	0.38	0.26	0.22	0.17	−0.03	0.41
Quality	0.20	0.24	0.27	0.22	0.10	0.10
Quality – Junk	0.82	0.76	0.66	0.63	0.62	
t-statistics for adjusted returns to Amihud (2002, 2014) and junk-sorted portfolios						
Junk	−3.54	−3.43	−2.54	−3.03	−3.95	−0.58
2	0.03	−0.49	−0.81	−0.56	−3.31	1.75
3	1.87	0.19	1.29	1.02	−1.72	2.44
4	3.14	2.77	2.69	2.40	−0.48	2.91
Quality	1.77	2.50	3.10	2.91	1.86	0.71
Quality – Junk	5.20	5.38	4.64	4.81	3.90	

(*Continued*)

Table 9 (***Continued***)

Portfolio	Illiquid	2	3	4	Liquid	Illiquid – Liquid
Adjusted returns to price impact and junk-sorted portfolios (%)						
Junk	−0.31	−1.46	−1.12	−0.85	−1.18	0.87
2	0.13	−0.19	−0.43	−0.25	−0.60	0.72
3	−0.18	0.06	−0.17	0.03	−0.39	0.22
4	0.24	0.11	0.04	0.18	−0.35	0.59
Quality	0.23	0.12	0.07	0.27	−0.23	0.47
Quality – Junk	0.54	1.58	1.19	1.12	0.94	
t-statistics for adjusted returns to price impact and junk-sorted portfolios						
Junk	−0.54	−4.04	−3.74	−3.20	−4.33	1.56
2	0.36	−0.77	−2.21	−1.44	−4.52	2.04
3	−0.83	0.33	−1.04	0.23	−3.54	1.05
4	1.19	0.66	0.26	1.21	−4.19	2.91
Quality	1.02	0.80	0.45	2.10	−2.77	1.87
Quality – Junk	0.95	4.63	4.45	4.70	3.21	

effect is fairly weak and, therefore, is not very difficult to explain with other correlated factors such as liquidity risk. However, controlling for quality, the size premium is substantially stronger, raising the question of whether liquidity risk can help explain these much larger returns.

To address this question, we regress the return to the size factor on standard factors, quality, and liquidity factors. A popular liquidity measure is the price impact measure of Amihud (2002), which is used by Acharya and Pedersen (2005) to construct a liquidity risk factor. To easily interpret the alphas, we use the traded version of this liquidity factor, relying on the illiquid minus liquid (IML) factor of Amihud (2014).We also construct a simple liquidity factor that we denote LIQ, which is the equal-weighted average of the decile return spread in portfolios sorted on turnover and on bid-ask spreads. [We do not use the Pastor and Stambaugh (2003) measure because it is not available over the full sample period, though we obtain similar results over the shorter sample period.]

Panel A of Table 9 reports the results of these regressions over the full quality sample, 1957–2012. The first regression shows that when we control only for the market, value, and momentum, the size effect is weak (*t*-statistic = 1.7). When we control for liquidity risk (in the second regression of the table) based on IML, we find a negative and significant alpha.

Next, the third and fourth regressions show that, when we control for quality, the alpha becomes significantly positive at 65 bps (with a *t*-statistic of 6.64). However, when we control for liquidity risk using the IML factor, we see a large drop in the alpha from 65 bps to only 16 bps. Controlling for both liquidity risk factors, IML and LIQ, we see a further drop in the alpha (the last regression in the table) to 12 bps, which is marginally statistically significant (*t*-statistic of 2.03). This marginally positive alpha could suggest that liquidity risk is not the full explanation for the size premium (once we control for quality) or that the liquidity proxies we use are measured with error and perhaps more precise liquidity measures would drive the alpha to zero. In addition, Acharya and Pedersen (2005) argue that investors want compensation for both the average illiquidity level and liquidity risk of an asset. Controlling for liquidity risk only is often a quick way to capture both liquidity levels and risk (because stocks that have high liquidity risk are likely stocks that are also illiquid

so the estimated liquidity risk premium can be viewed as the sum of compensation for liquidity level and risk), but it is also possible that the level effect is not fully captured here.

Panel B of Table 9 reports results from double sorted portfolios based on illiquidity and quality or junk, in which stocks are sorted independently based on illiquidity and quality into quintiles and the returns in excess of the market are reported on the 25 portfolios formed from the interaction between these sorts. Two measures of illiquidity are used: Amihud (2002), which forms the basis of the IML factor in Amihud (2014), and the market impact measure of Frazzini et al. (2013).The former covers the full sample period from 1957 to 2012, and the latter covers the 1998 to 2012 period. For both measures of illiquidity, evidence exists of an illiquidity premium within each quality category that is directionally consistent with the size premium within each quality category from similar double-sorts in Tables 3 and 5. A significant quality premium remains within each illiquidity category, just as within each size category. Hence, this evidence suggests that measures of illiquidity interact with quality or junk much like the size measures do and that size and illiquidity are very much related and possibly picking up the same return effects.

In summary, small stocks have high bid–ask spreads, have high market impact, and face high liquidity risk. These liquidity level and risk effects can help explain most, or perhaps even all, of the size effect, even when controlling for quality.

5. Conclusion

Size matters, and in a much bigger way than previously thought, after controlling for quality or junk. We find that previous evidence on the variability of the size effect is largely due to the volatile performance of small, low-quality junky firms. Controlling for junk, a much stronger and stable size premium emerges that is robust across time (including periods when the size effect seems to fail), monotonic in size and not concentrated in the extremes, prevalent across months of the year, existent even for non-market-price-based measures of size, and present internationally across nearly two dozen countries. These results are robust across a variety of quality measures, sample periods, within industries, and 23 international markets.[21]

Our evidence that shows a resurgence of the size effect after controlling for quality helps distinguish among many theories for why a size premium could exist. The fact that the size premium rises after controlling for other known risk factors, including growth factors, seems at odds with a simple risk-based explanation or rational explanation based on growth options. The fact that non-price-based measures of size deliver a return premium at least as large as that from market capitalization–based measures when controlling for quality or junk is inconsistent with time-varying risk premia models for the size effect as suggested by Berk (1995a). The fact that measures of limited arbitrage activity such as shorting costs and measures of disagreement such as analyst forecast dispersion suggest that small stocks should be overpriced, implying the opposite of the strong positive size premium we find. Hence, behavioral theories based on limits to arbitrage and over-valuation do not appear to explain our findings. Finally, we show that measures of liquidity and liquidity risk are strongly correlated to and significantly reduce the size premium, even after controlling for quality or junk. This evidence is consistent with liquidity risk-based theories for the size effect. However, a smaller but still significant size premium remains even after controlling for liquidity and liquidity risk, which suggests that either our liquidity measures contain error or liquidity is only a partial explanation for the size effect.

Finally, an interesting avenue for further research is to explore why small, junk stocks in particular do so poorly and why size is strongly negatively correlated with quality. Research by Asness et al. (2014) and Bouchard et al. (2016) seeks to explain the quality premium, with potential implications for its interaction with size.

Our focus in this paper is on the size effect, and our examination of its interaction with quality is aimed at providing further evidence in favor of or against various theories for the size effect. Our results revive the size anomaly, putting it on a more equal footing with other anomalies such as value and momentum in terms of its efficacy, and dismiss several previous explanations and challenges to the size effect. Thus, size, controlling for quality or junk, should be restored as one of the central cross-sectional empirical anomalies for asset pricing theory to explain.

Endnotes

[1]See Dichev (1998), Chan et al. (2000), Horowitz et al. (2000), Gompers and Metrick (2001), Van Dijk (2013), Israel and Moskowitz (2013), Mclean and Pontiff (2016), and Chordia et al. (2014). Schwert (2003) suggests that the small firm anomaly disappeared shortly after the initial publication of the papers that discovered it and coincided with an explosion of small-cap-based funds and indices. Gompers and Metrick (2001) argue that institutional investors' continued demand for large stocks in the 1980s and 1990s increased the prices of large companies relative to small companies, which accounts for a large part of the size premium's disappearance over this period.

[2]Horowitz et al. (2000) find that removing stocks with less than $5 million in market cap eliminates the small firm premium. Crain (2011) and Bryan (2014) find that the small stock effect is concentrated among the smallest 5% of firms. For more about the January effect, see Keim (1983), Reinganum (1983b), and Roll (1983).Gu (2003) and Easterday et al. (2009) also find that the January effect has declined over time, coinciding with the decline in the small firm premium. Van Dijk (2013) finds the same in a review of the size literature.

[3]A variety of quality measures have been proposed in the literature, including profitability (Graham and Dodd, 1934; Novy-Marx, 2013), investment (Fama and French, 2015 ; Hou and Van Dijk, 2017), growth (Lakonishok et al., 1994; Mohanram, 2005), low asset growth (Cooper et al., 2008), low use of accruals (Sloan, 1996; Richardson et al., 2005), payout (Baker and Wurgler, 2002; Pontiff and Woodgate, 2008), low risk (Ang et al., 2006, 2009; Black et al., 1972; Frazzini and Pedersen, 2014), low leverage (George and Hwang, 2010; Penman et al., 2007), low credit risk (Altman, 1968; Ohlson, 1980; Campbell et al., 2008), and good governance (Gompers et al., 2003; Bebchuk et al., 2009; Core et al., 2006; Cremers and Nair, 2005; Giroud and Mueller, 2011; Johnson et al., 2009); see also Cremers and Ferrell (2014) and Larcker et al. (2015) for reviews. Asness et al. (2014) summarize various measures and dimensions of quality and construct a composite quality index based on a number of proposed variables. They find that higher quality is associated with higher average prices but also with higher expected returns

[4]Our SMR factor is a slightly different version of the BAB strategy from Frazzini and Pedersen (2014) constructed following the methodology of Fama and French (1993). For consistency with the other portfolios used in this paper, we rank stocks based on their ex-ante beta and form a portfolio that is long low beta and short high beta stocks based on the intersection of six size- and beta-sorted portfolios.

[5]The Asness, Frazzini, and Pedersen composite index can be downloaded at https://www.aqr.com/library/data-sets/quality-minus-junk-factors-monthly.

[6]Common stocks are identified by a Compustat issue code (TPCI) of 0. We also drop stocks traded on over-the-counter (OTC) exchanges.

[7]To obtain shareholders' equity we use stockholders' equity (SEQ). If it is not available, we use the sum of common equity (CEQ) and preferred stock (PSTK). If both SEQ and CEQ are unavailable, we proxy shareholders' equity by total assets (TA) minus the sum of total liabilities (LT) and minority interest (MIB). To obtain book equity, we subtract from shareholders' equity the preferred stock value (PSTKRV, PSTKL, or PSTK depending on availability). Finally, to compute book value per share (B), we divide by common shares outstanding (CSHPRI). If CSHPRI is missing, we compute company-level total shares outstanding by summing issue-level shares (CSHOI) at fiscal year-end for securities with an earnings participation flag in the security pricing file.

[8]Harvey et al. (2014) note that *t*-statistics greater than 3.0 are likely required to pass the 5% significance test in the presence of the data mining that has taken place by researchers poring over the same return series, but we note two caveats. On the one hand, the size effect is simple and was discovered several decades ago so it could be less subject to data mining. On the other hand, our resurrection of size is based on the interaction with quality and, as Harvey et al. (2014) point out, conditional strategies could have different, potentially higher, statistical cutoffs to judge significance in light of potential data mining.

[9]Over time, CRSP has fixed many data errors, which are more common among the smallest firms, and these could have contributed positively to the returns of size. One such error was a delisting bias as noted by Shumway (1997), who shows that many studies focusing on small stocks had inflated returns due to mistreatment of the delisting returns to these stocks.

[10]Mclean and Pontiff (2016) claim that many anomalies provide their best returns over the sample period in which they were originally discovered.

[11]For the period from January 1931 to June 1957, which is completely out of sample from the period when all of the QMJ variables are measured, we find that BAB increases the size alpha from 3 to 16 bps, though the alpha is statistically insignificant given the smaller sample size. The coefficient on BAB is −0.35 with a *t*-statistic of −4.99, indicating that size and BAB are strongly related in the out-of-sample period as well and BAB helps resurrect the size premium as a result.

[12]To summarize these results, we also create an industry-neutral size factor by averaging across all of the industry SMB portfolios (equal-weighted across industries) and regress the returns on the market, lagged market, and industry-neutral versions of the HML, UMD, and QMJ factors all constructed in the same manner. Regressing SMB industry-neutral on these other industry-neutral factors, excluding QMJ, we find an insignificant alpha of 1 bp with a *t*-statistic of 0.13. Adding the QMJ industry-neutral factor to the regression resurrects SMB's alpha to 38 bps with a highly significant *t*-statistic of 5.00. The coefficient of industry-neutral SMB on QMJ is −0.69 with a *t*-statistic of −15.64. Hence, industry-neutral versions of all the factors show the same pattern, where the size premium is resurrected by controlling for quality even when all industry variation is removed from all of the factors.

[13]Fig. OA1 in the Online Appendix plots the year-to-year autocorrelation of portfolio weights based on quality characteristics of stocks for the four different measures of quality: profitability, growth, safety, and payout. For each year, we take the set of stocks that exist at year *t* and *t*−1 and compare the autocorrelation of their portfolio weights, which are based on quality rankings from year *t*−1 to t. The plot shows tremendous persistence to the quality measures, especially for the first three, which helps explain why size's exposure to quality is stable through time.

[14]Also, our comments largely refer to the bulk of the sample occurring after the initial period (after about 1960–1965). The very early part of the sample shows a somewhat more even distribution of size amongst both high-quality and junk stocks.

[15]Although the portfolios in Table 6 are constructed from independent sorts on quality and size, dependent-sorted portfolios based on five quintiles of quality and then, within those quintiles, another set of quintile portfolios based on size, yield similar results.

[16]This evidence also contradicts theories for the seasonality of size related to turn-of-the-year price pressure due to tax-loss selling, window dressing by institutions, or cash infusion of investors.

Controlling for quality, the turn-of-the-year evidence on size is much weaker and the size effect is more evenly distributed throughout the year. Hence, while price pressure effects in December and January could still be present, they are less about small versus large stocks than previously thought.

[17]Repeating these regressions over the three subperiods (golden age, embarrassment, and resurrection), the non-price-based size measures fail to generate significant alphas on their own. However, once we control for quality, all non-price-based size measures yield very significant alphas. Controlling for quality also makes the non-price-based size premia stable across the different subsamples. In addition, Fig. OA2 in the Online Appendix reports results from the intra-industry exercise we conducted earlier, but using SMB portfolios formed from the non-price-based size measures instead of market capitalization. Within each of the 30 industries, we form SMB portfolios based on book assets, sales, book equity, PP&E, and number of employees. We then regress the non-price-based SMB returns on the market, its lagged value, and HML and UMD factors and regress the SMB returns on these same factors plus a quality factor. The differences between the alphas are then plotted industry by industry in Fig. OA2, representing the improvement in SMB performance from controlling for quality or junk. Fig. OA3 plots the change in alpha before and after controlling for quality by country for non-price-based measures of size using book assets, sales, book equity, PP&E, and employees to sort stocks within each country. For the vast majority of countries, there is a significant size premium even for non-price-based measures of size once we control for quality. These results provide even more evidence of a robust size effect internationally as well as a large number of out-of-sample tests for non-price-based size measures that help alleviate any data mining concerns.

[18]We obtain similar results using other measures of shorting costs from Markit Data Explorers. We use other cost measures, such as the average fee of stock borrow transactions from hedge funds, expressed as a percentage (SAF), and the value-weighted average cost, a number from one to five indicating the fee for each stock (VWAF).

[19]The price impact trading cost data used by Frazzini et al. (2014) to calibrate their model and estimate its parameters pertains to stocks that are predominantly in the top 60% of market capitalization, because their execution database does not trade stocks below this threshold. While the model includes a size factor to allow the trading costs to scale with the size of the firm and its average daily trading volume, the functional form of this relation is estimated only from the larger stocks in their database. Therefore, extrapolating estimates of price impact costs for the smallest, particularly micro-cap stocks, could be less accurate.

[20]The only place where a hint of variation could exist in liquidity measures across quality is among the smallest stocks based on bid–ask spread and among the largest stocks for price impact, in which the junkiest firms appear a tiny bit less liquid than the quality firms. The relation between quality and these liquidity measures, however, is not consistent within a size group. For bid-ask spread, it is the highest quality firms among the smallest quintile of stocks that appear different from the other quality quintiles within that group. For price impact, it is the lowest quality firms among the largest stocks that appear different from the other quality quintiles. Moreover, no consistency is evident across the other size quintiles (2 through 4), in which no discernible relation between quality and liquidity is apparent.

[21]Combining these results, Figures OA4 and OA5 in the Online Appendix show seasonal patterns in size controlling for quality across time, among the smallest stocks, and using non-price-based size measures. Controlling for quality or junk has the effect of smoothing the returns to size, establishing a clear size premium that is no longer concentrated in January and, most importantly, is significant outside of January and no longer concentrated in certain time periods or for certain measures. The behavior of small, junk firms varies substantially and is chiefly responsible for diluting the size effect at certain times, for certain months, and for certain measures and exaggerating it at other times. Controlling for these firms through exposure to a quality factor, the size premium is robust over time, and for different measures of size, even those not based on market prices.

References

Acharya, V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. J. Financ. Econ. 77, 375–410.

Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Finance 23, 589–609.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time series effects. J. Financ. Mark. 5, 31–56.

Amihud, Y., 2014. The Pricing of the Illiquidity Factor's Systematic Risk. New York University, Stern School of Business Unpublished working paper.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. J. Financ. Econ. 17, 223–249.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. J. Finance 61, 259–299.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. J. Financ. Econ. 91, 1–23.

Asness, C.S., Frazzini, A., 2012. The devil in HML's details. J. Port Manage. 39, 49–69.

Asness, C.S., Frazzini, A., Pedersen, L.H., 2014. Quality minus junk. AQR Capital Management, Greenwich, CT Unpublished working paper.

Asness, C.S., Frazzini, A., Pedersen, L.H., 2017. Quality Minus Junk. AQR Capital Management, Greenwich, CT Unpublished working paper.

Asness, C.S., Krail, R.J., Liew, J.M., 2001. Do hedge funds hedge? J. Portf. Manag. 28, 6–19.

Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. J. Finance 68, 929–985.

Baker, M., Wurgler, J., 2002. Market timing and capital structure. J. Finance 57, 1–32.

Ball, R., 1978. Anomalies in relations between securities' yields and yield-surrogates. J. Financ. Econ. 6, 103–126.

Banz, R., 1981. The relationship between return and market value of common stocks. J. Financ. Econ. 9, 3–18.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. J. Financ. Econ. 49, 307–343.

Bebchuk, L., Cohen, A., Ferrell, A., 2009. What matters in corporate governance? Rev. Financ. Stud. 22, 783–827.

Berk, J., 1995a. A critique of size-related anomalies. Rev. Financ. Stud. 8, 275–286.

Berk, J., 1995b. An Empirical Reexamination of the Relation Between Firm Size and Return. University of Washington Unpublished working paper.

Berk, J., 1997. Does size really matter? Financ. Anal. J. 53, 12–18.

Black, F., 1972. Capital market equilibrium with restricted borrowing. J. Bus. 45, 444–455.

Black, F., 1992. Beta and return. J. Portf. Manag. 20, 8–18.

Black, F., Jensen, M.C., Scholes, M., 1972. The capital asset pricing model: some empirical tests. In: Jensen, M.C. (Ed.), Studies in the Theory of Capital Markets. Praeger Publishers, New York, pp. 79–121.

Bouchaud, J.-P., Ciliberti, S., Landier, A., Simon, G., Thesmar, D., 2016. The Excess Returns of "Quality" Stocks: A Behavioral Anomaly. HEC Paris, Paris, France Research paper no. FIN-2016-1134.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. J. Financ. Econ. 49, 441–464.

Bryan, A., 2014. Does the Small-Cap Premium Exist?. Morningstar Inc., Chicago, IL.

Campbell, J., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. J. Finance 63, 2899–2939.

Carlson, M., Fisher, A., Giammarino, R., 2004. Corporate investment and asset price dynamics: implications for the cross-section of returns. J. Finance 59, 2577–2603.

Chan, L.K.C., Karceski, J., Lakonishok, J., 2000. New paradigm or same old hype in equity investing? Financ. Anal. J. 56, 23–36.

Chordia, T., Subrahmanyam, A., Tong, Q., 2014. Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? J. Account. Econ. 58, 41–58.
Cooper, M., Gulen, H., Schill, M., 2008. Asset growth and the cross-section of stock returns. J. Finance 63, 1609–1651.
Core, J.E., Guay, W.R., Rusticus, T.O., 2006. Does weak governance cause weak stock returns? An examination of firm operating performance and investors' expectations. J. Finance 61, 655–687.
Crain, M.A., 2011. A Literature Review of the Size Effect. Financial Valuation Group, Tampa, FL Unpublished working paper.
Cremers, M., Ferrell, A., 2014. Thirty years of shareholder rights and firm value. J. Finance 69, 1167–1196.
Cremers, M., Nair, V., 2005. Governance mechanisms and equity prices. J. Finance 60, 2859–2894.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. A theory of overconfidence, self-attribution, and security market under and over-reactions. J. Finance 53, 1839–1885.
De Bondt, W.F.M., Thaler, R., 1985. Does the stock market overreact? J. Finance 40, 793–805.
Dichev, I.D., 1998. Is the risk of bankruptcy a systematic risk? J. Finance 53, 1131–1148.
Diether, K., Malloy, C., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. J. Finance 57, 2113–2141.
Easterday, K., Sen, P., Stephan, J., 2009. The persistence of the small firm/January effect: is it consistent with investors' learning and arbitrage efforts? Q. Rev. Econ. Finance 49, 1172–1193.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.
Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. J. Finance 51, 55–84.
Fama, E.F., French, K.R., 1997. Industry costs of equity. J. Financ. Econ. 43, 153–193
Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. J. Financ. Econ. 105, 457–472.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116, 1–22.
Fama, E.F., French, K.R., 2016. Dissecting anomalies with a five-factor model. Rev. Financ. Stud. 29, 69–103.
Frazzini, A., Israel, R., Moskowitz, T.J., 2013. Trading costs of asset pricing anomalies. University of Chicago, Chicago, IL Unpublished working paper.
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. J. Financ. Econ. 111, 1–25.
Garleanu, N., Panageas, S., Yu, J., 2012. Technological growth and asset pricing. J. Finance 67, 1265–1292.
George, T.J., Hwang, C.Y., 2010. A resolution of the distress risk and leverage puzzles in the cross section of stock returns. J. Financ. Econ. 96, 56–79.
Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121–1152.
Giroud, X., Mueller, H.M., 2011. Corporate governance, product market competition, and equity prices. J. Finance 66, 563–600.
Gompers, P., Metrick, A., 2001. Institutional investors and equity prices. Q. J. Econ. 116, 229–259.
Gompers, P., Ishii, J., Metrick, A., 2003. Corporate governance and equity prices. Q. J. Econ. 118, 107–156.
Graham, B., Dodd, D., 1934. Security Analysis. McGraw-Hill, New York.
Gu, A.Y., 2003. The declining January effect: evidences from the U.S. equity markets. Q. Rev. Econ. Finance 43, 395–404.
Harvey, C., Liu, Y., Zhu, H., 2014. ..... and the Cross-Section of Expected Returns. Duke University, Durham, NC Unpublished working paper.
Hong, H., Stein, H., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. J. Finance 54, 2143–2184.

Horowitz, J.L., Loughran, T., Savin, N.E., 2000. Three analyses of the firm size premium. J. Empir. Finance 7, 143–153.
Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross–section of expected returns. Rev. Financ. Stud. 18, 981–1020.
Hou, K., Van Dijk, M.A., 2017. Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns. Ohio State University, Columbus, OH Unpublished working paper.
Ibbotson, R., Chen, Z., Kim, D.Y.J., Hu, W., 2013. Liquidity as an investment style. Financ. Anal. J. 69, 30–44.
Israel, R., Moskowitz, T.J., 2013. The role of shorting, firm size, and time on market anomalies. J. Financ. Econ. 108, 275–301.
Johnson, S., Moorman, T., Sorescu, S., 2009. A reexamination of corporate governance and equity prices. Rev. Financ. Stud. 22, 4753–4786.
Keim, D.B., 1983. Size-related anomalies and stock return seasonality: further empirical evidence. J. Financ. Econ. 12, 13–32.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation and risk. J. Finance 49, 1541–1578.
Larcker, D.F., Reiss, P.C., Xiao, Y., 2015. Corporate Governance Data and Measures Revisited. Stanford Graduate School of Business, Stanford, CA Unpublished working paper no. 3379.
Lo, A., MacKinlay, C., 1988. Stock market prices do not follow random walks: evidence from a simple specification test. Rev. Financ. Stud. 1, 41–66.
Mclean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? J. Finance 71, 5–32.
Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. J. Finance 32, 1151–1168.
Mohanram, P.S., 2005. Separating winners from losers among low book–to-market stocks using financial statement analysis. Rev. Account. Stud. 10, 133–170.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108, 1–28.
Ohlson, J., 1980. Financial ratios and the probabilistic prediction of bankruptcy. J. Account. Res. 1980, 109–131.
Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111, 642–685.
Penman, S., Richardson, S., Tuna, I., 2007. The book-to-price effect in stock returns: accounting for leverage. J. Account. Res. 45, 427–467.
Pontiff, J., Woodgate, W., 2008. Share issuance and cross-sectional returns. J. Finance 63, 921–945.
Reinganum, M., 1983a. Portfolio strategies based on market capitalization. J. Portf. Manag. 9, 29–36.
Reinganum, M., 1983b. The anomalous stock market behavior of small firms in January: empirical tests for tax-loss selling effects. J. Financ. Econ. 12, 89–104.
Richardson, S., Sloan, R.G., Soliman, M., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. J. Account. Econ. 39, 437–485.
Roll, R., 1983. The turn-of-the-year effect and the return premia of small firms. J. Portf. Manag. 9, 18–28.
Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: the role of liquidity risk. J. Financ. Econ. 80, 309–349.
Schwert, G.W., 2003. Anomalies and market effciency. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance, 1. Elsevier, pp. 939–974.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. J. Finance 52, 35–55.
Shumway, T., 1997. The delisting bias in CRSP data. J. Finance 52, 327–340.
Sloan, R.G., 1996. Do stock prices reflect information in accruals and cash flows about future earnings? Account. Rev. 71, 289–315.
Van Dijk, M.A., 2013. Is size dead? A review of the size effect in equity returns. J. Bank. Finance 35, 3263–3274.

PART VI

Taking Research to the Real World

The Devil in HML's Details
Fact, Fiction, and Momentum Investing
Fact, Fiction, and Value Investing
Craftsmanship Alpha

This page intentionally left blank

The Devil in HML's Details

Clifford Asness and Andrea Frazzini

Few papers focus solely on subjects as seemingly innocuous as timely, frequent updates of price (P) in calculations of book-to-price ratio (B/P). But rarely is so innocuous a choice worth between 305 and 378 basis points annually of statistically significant alpha, plus an ability to illuminate important aspects of the dynamics between value and momentum strategies.[1]

This article focuses on a seemingly small detail in the construction of portfolios that are long value stocks and short growth stocks, often referred to as HML (high minus low).[2] The most common construction, as pioneered by Fama and French [1992], uses B/P as the proxy for value, and forms a portfolio that is long high-B/P firms and short low-B/P firms. A high B/P means a stock is cheap (or high risk, to efficient market fans) and has a high expected return. A low B/P means the opposite.

In calculating B/P for each stock and forming a value strategy, this method updates value once a year on June 30, using book and price as of the prior December 31. It then holds those values (and portfolio holdings) constant until rebalancing the portfolio the following June 30. In other words, both the book and price data used to form B/P and value portfolios are always between six and 18 months old.

Fama and French [1992] made these conservative construction choices to make sure that book value would actually be available at the time of portfolio construction and/or rebalancing. They then presumably chose to use price from the same date as book, based on common sense. To measure B/P, using book and price from the same date might be the obvious choice.

We believe that this was entirely reasonable, particularly in the early days of the literature, when momentum was not a literal or figurative factor. Now, however, it is suboptimal.

Most of this article focuses on the question of whether we should lag price in constructing valuation ratios. Unlike book value, we know with certainty that the June 30 price is available on the June 30 rebalance date, giving us a choice of computing valuation ratios based on lagged fiscal year-end prices or on current prices. We show that using a more-current price is superior to the standard method of using prices at fiscal year-end as a proxy for the true B/P ratio, and superior in five-factor model regressions. This improvement can lead to a significantly better portfolio combined strategy, and also sheds light on the dynamic relationship between value and momentum. When we use factor models to judge other strategies or for performance attribution, this strategy implicitly raises the bar.

Reprinted from *The Journal of Portfolio Management*, Summer 2013, Vol. 39 (4), with permission.
Recipient of the 2014 Bernstein Fabozzi/Jacobs Levy Best Article Award.

Consider a stock with a December fiscal year-end date and a price that fell 75 percent between December 31 and the June 30 rebalance date, when you must decide if this is a value stock. Does the fallen price make this more likely, less likely, or have no effect on whether this should be considered a value stock?

The answer depends on how much variation in B/P ratios is due to expected returns and how much is due to changes in future book values. Our findings show that true value stocks often show such price drops, and a measure that takes this fall into consideration, as our proposed method does, is superior to one that ignores it, as the standard method does. It is superior not because we create a better stand-alone value strategy—one might naively think that more timely updating improves any stand-alone strategy[3]—but because it better handles the complex relationship between value and momentum strategies.

Data, Methodology, and Terminology

Data Sources

Our U.S. equity data includes all available common stocks on the merged CRSP/XpressFeed data between July 1950 and March 2011. Our global equity data includes all available common stocks on the XpressFeed Global database for 19 developed markets. The international data runs from January 1983 to March 2011.[4] We report our sample's summary statistics in the appendix.

To compute total book value of equity (BE), we prefer stockholders' equity (SEQ). If that is unavailable, we use the sum of common equity (CEQ) and preferred stock (PSTK). If both SEQ and CEQ are unavailable, we proxy book equity by total assets (AT), minus the sum of total liability (LT), minority interest (MIB), and preferred stocks (PSTKRV, PSTKL, or PSTK, depending on availability). To compute book value per share (B), we divide by common shares outstanding (CSHPRI). If CSHPRI is missing, we compute company-level total shares outstanding by summing issue-level shares (CSHOI) at fiscal year-end for securities with an earnings participation flag in the security-pricing file. Following Fama and French [1992], we assume that accounting variables are known with a minimum six-month gap, and align the firm's book price at fiscal year-end, which is anywhere in year $t-1$ to June of calendar year t. To be included in any of our tests, a firm must have a non-negative book price and non-missing price at fiscal year-end, as well as in June of calendar year t.

Constructing Value Measures

We focus on a seemingly small modification to standard practice—one that we think is not so small in its impact. We compute three measures of B/P. The first is Fama and French's [1992] standard approach, with B/P equal to the book value per share (B) divided by price at fiscal year-end (P_{fye}), both in local currency:[5]

$$bp_t^{annual,lagged} \equiv bp_t^{a,l} = \log(B/P_{fye})$$

We label this measure *annual* (indicated by the superscript *a*), as it is updated once a year, and *lagged* (indicated by the superscript *l*), as at the update, it uses prices from six to 18 months ago, not current prices.[6]

The second measure is equal to book value per share (adjusted for splits, dividends, and other corporate actions between fiscal year-end and portfolio formation dates), divided by current price P_t, both in local currency:

$$bp_t^{annual,current} \equiv bp_t^{a,c} = \log(B^* / P_t)$$

where $B^* = B \times Adj_t / Adj_{fye}$ and *Adj* is the cumulative adjustment factor. This alternative measure holds Fama and French's method [1992] constant, save for choosing date to use for price.[7] We call this measure *annual,* as it is updated once a year (indicated by the superscript *a*) and *current,* as it uses the most recent available price as of the June 30 rebalance date (indicated by the superscript *l*). "Current" refers only to price at time of portfolio formation, not to book value, which our measures always lag.

Our last measure is equal to book value per share, divided by current price and updated monthly:

$$bp_t^{monthly,current} \equiv bp_t^{m,c} = \log(B^* / P_t)$$

In our naming convention (indicated by the superscript *m*), *monthly* applies only to price. Our convention for book value remains the same as the standard for all three measures. This measure is equal to $bp_t^{a,c}$ in June of each year, but is updated every month using current prices, as opposed to staying constant through the year. Through the paper we will maintain this notation convention: The first superscript indicates the refreshing frequency (annual *a* or monthly *m*); the second superscript indicates the lag used to update price (lagged *l* or current *c*).

Exhibit 1 illustrates the three approaches for a firm with a fiscal year ending in December 2000. To summarize each of the three measures, use the same measure of book value (lagged at least six months at portfolio formation date), but vary the lag used to update price. Note that $bp_t^{a,l}$ is the widely used method in academic finance; we refer to it as the *standard method*. $bp_t^{a,c}$ is the same measure using price as of June 30, not as of the prior December 31, then leaving both book and price unchanged for the next 12 months. $bp_t^{m,c}$ is the same ratio with price updated monthly.

The three measures are mechanically related:

$$bp_t^{a,c} = bp_t^{a,l} - r_{fye \to t}$$
$$bp_t^{m,c} = bp_t^{a,c} - r_{t \to t+k}$$

where $r_{t \to s} = \log\left(1 + R_{t \to s}\right)$ is equal to the total log return between date *t* and *s* > *t*. Hence the choice between the different measures is equivalent to choosing whether we should ignore or include recent returns when building value portfolios.

This choice matters most when we combine these portfolios with momentum or short-term reversal portfolios, which are themselves direct bets on recent returns. We'll form value strategies using all three methods for estimating B/P. We describe the details of portfolio construction in the next subsection.

Portfolios

Our portfolio construction closely follows Fama and French [1992, 1993, 1996]. Our global factors are country neutral. That is, we form one set of portfolios in each country and compute a global factor by weighting each country's portfolio by the country's total (lagged) market capitalization.

The market factor MKT is the value-weighted return on all available stocks, minus the one-month Treasury bill rate.

We construct the size and value factors using six value-weighted portfolios formed on size and B/P. At the end of June of year *t,* stocks are assigned to two size-sorted portfolios, based on their market capitalization. For the U.S., the size breakpoint is median NYSE market equity. For the international sample, the size breakpoint is the 80th percentile

$BP^{a,l}_{June\ 30th\ 2001}$ = Book (December 31, 2000)/Price (December 31, 2000)

$BP^{a,c}_{June\ 30th\ 2001}$ = Book (December 31, 2000)/Price (June 30, 2001)

$BP^{m,c}_{June\ 30th\ 2001}$ = Book (December 31, 2000)/Price (June 30, 2001, updated monthly thereafter)

Exhibit 1 Example: B/P Calculation for a Firm with Fiscal Year Ending in December 2000. This exhibit illustrates the three approaches used to compute B/P for a firm with a fiscal year ending in December 2000. $bp_t^{a,l}$ is equal to the book value per share (B) divided by price at fiscal year-end (P_{fye}) both in local currency, $bp_t^{a,c}$ is equal to book value per share (adjusted for splits, dividends, and other corporate actions between fiscal year-end and portfolio formation dates). $bp_t^{m,c}$ is equal to book value per share divided by current price, updated monthly. In the name convention, the first superscript indicates the refreshing frequency (annual *a* or monthly *m*), and the second superscript indicates the lag used to update price (lagged *l* or current *c*).

by country.[8] Portfolios are value-weighted, refreshed every June, and rebalanced every calendar month to maintain value weights. The size factor SMB (small minus big) is the average return on the three small portfolios, minus the average return on the three big portfolios:[9]

$$\begin{aligned}\text{SMB} = {} & 1/3\ \left(\text{Small Value} + \text{Small Neutral} + \text{Small Growth}\right) \\ & -1/3\ \left(\text{Big Value} + \text{Big Neutral} + \text{Big Growth}\right)\end{aligned}$$

The value factor's HML is the average return on the two value portfolios, minus the average return on the two growth portfolios:

$$\begin{aligned}\text{HML} = {} & 1/2\left(\text{Small Value} + \text{Big Value}\right) \\ & -1/2\left(\text{Small Growth} + \text{Big Growth}\right)\end{aligned}$$

We construct a version of HML for each annual measure: $\text{HML}^{\text{annual,lagged}} \equiv \text{HML}^{a,1}$ and $\text{HML}^{\text{annual,current}} \equiv \text{HML}^{a,c}$. Finally we construct a version of HML for our monthly B/P measure, $\text{HML}^{\text{monthly,current}} \equiv \text{HML}^{m,c}$, in the same manner, but this portfolio is refreshed monthly. All our portfolios are rebalanced monthly, to keep value weights. *Refreshed* refers to the date that we update value and size breakpoints—once a year in for the annual measures, every month for the monthly measure—not to the rebalancing frequency for value weighting, which is the same for all three.[10]

We construct the momentum and short-term reversal portfolios in a similar way. We use six value-weighted portfolios, formed on size and prior returns. The portfolios are the intersections of two portfolios formed on size and three portfolios formed on prior returns. We use one-year return (in local currency), skipping the most recent month for momentum (UMD) and (minus) the local currency return in the most recent month for short-term reversal (STR):

$$\begin{aligned}\text{UMD} &= 1/2\left(\text{Small High} + \text{Big High}\right) \\ &\quad -1/2\left(\text{Small Low} + \text{Big Low}\right) \\ \text{STR} &= 1/2\left(\text{Small Low} + \text{Big Low}\right) \\ &\quad -1/2\left(\text{Small High} + \text{Big High}\right)\end{aligned}$$

We refresh both portfolios every calendar month, and rebalanced monthly to maintain value weights.

All portfolio returns are in U.S. dollars; excess returns are above the U.S. Treasury bill rate.[11] Because some of our variables are computed from closing prices, we skip one trading day between portfolio formation and investment in all portfolios, both when refreshing the breakpoints and when rebalancing stocks in the portfolio.[12]

What Proxies Best for the True Unobservable B/P?

In this section, we run a horse race using cross-sectional regressions of current B/P on lagged B/P and highlight the relative forecasting power of the different measures. Imagine you're standing at December 31, 2000, and you want to form a value portfolio based on B/P. The measure you want is:

$$B^{Unobservable}_{as\ of\ Dec\,2000} \equiv BP^{a,l}_{t}$$
$$= \text{Book (December 31, 2000) / Price (December 31, 2000)}$$

But that measure isn't available, because book value as of December 31, 2000, is not known until sometime after that date. (Hence the standard six-month lag and our "Unobservable" superscript.) But as of December 31, 2000, you have two available measures:

$$BP^{a,l}_{t-1} = \text{Book (December 31, 1999) / Price (December 31, 1999)}$$
$$BP^{a,c}_{t-1} = \text{Book (December 31, 1999) / Price (June 30, 1999)}$$

Use either or any combination of both to form a forecast of $BP^{Unobservable}$.

We directly test this question: Does the standard method, $BP^{a,l}_{t-1}$, which aligns price and book, or our proposed method, $BP^{a,c}_{t-1}$, which uses more current prices (thus incorporating more recent returns), make a better proxy for unobservable B/P? As we are not yet testing our monthly refreshed method, both of these use a lagged price here, as both portfolios are only refreshed annually at the end of June, and we're examining the end of December. Our proposed $BP^{a,c}_{t-1}$ is simply less lagged than the standard $BP^{a,l}_{t-1}$.

We run Fama and MacBeth [1973] regressions of the unobservable B/P on competing versions of past B/P, plus an error-correction term:

$$bp^{a,l}_{t} = \gamma_0 + \gamma_1 bp^{a,l}_{t-1} + \gamma_2 \left(bp^{a,c}_{t-1} - bp^{a,l}_{t-1} \right) + \in_t$$

We test which of two observable proxies does a better job of explaining the unobservable, as of December 31.[13]

We run cross-sectional regressions each year for all firms in our universe. The left side is the unobservable, true B/P, for which we'd like to get the closest proxy. We can interpret coefficient γ_1 as the weight we would put on the standard B/P version in the literature. We can also interpret coefficient γ_2 as the amount by which we would move away from this standard version towards our new version, which differs by its more timely, less-lagged use of price. With some rearranging, we can interpret γ_2 and $\gamma_1 - \gamma_2$ as the linear weights we

would put on the different measures in a linear forecast. (The measure we cast as the starting point is, of course, irrelevant.):

$$\widehat{bp_{t+1}^{a,l}} = \gamma_0 + \left(\gamma_1 - \gamma_2\right) bp_t^{a,l} + \gamma_2 bp_t^{a,c}$$

Exhibit 4 reports the time-series averages of the cross-sectional estimates of γ_1, γ_2, and $\gamma_1 - \gamma_2$ and the corresponding *t*-statistics of the time-series of point estimates. We also report γ_2/γ_1, interpreted as the fraction of linear forecast attributed to our more timely $bp^{a,c}$. (We attribute the remainder to the standard method $bp^{a,l}$).

We focus on the all-sample U.S. results in the first row of Exhibit 4, panel A. In the appendix we report robustness checks across fiscal year, industry, size deciles, and time. The point estimate for γ_2 is 0.86, meaning that we would move 86 percent of the distance from the standard lagged B/P towards our proposed current B/P. The *t*-statistic for this move is 38.9. Alternatively, had we switched the order, started with our new measure, and reported how far to move towards the standard, we could still reject the null hypothesis of no incremental value of the standard lagged B/P versus our proposed current B/P, as measured by $\gamma_1 - \gamma_2$, but the effect is negligible (0.05 with a *t*-statistic of 3.14). The rightmost column gives a more intuitive way of looking at the results. Scaled to 100 percent, we would base

Exhibit 2 HML: Global Sample, Cumulative Five-Factor Alphas, 1950–2011. This exhibit plots cumulative portfolio alphas. We run time-series regressions on monthly excess returns of value portfolios (HML) and on monthly excess returns of a set of explanatory portfolios, then plot cumulative alphas.

Exhibit 3 Case Study: HML and UMD in 2009, Global Sample, Total Returns. This exhibit plots total returns. We plot cumulative returns of value (HML) and momentum (UMD) portfolios (HML) between February 2009 and July 2010.

94 percent of our linear forecast of the unobservable goal on our proposed current B/P, and only 6 percent on the standard method. All robustness checks are reasonably close to the all-sample results. Essentially, in a simple evaluation of what measure best proxies for the clear but unobtainable goal—true, timely B/P)—our proposed change is a clear winner.

The international results in panel B are strikingly consistent with our U.S. results and highly support our proposed method of computing B/P over the standard specification (although, of course, based on a shorter sample). In our international sample, we would base between 84 percent and 96 percent of our forecast of the unobservable goal on our current method.

We report a series of robustness checks in the Appendix. All the results tell a consistent story: recent returns matters, i.e., to proxy for the unobservable, true B/P, our new, more-timely measure is superior to the standard measure that unnecessarily lags price to match the necessary lag in book.

Exhibit 5 and Exhibit A3 in the appendix provide some information about the reasons that standard B/P is a worse proxy for the true, unobservable B/P. We run Fama and MacBeth's [1973] regression of log changes in book price per share on log returns over the past three years:[14]

$$\Delta b^*_{t-12\to t} = \theta_0 + \theta_1 r_{t-12\to t} + \theta_2 r_{t-24\to t-12} + \theta_3 r_{t-36\to t-12} + \in_t$$

In other words, we study how much a given price change translates into a change in book value. The all-sample results show that, in a given year, somewhere around 22 percent of a price move in the prior 12 months is reflected in a contemporaneous change in book price. Eventually, including all three lags, this total rises to a percentage in the mid-40s. Looking at the international sample, we find similar results, with the three-year totals almost hitting 50 percent.[15] To summarize, current and prior returns predict future changes in book value, but in an attenuated fashion, with coefficients of well below 100 percent.

How does this provide intuition for our Exhibit 5 results? For impacts of one or three-plus years, between 20 percent and 40 percent of price movements seem to be currently or eventually reflected in book value. Therefore, if someone told us about a strong price move, our first guess would not be that true B/P was unaffected. Rather, we would guess that, if price fell sharply, true B/P would rise sharply, though not quite to the full extent of the price move.

Thus, the standard method of measuring B/P, which unnecessarily lags price to match the necessary lag in book, is not our best guess of true B/P. Our best guess of true B/P would use most of any observed price move, even if that move was not aligned with the latest observable book value.

Had these coefficients summed to near 100 percent, our best guess of true B/P would indeed be approximately the standard method. Our more timely method did not have to create the large improvement we observe. But price moves much more than book, causing the standard method to miss important information.[16]

Does the Standard or New, More Timely B/P Form a Better Value Portfolio?

We have shown that our proposed more timely measure is a better proxy for true value than is the traditional measure. If the goal is value investing, one could advocate using our more timely proxy on first principles. The rest of this article is an attempt to discover how much this first principle really matters.

In Exhibit 6, we examine portfolio returns. We run time-series regressions and test whether each version of HML adds value in the presence of the other competing HML, the market (MKT), a size factor (SMB), and a short-term reversal factor (STR). We discuss results for the U.S. but also report tests for our international sample, as well for the full set of countries aggregated in a global portfolio.[17]

Columns 1 and 2 report results for our two annual value measures. They are run against five-factor models, including the other competing value measure. When fully controlling for factor exposure, the standard $HML^{a,l}$ approach subtracts –58 basis points (bps) annually (which is statistically insignificant). Our more timely $HML^{a,c}$ factor adds 143 bps (which is statistically significant) over the traditional four-factor model, augmented with short-term reversal.[18] In other words, in the presence of the other factors, our newer, more timely approach is clearly better than the standard lagged approach.

In the previous section we showed that, in ignoring returns and focusing only on proxies for true ex post B/P, our more timely measure is superior. In this section we add that, in the presence of momentum, and for logical reasons having to do with the overlap of our value measure and the period used to form momentum, our more timely value measure also outperforms the more standard lagged measure.

Exhibit 4 Cross Sectional Regressions: Forecasting B/P Ratios

This exhibit reports Fama-MacBeth regression of B/P ratios on past ratios and an error correction adjustment. The left side is equal to book value per share, divided by price at fiscal year-end. The right side is lagged book value divided by price at fiscal year-end and lagged book value divided by current price as of the previous June: $bp_t^{al} = \gamma_o + \gamma_1 bp_{t-1}^{a,l} + \gamma_2 \left(bp_{t-1}^{a,c} - bp_{t-1}^{a,l} \right) + \in_t$. The rightmost column reports γ2/γ1, the fraction of the linear forecast attributed to $bp^{a,c}$.

	γ^1		γ^2		$\gamma^1 - \gamma^2$			
	Coeff	***t*-stat**	**Coeff**	***t*- stat**	**Coeff**	***t*-stat**	**R2**	γ^2/γ^1
Panel A: U.S. Sample								
All sample	**0.91**	101.3	**0.86**	38.9	**0.05**	3.14	0.73	0.94
Large Cap (above NYE median)	**0.93**	101.5	**0.98**	43.1	**–0.04**	–2.40	0.78	1.05
Small Cap (below NYE median)	**0.88**	110.1	**0.80**	42.4	**0.08**	4.70	0.70	0.91
Panel B: International Sample								
All sample	**0.88**	68.9	**0.75**	37.3	**0.12**	5.92	0.67	0.86
Large Cap (above 80th percentile)	**0.91**	66.4	**0.87**	33.3	0.04	1.66	0.72	0.96
Small Cap (below 80th percentile)	**0.86**	61.1	**0.73**	31.8	**0.14**	6.14	0.65	0.84

Columns 3 and 4 show results for our monthly updated measure. Again we compete with the standard measure (annual lagged), but unlike in columns 1 and 2, in columns 3 and 4 we now also compete with our monthly updated measure. In column 33 we see a strong positive loading on UMD and a negative intercept of –161 bps a year, with a t-statistic of 3.92. In other words, the standard measure economically and statistically subtracts return, given five-factor exposure and including our monthly value measurement. The results are more dramatic in the other direction. Regressing our timely $HML^{m,l}$ on MKT, SMB, UMD, and standard $HML^{a,l}$ we see a large negative loading on UMD, as our timely measure is far more negatively correlated with momentum than is the standard measure, and a very significant intercept of 305 bps a year with a +5.92 *t*-statistic. Essentially, in the presence of MKT, SMB, standard value, and UMD, our most timely value measure is clearly superior.

The international and global results are consistent with the U.S., in particular. Using a monthly updated version of value with current prices adds between 305 and 378 bps of alpha, even after controlling for other value measures. The only exception are the international portfolio results with annual updated measure (column 6), where we are unable to reject the null hypothesis of no value added. (Column 8 still shows very significant results, and our largest intercept in basis points, for the international sample when the monthly HML is employed.)

Exhibit 2 summarizes the results, plotting the cumulative alphas from Exhibit 6, columns 10, 11, and 12. Cumulative alphas are the monthly alpha, plus the error term from the regression. The exhibit shows the large advantage from updating price when constructing

Exhibit 5 Cross Sectional Regressions: Forecasting Changes in Book per Share

This exhibit reports Fama-MacBeth regression of changes in the log of book price per share on log returns over the prior three years. $\Delta b^*_{t-12\to t} = \theta_0 + \theta_1 r_{t-12\to t} + \theta_2 r_{t-24\to t-12} + \theta_3 r_{t-36\to t-12} + \in_t$ The left side is equal changes in book value per share. Lowercase indicates logs = log(B), the asterisk * indicates that the quantity is adjusted for splits between the two dates, and $r_{t\to s} = \log\left(1 + R_{t\to s}\right)$ is equal to the total log return between date t and $s > t$. The lags are in months. Cross sectional regressions are run every fiscal year.

	Coefficient			*t*-statistics			
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	R^2
Panel A: U.S. Sample							
All sample	**0.22**	**0.15**	**0.08**	17.42	19.38	15.24	0.15
Large Cap (above NYE median)	**0.17**	**0.13**	**0.08**	14.13	14.93	11.46	0.14
Small Cap (below NYE median)	**0.23**	**0.15**	**0.07**	18.89	17.81	12.77	0.16
Panel B: International Sample							
All sample	**0.26**	**0.15**	**0.09**	13.25	9.29	3.79	0.11
Large Cap (above 80th percentile)	**0.17**	**0.12**	**0.10**	8.61	6.34	7.40	0.11
Small Cap (below 80th percentile)	**0.28**	**0.16**	**0.09**	**12.91**	9.29	3.18	0.12

value portfolios and combining with other known factors. The exhibit also shows that, though gains to our new factors were small (but the right sign) before 1970, they have been steady and not period specific after 1970.[19]

The appendix reports a battery of robustness checks. We run time-series regression of each value measure on the full set of factors, including the other value measure. For each sample (U.S., international, and global), we report results separately for firms with fiscal years ending and not ending in December. We split the sample into large and small firms, based on the NYSE median market cap for the U.S. sample or the top 80th percentile by country for the international sample, and we report results for different time periods. The robustness checks are consistent with our main results: Value portfolios constructed using more current prices earn higher abnormal returns, even after controlling for the other lagged standard value measure, on average between 121 and 378 bps of alpha.

We asked ourselves why a value portfolio based on more current prices does so much better when combined with momentum and other factors. The short answer is that failing to update prices when computing B/P ratios is not only an inferior measure of true unobservable B/P, but is also an inefficient way to load momentum into a portfolio (or, for stand-alone value, to load less negatively). If price has fallen sharply in the last six months, it is natural and empirically clear from our earlier results that the stock usually has also cheapened, or gotten more attractive on value measures. Also, if the price has fallen sharply in the last six months, then monthly momentum has almost always gotten worse.

Exhibit 6 HML: Time Series Regressions, 1950–2011

This table reports portfolio returns and multivariate loadings. We run time-series regressions on monthly excess returns of value portfolios (HML) and on monthly excess returns of a set of explanatory portfolios. This table includes all available stocks in our U.S. and international samples. The sample period runs from 1950 to 2011. Country portfolios are aggregated into international and global portfolios using the country's total market capitalization as of the prior month. Alpha is the intercept in a regression of monthly excess return. Alphas are annualizedics, *t*-statistics are reported below the coefficient estimates, and a five percent statistical significance is indicated in bold.

	U.S.				International				Global			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Refreshing frequency	Annual	Annual	Annual	Monthly	Annual	Annual	Annual	Monthly	Annual	Annual	Annual	Monthly
Method to lag price	Lagged	Current	Lagged	Current	Lagged	Current	Lagged	Current	Lagged	Current	Lagged	Current
Alpha	–0.58	**1.43**	**–1.61**	**3.05**	1.05	0.56	–1.24	**3.78**	–0.43	**1.51**	**–1.72**	**3.36**
	(–1.35)	(3.42)	(–2.92)	(5.92)	(1.52)	(0.71)	(–1.59)	(4.58)	(–1.07)	(3.69)	(–3.48)	(7.14)
MKT	0.01	**–0.03**	**–0.02**	–0.01	–0.02	0.00	**–0.03**	0.01	0.00	**–0.03**	–0.02	–0.01
	(0.93)	(–3.38)	(–2.09)	(–0.82)	(–1.38)	(–0.30)	(–2.36)	(0.63)	(0.26)	(–3.20)	(–1.87)	(–1.53)
SMB	**–0.04**	0.02	**–0.04**	0.01	–0.02	0.03	–0.01	0.03	**–0.03**	0.02	–0.02	0.00
	(–3.32)	(1.78)	(–2.50)	(0.46)	(–0.67)	(1.22)	(–0.33)	(0.91)	(–2.48)	(1.77)	(–1.28)	(0.31)
STR	–0.01	0.02	**–0.07**	**0.08**	–0.02	0.01	**–0.07**	**0.07**	–0.02	0.03	**–0.08**	**0.10**
	(–1.13)	(1.85)	(–4.19)	(5.58)	(–0.88)	(0.61)	(–3.22)	(3.28)	(–1.25)	(1.92)	(–5.13)	(6.52)
UMD	**0.17**	**–0.19**	**0.38**	**–0.43**	**0.14**	**–0.18**	**0.35**	**–0.45**	**0.17**	**–0.20**	**0.37**	**–0.42**
	(17.24)	(–21.46)	(26.12)	(–39.28)	(8.68)	(–10.68)	(17.33)	(–25.38)	(16.42)	(–20.37)	(25.99)	(–38.00)
HML$^{\text{annual,lagged}}$		**0.92**		**0.85**		**1.02**		**0.95**		**0.94**		**0.87**
		(70.41)		(53.14)		(35.17)		(31.50)		(65.58)		(52.83)
HML$^{\text{annual,current}}$	**0.95**				**0.78**				**0.91**			
	(70.41)				(35.17)				(65.58)			
HML$^{\text{monthly,current}}$			**0.94**				**0.80**				**0.91**	
			(53.14)				(31.50)				(52.83)	
R^2	0.89	0.90	0.82	0.89	0.80	0.81	0.76	0.85	0.87	0.89	0.82	0.89

In other words, skipping six months, as done in the standard $HML^{a,l}$, reduces the natural negative correlation of value and momentum. On the other hand, as of June 30, our more timely value measure $HML^{a,c}$ fully accounts for the negative correlation with momentum, including the impact of the prior six months.

We originally lagged the standard value factor $HML^{a,l}$ to make sure book was available, with price lagged more matter-of-factly to match book. Correlation or overlap with UMD was not a decision factor at that point, as the research on momentum was still in the future. We argue the lag in price was unjustified on first principles (again, when price falls, book does not fall as much, and our best guess is the stock has cheapened), without considering momentum. But if momentum were never discovered, the choice would have been fairly innocuous. As it is, the choice is anything but innocuous.

Effectively, the standard $HML^{a,l}$ looks like a portfolio of the more timely HML, plus UMD, plus noise. In fact, running this regression directly, i.e., without the other factors, results in:

$$\begin{aligned} &HML^{a,l} \\ &\quad = -0.77 + 0.96 \times HML^{a,c} + 0.18 \times UMD \quad R2 = 89\% \\ &\qquad (-1.86) \qquad (74.64) \qquad\qquad (19.18) \\ &HML^{a,l} \\ &\quad = -2.51 + 0.95 \times HML^{m,c} + 0.42 \times UMD \quad R2 = 89\% \\ &\qquad (-4.59) \qquad (55.20) \qquad\qquad (29.30) \end{aligned}$$

Effectively, the regression loads very positively on the highly correlated (but more timely) $HML^{a,c}$ and also very positively on UMD. If we accept, as our earlier evidence showed, that $HML^{a,c}$ is a better, purer proxy for true value, than we can view the standard $HML^{a,l}$ as a portfolio of more accurate value, momentum, and noise. Furthermore, examining the intercept shows a somewhat inferior portfolio—a little or a lot, depending on the annual or monthly method employed. We can only suppose this comes from the fact that ignoring six to 18 months of return is not the best way to account for momentum. It is not the same as the clean addition of a momentum factor, but rather a noisy proxy for it.

Although the general intuition is useful, and we have already shown that the result is strong and robust over time and geography, it is still useful to examine some specific examples.

Example 1: The 2009 Momentum Crash

After being battered by the financial crisis, markets sharply reversed in March 2009 and the momentum strategy suffered greatly. The three-month additive spread return on UMD from March to May, a 14 percent annual volatility series since 1950, was –56 percent. Although this was very painful, it did little to change the momentum strategy's record of long-term efficacy. Still, it's instructive to look at how much of that pain actually had to be borne by a value-plus-momentum investor.

Looking at the standard value portfolio $HML^{a,l}$, our more timely but still annual $HML^{a,c}$, and our very timely $HML^{m,c}$, we see spread returns of +2 percent, +7 percent, and +34 percent, respectively, over the same three months. The standard value portfolio didn't help at all, while our $HML^{m,c}$ offset much of the momentum pain (if indeed one were balanced 50/50 between value and momentum).

This can be seen in Exhibit 3, where we plot total returns for the different HML measures and UMD for our global sample. This is not an accident. March to May 2009 saw a momentum debacle, as momentum tends to severely underperform when the world reverses its actions of the last year—see Daniel [2011]—and this reversal was epic in size. A negatively correlated factor, such as value, could be there at such times to offset such a crash. The standard method could not. But with the simple and intuitive act of updating price in a timely way, our $HML^{m,c}$, value was there to save the day.

Example 2: Value and Momentum in Japan

Japan is a particularly constructive place to examine, as it is widely known as a country where the momentum strategy has failed (Asness [2011]).

Consider Exhibit 7. In columns 1 and 2, we see Japanese UMD adjusted for the market model and the traditional four-factor model, using the standard lagged and annual definition of HML, augmented with the short-term reversal factor. The result is 23 years of economically small and statistically insignificant alpha.

Many observe that momentum has failed in Japan, and they are correct when we view momentum through the standard lens. Furthermore, in column 2 we see that momentum is only marginally correlated with standard HML (and the wrong sign!). That is not intuitive. Recall that one of the problems with standard HML is that it radically reduces the natural negative correlation of a true value and true momentum strategy.

Column 3 replaces standard HML with our annual but unlagged HML, and column 4 replaces standard HML with our monthly unlagged HML. We focus on column 4, as the story it tells is stronger and clearer, but column 3 shows an attenuated version of the same effect.

In column 4, we see an economically and statistically large intercept for UMD in Japan, driven by a economically and statistically large, negative coefficient on monthly unlagged HML. When we adjust for the very strong negative correlation of UMD with monthly unlagged HML, we see tremendous value added, even in Japan.

This is all quite intuitive. In Japan, from 1988 to 2011, univariate value was quite strong, and univariate momentum was a complete dud (around zero univariate return). When we use the standard measure of HML, which downplays the negative correlation of momentum and true value, UMD remains a dud. But, when we use monthly unlagged value, itself a very strong strategy in Japan over this period, UMD is resurrected. Being very negatively correlated with a strong strategy, such as monthly HML in Japan, but not losing, is indeed value added, as risk can be reduced at a low cost to expected return. This reality in Japan is masked by the standard measure of value, but shown remarkably clearly by our much more timely measure of true value: $HML^{m,c}$.

Exhibit 7 Case Study: Momentum in Japan, 1988–2011

This exhibit reports portfolio returns and multivariate loadings. We run time-series regressions on monthly excess returns of momentum portfolios (UMD) and on monthly excess returns of a set of explanatory portfolios. This exhibit includes all available stocks in our Japanese sample. The sample period runs from 1988 to 2011. Alpha is the intercept in a regression of monthly excess return. The left side is momentum (UMD) returns. The explanatory variables are market excess returns (MKT), a size portfolio (SMB), a value portfolio (HML) and a short-term reversal (STR) portfolio. Alphas are annualized, *t*-statistics are reported below the coefficient estimates, and a five percent statistical significance is indicated in bold.

Panel A: UMD, Japan, 1988–2011

	(1)	(2)	(3)	(4)
Alpha	1.15	2.26	**7.34**	**12.04**
	(0.34)	(0.68)	(2.37)	(4.49)
MKT	**–0.23**	**–0.16**	**–0.27**	**–0.24**
	(–5.34)	(–3.38)	(–6.41)	(–7.13)
SMB		**–0.19**	–0.05	0.03
		(–2.54)	(–0.63)	(0.42)
STR		**–0.32**	**–0.31**	**–0.18**
		(–4.73)	(–4.95)	(–3.34)
$HML^{annual,lagged}$		0.14		
		(1.21)		
$HML^{annual,current}$			**–0.58**	
			(–6.60)	
$HML^{monthly,current}$				**–0.80**
				(–12.66)
R2	0.09	0.19	0.30	0.49

Conclusion

The standard approach to calculating HML, itself the standard value strategy, updates portfolios once a year, using prices lagged six months from the update. Thus, by the next update, the price used to determine value is 18 months old.

We show on first principles that, if the goal is approximating the true, unobservable B/P, a technique that uses an unlagged price comes much closer. We recommend a change to the standard approach, based only on this idea and before examining returns. We show that, in the context of a five-factor model that includes momentum, this logically superior value measure is actually far superior in terms of returns. We further extend this to a monthly updated value strategy and find that, for precisely analogous reasons, the return advantage grows far stronger.

The bottom line is that the standard approach to value was a reasonable, conservative choice that has served the field well. But it is not the best possible choice. Moving in very simple ways, based on first principles, to the choices we study here can make a big difference in the efficacy of combined portfolio strategies, helping us set a higher bar by using value and momentum for risk-adjustment and performance attribution.

Appendix

Additional Empirical Results And Robustness Tests

This appendix contains additional empirical results and robustness tests.

- Exhibit A1 reports summary statistics.
- Exhibit A2 reports results of Fama-MacBeth regression of book-to-price ratios on past ratios.
- Exhibit A3 reports results of Fama-MacBeth regression of changes in log of book price per share on log returns over the prior three years.
- Exhibit A4 reports returns of HML portfolios.
- Exhibit A5 reports five-factor alphas of HML portfolios across different subsamples.
- Exhibit A6 reports *t*-statistics of five-factor alphas of HML portfolios by country.

Exhibit A1 Summary Statistics

This exhibit shows summary statistics as of June of each year. The sample includes all common stocks on the CRSP/XpressFeed data between 1950 and 2011 and all common stocks on the Xpress-Feed Global data between 1983 and 2011. "Number of stocks—mean" is the average number of stocks per year. "Mean ME" is the average firm's market value of equity, in billion USD. Means are pooled averages (firm-year) as of June of each year.

Country	Number of Stocks—Total	Number of Stocks—Mean	Mean ME (firm, Billion USD)	Average Weight in International Portfolio	Average Weight in Global Portfolio	Start Year	End Year
Australia	2,951	808	0.56	0.031	0.018	1989	2011
Austria	207	76	0.72	0.004	0.002	1990	2011
Belgium	421	132	1.88	0.017	0.010	1989	2011
Canada	5,560	709	0.71	0.039	0.023	1983	2011
Switzerland	541	200	2.90	0.043	0.025	1989	2011
Germany	2,048	662	2.43	0.109	0.065	1989	2011
Denmark	412	137	0.79	0.008	0.005	1989	2011
Spain	415	141	2.66	0.026	0.016	1991	2011
Finland	288	105	1.38	0.010	0.006	1989	2011
France	1,765	555	2.09	0.084	0.049	1989	2011
United Kingdom	6,006	1,811	1.19	0.167	0.099	1988	2011
Hong Kong	1,670	602	1.13	0.046	0.027	1989	2011
Italy	600	219	2.09	0.034	0.020	1990	2011
Japan	4,952	2,847	1.21	0.291	0.172	1987	2011
Netherlands	407	167	3.25	0.040	0.024	1989	2011
Norway	648	148	0.74	0.007	0.004	1989	2011
New Zealand	312	94	0.79	0.005	0.003	1993	2011
Singapore	1,026	342	0.60	0.016	0.009	1989	2011
Sweden	997	243	1.28	0.022	0.013	1989	2011
United States	23,320	3,087	0.95		0.409	1950	2011

Exhibit A2 Cross Sectional Regressions: Forecasting Book-to-Price Ratios

This exhibit reports Fama-MacBeth regression of book-to-price ratios on past ratios and an error correction adjustment. The left-hand side is equal to book value per share divided by price at fiscal year-end. The right-hand side is lagged book value divided by price at fiscal year-end and lagged book value divided by current price as of the previous June:

$$bp_t^{a,l} = \gamma_0 + \gamma_1 bp_{t-1}^{a,l} + \gamma_2 \left(bp_{t-1}^{a,c} - bp_{t-1}^{a,l} \right) + \epsilon_t$$

The first superscript indicates the refreshing frequency (annual *a* or monthly *m*), the second superscript indicates the lag used to update price (lagged *l* or current *c*). The right-hand side variables are windsorized at 1% level and cross sectional regressions are run every fiscal year. The rightmost column reports γ_2/γ_1, the fraction of the linear forecast attributed to *bpa,c*. Panel A reports results for our U.S. sample. "All sample" reports results for the full sample. "(Non) December FYE" report results for firms with fiscal year (not) ending in December. "Industry Fixed Effect" reports results for regression including industry fixed effects based on 49-industry classification from Ken French's website. "ME-1" to "ME-10" reports results for each NYSE-based size percentiles. The last rows reports results by sample period. The sample period for the U.S. sample runs from 1950 to 2011. Panel B reports results for our International sample. "All sample" reports results for the full sample. "Large (Small) Cap" report results for firms above (below) the 80th percentiles (by country). The remaining rows report results by sample period and by country. The sample period for the International sample runs from 1983 to 2011. *t*-statistics are reported next to the coefficient estimates and five percent statistical significance is indicated in bold.

Panel A: U.S. results

	γ_1		γ_2		$\gamma_1 - \gamma_2$			
	Coeff	***t*-stat**	**Coeff**	***t*-stat**	**Coeff**	***t*-stat**	**R2**	γ_2/γ_1
All sample	**0.91**	101.3	**0.86**	38.9	**0.05**	3.14	0.73	0.94
December FYE only	**0.91**	86.2	**0.89**	35.7	0.02	1.14	0.72	0.98
Non-December FYE only	**0.92**	115.6	**0.83**	41.2	**0.09**	5.97	0.76	0.90
Industry Fixed Effects	**0.90**	97.9	**0.84**	41.3	**0.06**	3.87	0.69	0.93
ME-1 (small)	**0.89**	92.1	**0.79**	29.4	**0.10**	3.86	0.67	0.89
ME-2	**0.86**	78.2	**0.81**	34.8	**0.05**	2.48	0.71	0.94
ME-3	**0.87**	74.2	**0.85**	36.3	0.02	0.94	0.72	0.98
ME-4	**0.88**	75.2	**0.90**	29.6	–0.02	–0.64	0.74	1.02
ME-5	**0.90**	69.6	**0.91**	28.3	–0.01	–0.23	0.75	1.01
ME-6	**0.92**	69.5	**0.97**	33.6	**–0.05**	–1.99	0.77	1.05
ME-7	**0.91**	85.5	**0.97**	32.8	**–0.06**	–1.97	0.77	1.06
ME-8	**0.94**	75.0	**1.03**	33.8	**–0.09**	–3.16	0.76	1.10
ME-9	**0.94**	93.6	**1.03**	31.2	**–0.08**	–2.86	0.79	1.09
ME-10 (large)	**0.95**	86.4	**1.06**	31.4	**–0.11**	–3.63	0.80	1.12
1950–1970	**0.98**	87.1	**1.06**	45.6	**–0.09**	–4.08	0.81	1.09
1971–1990	**0.91**	78.7	**0.78**	38.8	**0.13**	7.99	0.75	0.86
1991–2000	**0.88**	61.5	**0.75**	37.8	**0.13**	7.85	0.64	0.86
2001–2011	**0.84**	47.2	**0.74**	21.0	**0.11**	2.48	0.65	0.87

(*Continued*)

Exhibit A2 (*Continued*)

Panel B: International Results

	γ_1		γ_2		$\gamma_1 - \gamma_2$			
	Coeff	***t*-stat**	**Coeff**	***t*-stat**	**Coeff**	***t*-stat**	**R2**	γ_2/γ_1
All sample	**0.88**	68.9	**0.75**	37.3	**0.12**	5.92	0.67	0.86
Large Cap	**0.90**	66.4	**0.87**	33.3	0.04	1.66	0.72	0.96
Small Cap	**0.86**	61.1	**0.73**	31.8	**0.14**	6.14	0.65	0.84
1983–1990	**0.82**	26.8	**0.79**	14.1	0.03	0.66	0.64	0.96
1991–2000	**0.93**	73.8	**0.78**	31.4	**0.15**	6.58	0.73	0.84
2001–2011	**0.87**	80.8	**0.71**	25.3	**0.16**	5.52	0.65	0.82
Australia	**0.82**	61.7	**0.69**	31.4	**0.13**	8.30	0.64	0.84
Austria	**0.83**	17.8	**0.75**	8.2	0.08	1.14	0.59	0.90
Belgium	**0.91**	2.0	**0.89**	3.2	0.02	0.06	0.68	0.98
Canada	**0.82**	65.1	**0.71**	31.3	**0.12**	4.68	0.59	0.86
Switzerland	**0.96**	43.5	**0.93**	20.5	0.03	0.66	0.67	0.97
Germany	**0.85**	39.6	**0.80**	21.1	0.04	1.08	0.61	0.95
Denmark	**0.98**	25.0	**1.02**	12.2	–0.04	–0.52	0.62	1.04
Spain	**0.84**	33.7	**0.92**	11.1	–0.08	–0.85	0.70	1.10
Finland	**0.93**	21.3	**0.81**	11.5	0.12	1.76	0.67	0.87
France	**0.88**	44.2	**0.87**	20.4	0.01	0.30	0.67	0.99
U.K.	**0.88**	54.9	**0.74**	17.7	**0.14**	4.16	0.66	0.84
Honk Hong	**0.92**	64.5	**0.80**	19.8	**0.12**	2.46	0.69	0.87
Italy	**0.90**	44.9	**0.73**	10.0	**0.17**	2.44	0.69	0.81
Japan	**0.95**	108.9	**0.72**	38.2	**0.23**	11.19	0.82	0.75
Netherlands	**0.92**	27.6	**0.95**	10.7	–0.03	–0.42	0.66	1.03
Norway	**0.83**	23.1	**0.69**	8.5	0.14	1.90	0.49	0.84
New Zealand	**0.91**	23.9	**0.70**	12.2	**0.21**	3.64	0.73	0.77
Singapore	**0.88**	47.4	**0.84**	19.5	0.04	0.94	0.68	0.96
Sweden	**0.91**	24.4	**0.91**	15.0	0.00	0.01	0.58	1.00

Exhibit A3 Cross-Sectional Regressions: Forecasting Changes in Book per Share

This exhibit reports Fama-MacBeth regression of changes in log book per share on log returns over the prior three years.

$$\Delta b^{*}_{t-12\to t} = \theta_0 + \theta_1 r_{t-12\to t} + \theta_2 r_{t-24\to t-12} + \theta_3 r_{t-36\to t-12} + \in_t$$

The left-hand side is equal changes in book value per share where lowercase indicated logs = log (B), the asterisk * indicates that the quantity is adjusted for splits between the two dates and $r_{t\to s} = \log\left(1 + R_{t\to s}\right)$ is equal to the total log return between date t and $s > t$. The lags are in months. Cross sectional regressions are run every fiscal year. Panel A reports results for our U.S. sample. "All sample" reports results for the full sample. "(Non) December FYE" report results for firms with fiscal year (not) ending in December. "Industry Fixed Effect" reports results for regression including industry fixed effects based on 49-industry classification from Ken French's website. "ME-1" to "ME-10" reports results for each NYSE-based size percentiles. The last rows reports results by sample period. The sample period for the U.S. sample runs from 1950 to 2011. Panel B reports results for our International sample. "All sample" reports results for the full sample. "Large (Small) Cap" report results for firms above (below) the 80th percentiles (by country). The remaining rows report results by sample period and by country. The sample period for the International sample runs from 1983 to 2011. t-statistics are reported next to the coefficient estimates and five percent statistical significance is indicated in bold.

Panel A: U.S. Results

	Coefficient			t-statistics			
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	R2
All sample	**0.22**	**0.15**	**0.08**	17.42	19.38	15.24	0.15
December FYE only	**0.22**	**0.15**	**0.08**	16.04	16.92	12.60	0.14
Non-December FYE only	**0.22**	**0.16**	**0.08**	20.21	19.10	13.29	0.19
Industry Fixed Effects	**0.20**	**0.15**	**0.05**	12.29	12.19	3.99	0.38
ME-1 (small)	**0.24**	**0.15**	**0.08**	19.86	14.82	8.54	0.19
ME-2	**0.21**	**0.14**	**0.05**	15.06	10.50	4.47	0.16
ME-3	**0.18**	**0.16**	**0.07**	15.22	12.63	7.93	0.17
ME-4	**0.16**	**0.14**	**0.07**	13.71	12.36	6.55	0.16
ME-5	**0.20**	**0.13**	**0.06**	12.01	10.64	5.80	0.18
ME-6	**0.18**	**0.13**	**0.09**	12.65	9.48	8.39	0.20
ME-7	**0.17**	**0.14**	**0.08**	9.43	8.24	5.82	0.20
ME-8	**0.17**	**0.13**	**0.07**	10.43	9.39	3.87	0.16
ME-9	**0.12**	**0.13**	**0.10**	6.89	7.42	7.16	0.13
ME-10 (large)	**0.13**	**0.13**	**0.10**	8.23	11.08	6.75	0.13
1950–1970	**0.11**	**0.09**	**0.05**	10.39	11.72	6.49	0.09
1971–1990	**0.22**	**0.16**	**0.08**	14.26	14.41	12.91	0.15
1991–2000	**0.30**	**0.20**	**0.10**	40.80	28.04	9.63	0.17
2001–2011	**0.31**	**0.22**	**0.11**	28.04	24.73	9.06	0.22

(*Continued*)

Exhibit A3 (*Continued*)

Panel B: International Results

	Coefficient			*t*-statistics			R2
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	
All sample	**0.26**	**0.15**	**0.09**	13.25	9.29	3.79	0.11
Large Cap	**0.17**	**0.12**	**0.10**	8.61	6.34	7.40	0.11
Small Cap	**0.28**	**0.16**	**0.09**	**12.91**	9.29	3.18	0.12
1983-1990	**0.29**	**0.17**	0.10	6.86	5.83	1.75	0.13
1991-2000	**0.24**	**0.13**	**0.08**	**14.13**	8.07	6.26	0.10
2001-2011							
Australia	**0.31**	**0.18**	**0.12**	**9.67**	11.66	5.15	0.23
Austria	**0.20**	**0.35**	0.10	2.63	3.18	0.90	0.06
Belgium	**0.32**	–0.02	**0.25**	**4.38**	–0.24	2.49	0.09
Canada	**0.33**	**0.19**	**0.09**	**15.25**	10.85	4.82	0.21
Switzerland	**0.41**	0.06	**0.08**	**5.07**	0.79	2.66	0.13
Germany	**0.31**	**0.17**	**0.09**	**14.00**	7.35	3.50	0.11
Denmark	**0.25**	**0.18**	0.02	6.16	4.75	0.58	0.10
Spain	0.11	**0.18**	**0.28**	**1.56**	3.25	3.83	0.17
Finland	**0.22**	0.05	**0.11**	**2.05**	0.84	2.45	0.21
France	0.21	0.16	0.05	6.74	5.63	1.27	0.09
U.K.	**0.19**	**0.18**	**0.10**	**3.53**	5.88	2.40	0.12
Honk Hong	**0.16**	**0.15**	**0.08**	**2.65**	6.98	3.62	0.13
Italy	**0.42**	0.12	**0.30**	**4.53**	1.01	2.18	0.15
Japan	**0.16**	**0.12**	**0.07**	**12.88**	12.30	6.81	0.08
Netherlands	**0.23**	**0.21**	**0.11**	**6.02**	4.71	2.36	0.14
Norway	**0.37**	**0.27**	0.03	6.39	7.85	0.73	0.20
New Zealand	**0.32**	**0.18**	–0.01	6.13	4.71	–0.14	0.19
Singapore	**0.17**	**0.15**	**0.04**	**9.67**	6.46	2.87	0.09
Sweden	**0.30**	**0.20**	**0.12**	**7.62**	6.90	5.34	0.29

This exhibit reports Fama-MacBeth regression of changes in log book per share on log returns over the prior three years. Cross sectional regressions are run every fiscal year. Panel A reports results for our U.S. sample. "All sample" reports results for the full sample. "(Non) December FYE" report results for firms with fiscal year (not) ending in December. "Industry Fixed Effect" reports results for regression including industry fixed effects based on 49-industry classification from Ken French's website. "ME-1" to "ME-10" reports results for each NYSE-based size percentiles. The last rows reports results by sample period. The sample period for the U.S. sample runs from 1950 to 2011. Panel B reports results for our International sample. "All sample" reports results for the full sample. "Large (Small) Cap" report results for firms above (below) the 80th percentiles (by country). The remaining rows report results by sample period and by country. The sample period for the International sample runs from 1983 to 2011. *t*-statistics are reported next to the coefficient estimates and five percent statistical significance is indicated in bold.

Exhibit A4 HML: Univariate Results, 1950–2011

This exhibit reports returns of value portfolios (HML). The value factors are constructed using three book-to-price (B/P) measures: The first measure is equal to book value per share divided by price at fiscal year-end both in local currency. We denote this value portfolio as $HML^{annual,lagged}$. The second measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price. We denote this value portfolio as $HML^{annual,current}$ Both annual measures are refreshed in June. The third measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price, updated monthly. We denote this value portfolio as $HML^{monthly,current}$ We construct portfolios within each country in our sample. At the end of June of year *t* (at the end of each calendar month for the monthly measure), stocks are assigned to two size-sorted portfolios based on their market capitalization. The size breakpoint for the U.S. sample is the median NYSE market equity. The size breakpoint for the international sample is the 80th percentile by country. Portfolios are value-weighted, refreshed every June (refreshed every month for the monthly measure), and rebalanced every calendar month to maintain value weights. The value factor HML is the average return on the two value portfolios minus the average return on the two growth portfolios. This exhibit includes all available stocks in our U.S. and International sample. The sample period runs from 1950 to 2011. Country portfolios are aggregated into Global portfolios using the country's total market capitalization as of the prior month. Returns, volatilities and Sharpe ratios are annualized, *t*-statistics are reported below the coefficient estimates and five percent statistical significance is indicated in bold.

	U.S.				International				Global			
Measure	**HML**	**HML**	**HML**	**UMD**	**HML**	**HML**	**HML**	**UMD**	**HML**	**HML**	**HML**	**UMD**
Refreshing frequency	**Annual**	**Annual**	**Monthly**	**Monthly**	**Annual**	**Annual**	**Monthly**	**Monthly**	**Annual**	**Annual**	**Monthly**	**Monthly**
Method to lag price	**Lagged**	**Current**	**Current**	**Current**	**Lagged**	**Current**	**Current**	**Current**	**Lagged**	**Current**	**Current**	**Current**
Full Sample												
Mean	**4.04**	**3.44**	**3.10**	**8.70**	**6.65**	**6.07**	**6.99**	**6.97**	**4.92**	**4.42**	**4.32**	**8.71**
t-statistics	3.35	2.70	2.09	4.83	4.53	3.50	3.48	2.65	4.76	3.94	3.36	5.49
Volatility	9.32	9.84	11.44	13.91	7.66	9.05	10.47	13.72	7.99	8.68	9.93	12.26
SR	0.43	0.35	0.27	0.63	0.87	0.67	0.67	0.51	0.62	0.51	0.44	0.71
Corr with UMD	–0.13	–0.39	–0.64	1.00	0.01	–0.27	–0.61	1.00	–0.09	–0.36	–0.62	1.00

(*Continued*)

Exhibit A4 (*Continued*)

	U.S.				International				Global			
Measure	**HML**	**HML**	**HML**	**UMD**	**HML**	**HML**	**HML**	**UMD**	**HML**	**HML**	**HML**	**UMD**
Refreshing frequency	**Annual**	**Annual**	**Monthly**	**Monthly**	**Annual**	**Annual**	**Monthly**	**Monthly**	**Annual**	**Annual**	**Monthly**	**Monthly**
Method to lag price	**Lagged**	**Current**	**Current**	**Current**	**Lagged**	**Current**	**Current**	**Current**	**Lagged**	**Current**	**Current**	**Current**
1951–1989												
Mean	**4.60**	**4.31**	**3.34**	**10.04**	**6.22**	**7.07**	**4.47**	**7.14**	**4.72**	**4.54**	**3.62**	**9.79**
t-statistics	3.62	3.12	2.23	6.05	1.74	1.59	0.92	1.52	3.71	3.28	2.40	5.91
Volatility	7.89	8.57	9.30	10.30	8.75	10.86	11.86	11.52	7.89	8.58	9.37	10.27
SR	0.58	0.50	0.36	0.97	0.71	0.65	0.38	0.62	0.60	0.53	0.39	0.95
Corr with UMD	–0.15	–0.42	–0.56	1.00	–0.17	–0.29	–0.47	1.00	–0.14	–0.40	–0.55	1.00
1990–2011												
Mean	3.04	1.87	2.65	6.27	**6.77**	**5.79**	**7.70**	**6.92**	**5.28**	**4.21**	**5.58**	**6.75**
t-statistics	1.22	0.73	0.84	1.54	4.25	3.14	3.53	2.23	2.98	2.18	2.37	2.05
Volatility	11.49	11.81	14.56	18.76	7.34	8.50	10.06	14.30	8.19	8.88	10.87	15.22
SR	0.26	0.16	0.18	0.33	0.92	0.68	0.77	0.48	0.65	0.47	0.51	0.44
Corr with UMD	–0.12	–0.37	–0.70	1.00	0.06	–0.27	–0.66	1.00	–0.04	–0.33	–0.70	1.00

Exhibit A5 Robustness Checks: 5-Factor Alphas

This exhibit reports portfolio returns. We run time series regressions on monthly excess returns of value portfolios (HML) on monthly excess returns on a set of explanatory portfolios. The value factors are constructed using three book-to-price (B/P) measures: The first measure is equal to book value per share divided by price at fiscal year-end both in local currency. We denote this value portfolio as $HML^{annual,lagged}$. The second measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price. We denote this value portfolio as $HML^{annual,current}$. Both annual measures are refreshed in June. The third measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price, updated monthly. We denote this value portfolio as $HML^{monthly,current}$. We construct portfolios within each country in our sample. At the end of June of year t (at the end of each calendar month for the monthly measure), stocks are assigned to two size-sorted portfolios based on their market capitalization. The size breakpoint for the U.S. sample is the median NYSE market equity. The size breakpoint for the international sample is the 80th percentile by country. Portfolios are value-weighted, refreshed every June (refreshed every month for the monthly measure), and rebalanced every calendar month to maintain value weights. The value factor HML is the average return on the two value portfolios minus the average return on the two growth portfolios. This exhibit includes all available stocks in our U.S. and International sample. The sample period runs from 1950 to 2011. Country portfolios are aggregated into International and Global portfolios using the country's total market capitalization as of the prior month. Alpha is the intercept in a regression of monthly excess return. The explanatory variables are market excess returns (MKT), a size portfolio (SMB), a momentum portfolio (UMD) and, short term reversal (STR) portfolio and the value measure (HML) indicated in the exhibit. Alphas are annualized, *t*-statistics are reported next to the coefficient estimates and five percent statistical significance is indicated in bold.

Panel A: U.S. sample

	Annual Lagged		**Annual Lagged**		**Annual Current**		**Monthly Current**	
Left hand side: HML	**Annual Current**		**Monthly Current**		**Annual Lagged**		**Annual Lagged**	
HML measure on right hand side	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**
All sample	–0.58	–1.35	**–1.61**	–2.92	**1.43**	3.4	**3.05**	5.92
December FYE only	–0.68	–1.44	**–1.91**	–3.07	**1.49**	3.4	**3.48**	6.09
Non-December FYE only	1.30	1.75	0.20	0.25	1.05	1.3	**3.21**	3.49
Large Cap	**–1.54**	–2.94	**–2.73**	–4.14	**2.14**	4.3	**3.79**	6.21
Small Cap	0.81	1.54	0.52	0.71	0.74	1.4	**2.74**	3.82
1950–1970	–0.89	–1.26	0.65	0.70	**1.97**	2.8	1.00	1.14
1971–1990	**–2.22**	–2.90	**–2.83**	–3.67	**3.25**	4.4	**3.86**	5.29
1991–2000	**2.30**	2.10	1.84	1.27	–0.75	–0.7	0.82	0.63
2001–2010	0.06	0.05	–1.70	–1.12	0.41	0.3	2.70	1.77

(*Continued*)

Exhibit A5 (*Continued*)

Left hand side: HML	Annual Lagged		Annual Lagged		Annual Current		Monthly Current	
	Annual Current		Monthly Current		Annual Lagged		Annual Lagged	
HML measure on right hand side	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**	**Alpha**	***t*-stat**
Panel B: International sample								
All sample	1.05	1.52	–1.24	–1.59	0.56	0.71	**3.78**	4.58
December FYE only	0.26	0.31	–1.55	–1.54	**1.78**	2.17	**4.73**	5.04
Non-December FYE only	0.78	0.79	–1.21	–1.05	1.14	1.10	**4.67**	3.88
Large Cap	1.28	1.42	–1.77	–1.89	0.32	0.32	**3.88**	4.04
Small Cap	0.88	1.05	0.35	0.33	1.73	1.89	**4.88**	4.23
1983–1990	1.12	0.62	1.33	0.72	1.56	0.70	**1.44**	0.63
1991–2000	0.23	0.27	**–2.88**	–3.22	0.67	0.78	**3.96**	4.77
2001–2010	1.43	1.49	–1.36	–1.29	0.35	0.30	**4.61**	3.70
Panel C: Global sample								
All sample	–0.43	–1.07	**–1.72**	–3.48	**1.51**	3.69	**3.36**	7.14
December FYE only	0.64	0.75	–0.59	–0.55	1.54	1.81	**4.34**	4.21
Non-December FYE only	1.24	1.20	–0.36	–0.30	1.02	0.92	**4.42**	3.42
Large Cap	**–1.53**	–3.13	**–3.11**	–5.38	**2.39**	4.99	**4.33**	8.03
Small Cap	0.84	1.74	0.50	0.76	0.87	1.74	**2.99**	4.56
1983–1990	–1.57	–1.94	**–2.57**	–3.15	**2.94**	3.66	**4.01**	5.04
1991–2000	1.19	1.65	–1.05	–1.21	–0.29	–0.39	**2.26**	2.90
2001–2010	0.47	0.48	**–2.40**	–2.26	0.52	0.51	**4.16**	3.70

Exhibit A6 Robustness Checks: *t*-statistics of 5-Factor Alpha by Country.

Exhibit 6 report *t*-statistics of abnormal portfolio returns. We run time series regressions on monthly excess returns of value portfolios (HML) on monthly excess returns on a set of explanatory portfolios. The value factors are constructed using three book-to-price (B/P) measures: The first measure is equal to book value per share divided by price at fiscal year-end both in local currency. We denote this value portfolio as $HML^{annual,lagged}$. The second measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price. We denote this value portfolio as $HML^{annual,current}$. Both annual measures are refreshed in June. The third measure is equal to book value per share (adjusted for splits, dividends and other corporate actions between fiscal year-end and portfolio formation dates) divided by current price, updated monthly. We denote this value portfolio as $HML^{monthly,current}$. We construct portfolios within each country in our sample. At the end of June of year (at the end of each calendar month for the monthly measure), stocks are assigned to two size-sorted portfolios based on their market capitalization. The size breakpoint for the U.S. sample is the median NYSE market equity. The size breakpoint for the international sample is the 80th percentile by country. Portfolios are value-weighted, refreshed every June (refreshed every month for the monthly measure), and rebalanced every calendar month to maintain value weights. The value factor HML is the average return on the two value portfolios minus the average return on the two growth portfolios. This exhibit includes all available stocks in our U.S. and International sample. We plot *t*-statistics of five-factor alphas. Alphas. Alpha is the intercept in a regression of monthly excess return. The left-hand sides are return of the $HML^{annual,current}$ factor or $HML^{monthly,current}$. The explanatory variables are market excess returns (MKT), a size portfolio (SMB), a momentum portfolio (UMD) and, short-term reversal (STR) portfolio and the value portfolio $HML^{annual,lagged}$.

Endnotes

We thank Aaron Brown, Kent Daniel, Tobias Moskowitz, Lars Nielsen, Lasse Pedersen, and seminar participants at the 2012 JOIM Conference for useful comments and suggestions. An early draft of this article circulated under the title "Lagging Value, AQR Capital Management Paper, 2011."

[1]From Exhibit 6.

[2]This is HML as used in its now-ubiquitous academic meaning, not as the Internet texting shorthand with a very different meaning. For a brief time centered around the 1999 tech stock bubble, the two meanings became interchangeable.

[3]In the appendix, we show that our more timely value portfolios earn lower raw returns than traditional value portfolios, but have larger four- and five-factor alphas.

[4]We assign individual issues to the corresponding market, based on the primary exchange's location. For international companies with securities traded in multiple markets, we use the primary trading vehicle that XpressFeed identifies.

[5]Throughout the paper, we use lowercase letters to indicate logs: bp=log(BP).

[6]For firms with fiscal years ending in December, this is the same measure as in Fama and French [1992]. For firms with fiscal year not ending in December, we use prices at the fiscal year-end date, while Fama and French [1992] use December prices for all firms, thus introducing a slight mismatch. Our results are unchanged if we adopt Fama and French's [1992] convention, or if we restrict our sample to firms with fiscal year-ends in December.

[7]The adjust factor adjusts for splits and other corporate actions between the fiscal year-end and the current date.

[8]Because some countries have a small cross-section of stocks in the early years of the sample period, for the international sample we use conditional sorts (first sorting on size, then on B/PB/P) to ensure we have enough securities in each portfolio (the U.S. sorts are always independent).

[9]We use the standard annual lagged method to compute SMB. Using either of the alternative methods of computing B/P has a negligible impact on SMB returns and on our main results.

[10]Corresponds to the standard HML factor used in the literature. From Exhibit 7, over our sample returned 4.0 percent a year, with an annualized volatility of 9.3 percent a year. For comparison, over the common sample period, returns of the HML factor from Ken French's data library were 4.5 percent a year, with 9.5 percent volatility. The correlation between the two series was 0.95. The small (and statistically insignificant) discrepancy between the two series is due to our choice of using price at fiscal year-end (as opposed to December price for all firms, as in Fama and French [1992]), and the fact that our portfolio skips one trading day between rebalancing and investment.

[11]We include delisting returns when available in CRSP. Delisting returns are not available for our international sample. If a firm is delisted but the delisting return is missing, we investigate the reason for disappearance. If the delisting is performance related, we follow Shumway [1997] and assume a –30 percent delisting return. This assumption does not affect any of the results.

[12]Skipping a day serves two purposes. First, it ensures that our portfolios are implementable, in that they use only information available at portfolio formation. Second, it avoids mechanic negative autocorrelation in returns induced by bid-ask bounce, which would tend to overstate returns to STR.

[13]We run annual regressions using annual measures, to put both forecasting variables on an equal footing. In practice on December 2000, we also observe Book (December 31, 1999)/Price (November 29, 2000). Regressions using our monthly measure yield even stronger results, but we prefer to report results based on the annual measure, in order to keep a clean comparison between the two alternatives.

[14]Indicates that the quantity is adjusted for splits between the two dates, and that lags are in months.

[15]For brevity, we do not report results for lags of more than three years, as the coefficients tend to be insignificant.

[16]Later we will argue that the standard method avoids too much shorting of the UMD factor, but in a suboptimal manner versus our more timely measures. This issue would remain even if the coefficients here summed to near 100 percent.

[17]As shown in Exhibit 4, this global portfolio is on average 40 percent U.S. and 60 percent international stocks, with less weight in the U.S. in the most recent period. Our international sample is quite short (starting in 1983 for Canada, with the full set of countries not available until the early 1990s). Because we are estimating expected returns, we tend to emphasize U.S. and global results that are based on a longer time series.

[18]We include STR purely for conservatism. Although we lag a day in constructing our portfolios, it is possible that using more timely measures of price introduces an exposure to the known one-month reversal factor, a factor that is more difficult to implement, and more open to microstructure biases, than our other factors. Our results are still very strong, but are very slightly and intuitively weakened by adding this factor. Skipping a day in portfolio construction and including this factor ensures that our results are not driven by exposure to this higher turnover factor. Repeating our tests on the more standard four-factor model would show slightly stronger results, with no changes in conclusion.

[19]For those used to looking at cumulative returns to value and seeing a big dip during the technology bubble of 1999, please note these are not returns to value investing, but returns to one form of value investing versus another form of value investing (and additional risk factors).

References

Asness, C. "Momentum in Japan." *The Journal of Portfolio Management,* 37 (2011), pp. 67–75.

Daniel, K. "Momentum Crashes." Working paper, Columbia University, 2011.

Fama, E., and K. French. "The Cross-Section of Expected Stock Returns." *The Journal of Finance,* 47 (1992), pp. 427–465.

_____. "Common Risk Factors in the Returns on Stocks and *Bonds." Journal of Financial Economics*, 33 (1993), pp. 3–56.

_____. "Multifactor Explanations of Asset Pricing Anomalies." *The Journal of Finance,* 51 (1996), pp. 55–84.

Fama, E., and J. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy,* 81 (1973), pp. 607–636.

Shumway, T. "The Delisting Bias in CRSP Data." *The Journal of Finance,* 52 (1997), pp. 327–340.

This page intentionally left blank

Fact, Fiction, and Momentum Investing

Clifford Asness, Andrea Frazzini, Ronen Israel, and Tobias Moskowitz

Momentum is the phenomenon that securities that have performed well relative to peers (winners) on average continue to outperform, and securities that have performed relatively poorly (losers) tend to continue to underperform.[1] The existence of momentum is a well-established empirical fact. The return premium is evident in 212 years (yes, this is not a typo, *two hundred and twelve years* of data from 1801 to 2012) of U.S. equity data,[2] dating back to the Victorian age in U.K. equity data,[3] in more than 20 years of out-of-sample evidence from its original discovery, in 40 other countries, and in more than a dozen other asset classes.[4] Some of this evidence predates academic research in financial economics, suggesting that the momentum premium has been a part of markets since their very existence, well before researchers studied them as a science.

The growth in popularity of momentum strategies has, not surprisingly, corresponded to an expanding body of research.[5] At the same time, myths around momentum have also proliferated. Some of the most common myths are that momentum is too small and sporadic a factor, works mostly on the shortside, works well only among small stocks, and doesn't survive trading costs. Furthermore, some argue that momentum is best used as a screen, not as a regular factor in an investment process. Others will go so far as to say that momentum investing is like a game of hot potato, implying that it isn't a serious investment strategy, with no theory or reasonable explanation to back it up.

Frankly, we're a little irked (if that was not clear) by those who should know better but continue to repeat these myths, stretching the limits of credulity. In this article, we address and refute these myths using academic papers (that have been widely circulated throughout the academic and practitioner communities, have been presented and debated at top-level academic seminars and conferences, and have been published in peer-reviewed journals) and the simplest data taken from Kenneth French's publicly available website, a standard

Reprinted from *The Journal of Portfolio Management's* 40th Anniversary Issue, 2014, Vol. 40 (5) 75–92 with permission.

data set used by both academics and practitioners. Anyone repeating these myths, in any dimension, after reading this piece is simply ignoring the facts.

We make no claim that momentum works all the time. In fact, of late (the last few years), momentum as a strategy has had a more difficult time. Still, the fact is that momentum is a risky variable factor (as most investment factors are) with an impressive long-term average return that survives all the attacks (myths) hurled against it. In this article, we defend momentum, including its use, both stand-alone (especially as a substitute for growth investing) and in combination with value, from these persistent attacks. We believe this—both myth busting and focusing on the long term—is especially important given momentum's recent performance which only wrongly reinforces the resilience of its attackers. At the same time, our goal is not to denigrate other factors, most specifically value. Although we occasionally note the irony that many of the myths we dispel come from value investors attempting to discredit momentum, several of these myths actually apply better to value investing itself. However, as we'll show in this article, value and momentum work better when used as complements, and it is the combination of the two we stress and most-strongly recommend. We are fans of both momentum and value but bigger fans of their combination (and not fans of myths at all).

Now, on to the myth busting.

Myth No. 1: Momentum Returns Are Too Small and Sporadic

Although we have already cited some dispelling evidence, given that this precisely worded myth has been used in print, further exploration of this most basic issue is called for. We start with gross of costs, long-short portfolios to establish baseline results. In later sections we debunk the myths surrounding shorting, transactions costs, and the general implementability of momentum for traditional long-only investors.

Momentum's presence and robustness are remarkably stable. By this we don't mean that it doesn't have long stretches of poor performance, as does any factor, or short stretches of extreme performance; we mean the overall evidence across very long periods of time and in many places. Again, momentum is present in U.S. stocks over very long time periods and, following its academic discovery in the early 1990s, has been shown to be robust out-of-sample (an important exercise we will repeat here), in the individual stocks of other countries, for stock markets, and for completely different asset classes, such as bond markets, currencies, commodities, and others. It has become one of the preeminent empirical regularities studied by academics and practitioners. To see why, we will provide evidence that anyone can replicate. Most of the analysis is based on factors from Professor Kenneth French's website and focuses on momentum within U.S. stocks. Some definitions are needed and we follow Professor French here:

- RMRF represents the equity market risk premium, or aggregate equity return minus the risk free (U.S. Treasury bill) rate. It is the return from simply being long equities at market-capitalization weights and, unlike the other factors, is not a spread return between one set of stocks and another but between all stocks and cash.
- SMB (small minus big) represents a portfolio that is long small stocks and short big stocks to capture the size effect.

- HML (high minus low) represents a portfolio that is long high book-to-price stocks and short low book-to-price stocks representing value investing.
- UMD (up minus down) represents a portfolio that is long stocks that have high relative past one-year returns and short stocks that have low relative past one-year returns[6] to capture momentum.[7]

For all factors, Kenneth French's data library provides returns of the long and short sides separately, for both large- and small-capitalization securities separately, all of which we use in this article. Most of our analysis focuses on UMD and its components.[8]

Exhibit 1 reports the annualized mean spread returns and Sharpe ratios for each of the difference portfolios described earlier over three different periods: 1) the longest period for which Kenneth French provides data on all factors (starting in January 1927 and running to the end of 2013), 2) beginning in July 1963, the start date of Fama and French's seminal papers [1992, 1993] on the three-factor model and running to the end of 2013, and 3) the out-of-sample period since the original momentum papers (Jegadeesh and Titman [1993] and Asness [1994]), beginning January 1991 and running to the end of 2013.[9] Gross returns and Sharpe ratios for momentum (UMD) are large and, in fact, larger than both value and size. This is true over the full-sample period of 87-plus years of data, from 1963 onward and in the out-of-sample period from 1991 to 2013.

Critics of momentum who complain it is volatile may be pointing to some of the evidence implied in Exhibit 1. Momentum's advantage over the other factors is somewhat smaller in Sharpe ratio terms than in raw spread returns. But even considering its higher volatility, momentum still comes out on top. Stepping back and explaining a bit more, and focusing on the full period from 1927–2013, the spread of small stocks over large stocks averaged 2.9% a year, the spread of cheap stocks over expensive stocks averaged 4.7% a year, and the spread of recent winners over recent losers averaged 8.3% a year, all calculated using analogous methods (and these correspond to Sharpe ratios of 0.26, 0.39, and 0.50, respectively). This ordering, by return or Sharpe ratio, is the same over the much shorter out-of-sample period, too.

As for the word sporadic included in myth No. 1, it is not clear if this needs any more coverage since we have included Sharpe ratios above (which are adjusted, of course, for volatility, an imperfect yet very useful measure of *sporadicness*).[10] But for those who like a more common-sense method of judging whether something is sporadic, we also present in Exhibit 2 the percentage of times each strategy generates positive returns (that is, the longs beat the shorts, so for value this is how often the cheap stocks beat the expensive stocks, for momentum it's how often winners beat losers, and so on). We focus on one- and five-year horizons, though results are not very sensitive to this choice.

Exhibit 1 Returns and Sharpe Ratios of Factor Portfolios

	Returns				Sharpe Ratios			
Sample	**RMRF**	**SMB**	**HML**	**UMD**	**RMRF**	**SMB**	**HML**	**UMD**
1927–2013	7.7%	2.9%	4.7%	8.3%	0.41	0.26	0.39	0.50
1963–2013	6.0%	3.1%	4.5%	8.4%	0.39	0.29	0.45	0.57
1991–2013	8.2%	3.3%	3.6%	6.3%	0.54	0.29	0.32	0.36

Exhibit 2 Rolling One-Year and Five-Year Hit Ratios of Factor Portfolios

	% Positive, 1-Year Rolling Returns				% Positive, 5-Year Rolling Returns			
Sample	**RMRF**	**SMB**	**HML**	**UMD**	**RMRF**	**SMB**	**HML**	**UMD**
1927–2013	71%	58%	63%	81%	82%	65%	89%	88%
1963–2013	72%	61%	64%	80%	77%	65%	87%	89%
1991–2013	78%	62%	61%	76%	73%	75%	74%	71%

At one-year rolling horizons, UMD is the most consistent over the longest period. At five-year horizons (any longer gets a bit silly for an out-of-sample period of 23 years), UMD is edged by HML, perhaps (statistics on this are not dispositive) because value has more negative long-term autocorrelation than does momentum, or perhaps, as discussed in Asness and Frazzini [2013], because this version of HML is really a portfolio of mostly value with a little oddly constructed momentum thrown in.[11] But we don't recommend one versus the other—we recommend using both value and momentum together; and neither, in any reasonable form, are what any knowledgeable analyst, economist, money manager, or academic should call sporadic.

Finally, although not the direct point of this section, we elaborate a bit more on using value and momentum together. In Exhibit 3 are the statistics for a portfolio that combines HML and UMD, with 60% of the weight on HML and 40% of the weight on UMD.[12] We believe the 60/40 HML/UMD column speaks for itself.

Critics and myth makers would do well to remember that even if a factor were sporadic, it's not the sporadicness of one factor that matters, but that of the portfolio, and therefore how that factor contributes to the overall portfolio. This is portfolio theory 101. Viewing momentum alone, the myth is wrong. Viewing momentum as part of a portfolio, the myth is very, very wrong.

Of course, again, the debate can still rage on about how much of the this (for each factor, not just momentum) can be captured by long-only investors,[13] after trading costs and, for some investors, taxes, and even how much history will repeat going forward in a possibly changing world. But, starting with the basic spreads between winners and losers, as do most other authors on this topic, it's undeniable that far from being small, momentum returns are large—large after basic risk-adjustment (Sharpe ratio), and larger than other major factors, even those occasionally being promoted by the exact same crowd calling momentum small and sporadic. If the mythmakers think momentum is small and sporadic, then size, value, and even the equity premium must appear tiny and positively flighty to them. So, to sum up, who you calling small and sporadic?

Exhibit 3 Persistence of Factor Portfolios and a Value and Momentum Combination

1927–2013	**RMRF**	**SMB**	**HML**	**UMD**	**60/40 HML/UMD**
Sharpe Ratios	0.41	0.26	0.39	0.50	0.80
% Positive, 1-Year Rolling Returns	71%	58%	63%	81%	81%
% Positive, 5-Year Rolling Returns	82%	65%	89%	88%	92%

Myth No. 2: Momentum Cannot Be Captured by Long-only Investors because Momentum Can Be Exploited Only on the Short Side

In other words, the UMD factor is long winners and short losers, and those repeating this myth are asserting that most or all of the returns we showed earlier for UMD come from being short the losers. This is patently and clearly false, which somehow does not stop it from being among the most-repeated momentum myths.

First, even if it were true (it's not), for a long-only investor, being underweight a security relative to the market is economically similar to being short the security (albeit with the constraint that your largest underweight can only be as large as a stock's weight in the benchmark or market). So, asking how much of a factor return comes from the long and short side is already only partially relevant. But, admittedly, if all of the returns came from the short-side, it would certainly weaken the factor's utility for long-only investors, because this constraint on the size you can underweight a stock could, depending on goals, be binding.

However, this is not the case, and disproving myth No. 2 is easy. Simply take Kenneth French's momentum factor, UMD, and look at the market-adjusted returns (alphas) of the up (U) and down (D) portfolios separately (market-adjusted returns are the intercept of a regression of returns in excess of the risk-free rate on market returns in excess of the riskfree rate).[14] Remember, Kenneth French's UMD portfolio is just a long (winners) portfolio plus a short (losers) portfolio, and now we're just going to examine these two sides separately (so if the short portfolio goes down, it records as a positive number here as that is its contribution to UMD).

As the left panel of Exhibit 4 indicates, there is little difference between the long and short sides of momentum. Historically, almost half of the UMD premium came from *up*. For instance, over the full period, the short side contributed 5.1% to UMD (remember, that means the short portfolio fell 5.1% more than its market beta would imply it should and because it's held short, −5.1% becomes a +5.1% contribution) and the long side contributed 5.5%, which sum to the 10.6% of UMD itself. The long side is every bit as profitable as the short side. Furthermore, the exhibit shows that whether you look at U versus D over the whole sample period or over subsample periods (including the out-of-sample period), you cannot find any reliable evidence that the short side is more important than the long side—in fact, it's evenly split between them. If you do not like regressions and prefer to simply look at average returns versus the market (abstracting from the difference in market

Exhibit 4 Long and Short Side Returns of Momentum

	UMD Market-Adjusted Returns				UMD Returns Minus Market			
Sample	Short Side	Long Side	UMD	%Long	Short Side	Long Side	UMD	%Long
1927–2013	5.1%	5.5%	10.6%	51.8%	2.2%	6.1%	8.3%	73.6%
1963–2013	3.8%	5.3%	9.1%	57.8%	2.5%	5.9%	8.4%	70.1%
1991–2013	3.8%	4.8%	8.7%	56.0%	1.1%	5.2%	6.3%	83.3%

beta of the long and short portfolios), the right panel of Exhibit 4 also indicates that, if anything, on average the long side of UMD has contributed to most of its returns, the *opposite* of what critics often assert.

We present only the data for momentum within U.S. stocks here. More formally, and with a plethora of tests and specifications, Israel and Moskowitz [2013a] show that the long and short side of momentum are equally profitable using 86 years of U.S. data as well as 40 years of international equity data, and another 40 years of data from five other asset classes outside of equities. Everywhere they looked and in every way, they could not find any evidence that the short side profits were systematically larger or more important than the long side. In other words, long-only momentum is quite profitable, equally so with the short side of momentum.

If you don't like what Israel and Moskowitz [2013a] do in their paper (or don't have time to read it), you can download Kenneth French's data and try it yourself as we have done. You will find what we find: momentum does not work better, or only, on the short side.

Myth No. 3: Momentum Is Much Stronger among Small-Cap Stocks Than Large Caps

Like the other myths, this is often claimed even more histrionically as "momentum exists *only* among small caps." And, like the other myths, it is false. But what it lacks in truth it makes up for with the amusing quality of being backward at least when uttered (as is often the case) by fans of value investing—this myth happens to be true if you replace the word "momentum" with "value" (yes, we still love value, despite its weakness among large caps).

For the most detailed study to date on this topic, see Israel and Moskowitz [2013a]. In their paper, they find little to no evidence that momentum is related to size; it is almost equally as strong among large caps as it is among small caps. However, in an interesting twist, they find that though the value premium is strong among small caps, it's virtually nonexistent among large caps. Although we ourselves are big proponents of value investing (we just believe the ubiquitous data that it is better alongside momentum), to argue that momentum is all about small stocks is completely inconsistent with the facts, and far more of an argument to lay at the feet of pure value investing. To promote this myth about momentum while simultaneously advocating value investing borders on absurd.

Returning to Kenneth French's data and carrying out some simple tests, Exhibit 5 looks at UMD small, which goes long winners and short losers only among small stocks, UMD big, which does the same only among large caps, and repeats the results for regular UMD, which is done over all stocks (see Endnote 7 for exact specification). It also does the analogous exercise for HML. The exhibit shows that momentum returns among big-cap stocks are large and only slightly smaller in magnitude than returns among small-cap stocks (most factor averages get larger among small stocks, either because risk premia are greater, inefficiencies are greater, or just because volatility is greater). Value returns are also smaller among large cap than small cap (comparing the HML big column to the HML small column).

Over the entire sample period, the return to value within small cap stocks is 5.9% per annum, and within large caps it's 3.5% per annum (and, as it turns out, not statistically different from zero once you adjust for market beta).[15] The return for momentum within small

is 9.8% per annum, and within large, it's 6.8% per annum (both highly statistically significant, even after adjusting for beta). Momentum is again better in both categories, with a smaller percentage drop-off in large versus small caps than for value.

Taking this a bit further reveals a dirty little secret of value investing. It turns out that value investing, as measured by HML, which is gross of everything and implemented long-short (the test usually most biased to find strong results), is highly sensitive to the fact that Fama and French chose to split the weight in HML half to large and half to small (try building that portfolio in real life by shorting expensive tiny stocks). HML constructed among only large capitalization stocks is, dependent on the time horizon, quite a dodgy proposition (for instance, if HML were done just among large-cap stocks in the original Fama and French [1993] paper over the time frame used at that point in history, they would not have found a very strong value effect at all). Even over what we call the full out-of-sample period, 1991–2013, HML among only large stocks is a paltry 0.06 Sharpe ratio versus 0.51 within small stocks. In contrast, UMD over this period among large stocks is a 0.24 Sharpe versus 0.45 in small stocks (yes, not as big in large caps, but much closer for momentum than for value). Momentum, unlike value, is far more robust among large versus small stocks (and again makes others getting this backward really odd—people who live in houses made of cheap stocks shouldn't …).

Putting it starkly: in-sample, out-of-sample, calculated in Greenwich Connecticut, Chicago, Boston, Palo Alto, Santa Monica, Austin, or in the library with a candlestick, wherever or however you want to look, along any dimension, those who make the claim that momentum fails for large caps, yet being supporters of value investing, are not simply mistaken, they have it backward.

You might wonder how myths No. 2 and No. 3 originated. Well, two papers in particular helped contribute to these myths, with one of the papers co-authored by an author of this article. In Hong et al. [2000] and Grinblatt and Moskowitz [2004], the authors, using a more limited sample period predominantly from the 1980s and 1990s, show that momentum is stronger among small stocks and on the short side (though to be clear, neither article ever claimed momentum was nonexistent among large caps or on the long side; somehow the original evidence became twisted into something more extreme). We also showed above that over an overlapping period from 1991–2013 momentum worked better in small than large caps (0.45 versus 0.24 Sharpe ratios, respectively), but it held up in large cap and fared far better than value did. It turns out even this difference, which is still a victory for momentum, is anomalously weak. These results have proven not to be robust out-of-sample by Israel and Moskowitz [2013a]. As we've shown, over the much longer out-of-sample period from 1927 to 2013, and Israel and Moskowitz [2013a] have shown in international markets and other asset classes, the returns to momentum are really no stronger on the short

Exhibit 5 Large-and Small-Cap Returns of Value and Momentum

	Momentum			Value		
Sample	**UMD Small**	**UMD Big**	**UMD**	**HML Small**	**HML Big**	**HML**
1927–2013	9.8%	6.8%	8.3%	5.9%	3.5%	4.7%
1963–2013	11.3%	5.5%	8.4%	6.4%	2.6%	4.5%
1991–2013	8.1%	4.5%	6.3%	6.5%	0.7%	3.6%

side and are not related to size. The only evidence Israel and Moskowitz [2013a] find for size and momentum is a slightly stronger effect for small-cap winners versus large-cap winners, but even that is shown to be pretty weak. There is very little effect for size among losers and a negligible effect overall for size on momentum returns. So, if one of the authors of the original papers claiming these facts can admit that they do not hold up out-of-sample, certainly those without their names on these papers should accept the facts.

This sample-specific effect of size on momentum also explains other results in the literature that claim the same facts. For example, Fama and French [2012] look at momentum internationally and conclude that it is stronger among small-cap stocks (again their evidence does show a healthy momentum premium among large caps, just not as strong as among small caps). However, their sample period is from 1989 to 2011, which is essentially the same period over which these other short-period papers, and our out-of-sample test, find a stronger small-cap momentum effect in the United States. Over the longer sample period, there is little statistical evidence that momentum is much stronger among small-cap stocks. And, most important, over no reasonable length sample period is there any evidence that momentum actually fails among large-cap stocks.

Finally, these two myths—that momentum is dominated by the short side and mostly among small caps—are often voiced together. The motivation (we think) being to convey that it will be practically difficult and costly to implement. First, even if this were true, a long-only investor would still benefit from under-weighting small-cap losers as mentioned above. Second, and far more important, it isn't true or even close to true. This leads us to the next myth.

Myth No. 4: Momentum Does Not Survive, or Is Seriously Limited by, Trading Costs

Momentum is a higher turnover strategy than some other strategies (for example, value) and hence the question arises as to whether the premium for momentum covers trading costs (a reasonable question for any strategy). Plus, if you believed in the myth that momentum was dominated by shorting small stocks, then trading costs might seem to be an even larger potential impediment. However, just like these previous myths, the statement that momentum does not survive trading costs is false.

Although much of our other myth-dispelling can be done with Kenneth French's data, disproving this particular myth requires real-world, net-of-costs data. To the best of our knowledge, the most comprehensive work to date that analyzes real-world trading costs of factors is Frazzini et al. [2013] (FIM), which uses trades from a large institutional investor (AQR Capital) over a long period of time. Using a unique data set containing more than a trillion dollars of live trades from 1998 to 2013 across 19 developed equity markets, the authors estimate real-world trading costs for momentum, value, and size-based strategies. Their conclusion is that per dollar trading costs for momentum are quite low, and thus, despite the higher turnover, momentum easily survives transactions costs.

Unlike testing real-world strategies as in FIM, most academic studies examine portfolios that do not consider transactions costs in their design and do not allow for tradeoffs that could lead to a reduction in trading costs. They simply rebalance as automatons ignoring

costs. Trading patiently (by breaking orders up into small sizes and setting limit order prices that provide, not demand, liquidity) and allowing some tracking error to a theoretical style portfolio can significantly reduce trading costs without changing the nature of the strategy. FIM show that allowing both innovations can result in trading cost estimates (and break-even fund sizes) that are significantly smaller (larger) relative to naïve implementations.

Where did this myth come from? Several academic papers (for example, Korajczyk and Sadka [2004] and Lesmond et al. [2003]) using trading cost estimates from daily or intra-daily data found much larger effects from transaction costs on the viability of momentum strategies. However, two key differences can explain the different results. First, the studies that find much larger trading costs do so because they estimate costs for the *average* investor using aggregated daily or transaction level data for all trades in the market, which turn out to be about ten times larger than the costs of a large institutional manager, which are the costs FIM implicitly measure. Second, as discussed above, these other studies examine portfolios that do not consider transactions costs in their design, which can significantly reduce turnover and therefore trading costs further. Both factors result in trading cost estimates (and break-even fund sizes) that are an order of magnitude smaller (larger) than previous studies suggest.

History provides an analogous myth. Decades ago, when the first academic studies on the size premium came out, many declared, "You can't trade it; the trading costs would wipe out any return premium." These statements were made without realistic trading-costs data and without allowing for cost minimization through real-world, practical implementation. Similar to FIM, a paper by Keim [1999] that used real-world transactions costs from a large institutional investor—Dimensional Fund Advisors (DFA)—showed that these previous studies were flawed and had grossly overestimated transactions costs. A firm like DFA would never face the same costs as the average investor and is far smarter than to trade blindly to a set of dynamically changing strategy weights when even small modifications can greatly reduce costs. As the industry has proven for decades after these papers, small-cap portfolios can indeed be traded in an efficient manner that does not wipe out their returns. Since the premium for momentum is much higher than it is for size, and the costs to trading momentum are slightly lower than those for size (momentum is higher turnover but small caps are more expensive to trade than other stocks), you don't have to do much math to realize that momentum can easily survive trading costs.

Myth No. 5: Momentum Does Not Work for a Taxable Investor

This myth is related to momentum's higher turnover relative to other strategies (for example, value), so at face value it may seem reasonable. However, high turnover does not necessarily equal high taxes.

Papers by Israel and Moskowitz [2013b]; Bergstresser and Pontiff [2013]; and Sialm and Zhang [2013] show that momentum, despite having five to six times the annual turnover as value, actually has a similar tax burden as value. At first blush this seems counter-intuitive, until you realize the following two facts: First, momentum actually has turnover that is biased to be tax advantageous—it tends to hold on to winners and sell losers—thus avoiding realizing short-term capital gains in favor of long-term capital gains and realizing

short-term capital losses. From a tax perspective this is efficient and effectively lowers the tax burden of momentum strategies. Second, value strategies, despite their low turnover, have very high dividend income exposure, which is (in most tax regimes in history) tax inefficient. Momentum, on the other hand, more often than not has low dividend exposure. On net, this makes value and momentum roughly equally tax efficient. Since the premium for momentum is quite a bit higher than for value, yet they face similar tax rates, the after-tax returns to momentum are also higher than for value.

One more twist is worth mentioning. The analysis above didn't consider any *smart* trading, but just implicitly implemented the strategies from Kenneth French's data. Israel and Moskowitz [2013b] also look into tax-optimized versions of these strategies by designing portfolios that attempt to minimize taxes while not incurring meaningful style drift. The authors find that tax optimization is much easier to achieve through capital gains than through dividend income, which makes intuitive sense. Pushing the realization of gains from short-term to long-term status (which may often require only delaying a trade by a month) has a very small effect on the portfolio, but a large tax effect given the difference in tax rates between short-and long-term capital gains. There is a similar trade-off between short-and long-term loss realizations. But the only way to reduce dividend income is to not hold dividend paying stocks, which has a much more significant impact on a value portfolio and induces substantial style drift. Hence, tax optimization considerably improves the tax efficiency of a momentum strategy, whose tax burden comes mostly from capital gains, whereas it has a more limited effect on a value strategy, whose tax burden is driven primarily by dividend exposure.

The bottom line is that momentum survives taxes and has a tax burden roughly equal to or smaller than lower-turnover strategies such as value, especially if run optimally. Even for a taxable investor, momentum offers a healthy after-tax return premium and larger than what is provided by other strategies.

Myth No. 6: Momentum Is Best Used with Screens Rather Than as a Direct Factor

A stronger form of this myth, and wording that has been used publicly, states "momentum is not useful as a factor in portfolio construction." Yet those who say this, including those who demean momentum as a hot potato, often leave the door open to use momentum in some other, ancillary way, typically as a screen.

Though a little confusing, we presume the position summed up as "momentum screens, yes; direct factor, no" means you wouldn't want to treat momentum like value (that is, use both value and momentum to come up with a method of evaluating companies on both measures). But under this particular myth, it still makes sense to use momentum as a screen where after deciding, based on value, what to buy or sell, momentum is allowed some influence over the implementation of this rebalance. This seems like an attempt to incorporate momentum, as anyone looking at the literature and wealth of evidence (or the results noted earlier) should want to do, but not quite being willing to admit that it's a real factor.

It is, in our opinion, an attempt to have your cake but denounce it too![16]

What's strange about using momentum as a screen but not a "real factor" is that it still requires a belief in momentum, albeit perhaps a milder one than ours. In other words,

despite not giving it due credit, perhaps for fear it detracts from the value story or perhaps detracts from an efficient-markets-only point of view (we are believers that both risk-based efficient market and behavioral reasons likely contribute to the success of all of these factors), advocates of the screen approach want to find a way to use a little bit of it because of the strong evidence in its favor. The problem is (as the saying goes) you can't be a little bit pregnant. Either you believe in momentum and acknowledge the data, or you don't.

Now, there is one possible way to save the screen story and indeed claim to be just partially with child. In some sense, the fable we are about to tell unifies a bunch of the myths we discuss under one untrue umbrella. The notion of using momentum as just a screen is consistent with some of the other myths we previously dispensed with: that momentum is mostly driven by the short side, works only among small-cap stocks, and doesn't survive trading costs. If all three of these hold, using momentum only as a trading screen becomes more valid (how valid would depend on how strong these effects were). For example, imagine a long-only investor who believed in momentum but thought (wrongly) that it worked only to underweight securities (that is, the short side), believed (wrongly) that it worked only in very small cap stocks, and believed (wrongly) that it would be too costly to implement alone. That investor might still look to avoid, or screen out, very small-cap stocks that had poor momentum from his purchase list, since not buying something is free (that is, no transactions costs), and still believe momentum has efficacy for shorting small stocks (the signal momentum is giving here). Using the momentum factor in a long-only context at low weight would also achieve a similar outcome as a screen, namely, not owning these stocks, but also have more influence on what is purchased (not simply what is not purchased). So it's possible, though still far from a certainty, that if all of these things were meaningfully true, a screen could be preferred.

But alas the mythmakers are batting 200 points below the Mendoza line across each of these three assertions (that's 0.000 for you non-baseball fans).[17] Since momentum is very strong, is just as strong on the long side as the short side, works equally well among large cap as it does among small-cap stocks, and is certainly profitable after trading costs, using momentum as a screen will be significantly suboptimal versus using momentum as a factor (there is no more reason to use a screen for momentum than for value; actually, there's even less given value's weakness among large-cap stocks). Frazzini et al. [2013] explore this issue empirically and show that a factor-based approach for momentum is superior to a screen-based approach.

Myth No. 7: One Should Be Particularly Worried about Momentum's Returns Disappearing

First, we find it odd that this is often said about momentum by supporters of other factors that also face this concern, since this concern can—and should—exist for any factor. We remember 1999–2000, when investors were abandoning value investing, many with the belief that it would never work again because the world had changed. Every investor worries that the future may not reflect the past, and that return expectations may be too optimistic. When others get this admittedly valid question about the future returns of their favorite factor, again for instance the value factor, I'm sure they roll their eyes and think *here we go again*. That they'd turn around and unabashedly ask it *only* of momentum is odd to say

the least, especially given the strength and stability of momentum's historical record. No other factor, save perhaps the market itself (and that is far from clear), has nearly as long a track record (remember, there is evidence of momentum for the past 212 years), as much out-of-sample evidence (including across time, geography and even security type) or as strong and reliable a return premium as momentum (see Exhibit 1).

Our guesses as to why people ask this question more frequently about momentum are that 1) momentum is a newer factor in terms of academic attention than size or value, and 2) behavioral explanations for its origin have been pushed more prominently (though not exclusively). The first reason does not make much sense once you've seen the data, since no other factor has as much evidence behind it. The second rationale is more plausible, yet it still requires a leap of faith in that it presumes that behavioral phenomena are somehow less likely to persist than risk-based ones, and that other factors are 100% risk versus behavioral based (we know few of even the most ardent believers in the risk story who thought, when NASDAQ hit 5,000 in the year 2000, that there was no behavioral component at all to the destruction being suffered by value). The idea is that if something is driven by investor behavior, then arbitrage forces may eventually eliminate it. This is, of course, possible, but it is far from certain, and a risk-based factor can also disappear if tastes for risk change or the price of risk changes (even supporters of a pure risk-based story readily admit that the price of risk can and does change substantially through time).

Moreover, since the average investor has to, by definition, own the market, not everyone can be tilted toward the same risk factors. That is, for every value investor, there has to be a growth investor. If money managers continue to push value on everyone, then prices for value stocks will have to rise and will eventually eliminate the value premium.[18] So, yes, any factor can fail to produce returns in the future, but that possibility of failure exists for behavioral factors and risk-based factors. And, remember, the jury is still out on whether momentum is a behavioral or risk-based factor (we have not given up hope on improving upon the risk-based explanations). Perhaps the most important point is that both theories—behavioral and risk-based—provide good reasons for why the premium should persist (more on this in myth No. 10). Considering the overwhelming long-term evidence for *both* value and momentum investing, the onus is on anyone claiming future risk premia or behavior will change to these factors' detriment. This challenge has not been met, not even closely.

Having said that, we are of course interested in trying to answer whether momentum's returns are likely to disappear. Israel and Moskowitz [2013a] take up this issue by looking at a host of out-of-sample periods for momentum (after the original momentum studies were published) to see if there was any degradation in its returns. They did not find any evidence of degradation. They also looked at whether momentum's returns decreased with declines in trading costs (a proxy for the cost of arbitrageurs) and the growth in hedge fund and active mutual fund assets (a proxy for arbitrage activity). Again, the answer was no on both counts.

So there is no evidence that momentum has weakened since it has become well known and once many institutional investors embraced it and trading costs declined. This doesn't mean momentum could never disappear, but at least in the more than 20 years since its original discovery, we've seen nothing to indicate that it is being arbitraged away. Israel and Moskowitz [2013a] also looked at value and size under the same light and found that these factors, especially size, had not fared as well as momentum out-of-sample (though,

at the risk of repetition that annoys the reader, we mention again that we remain fans of *combining* value with momentum).

But let's forget all that and leave caution to the wind. What if the expected return on momentum were truly zero? Suppose, despite all of the evidence to the contrary and our strong belief it's positive, momentum had a *zero* expected return going forward. Would it still be a valuable investment tool? The answer is clearly, though perhaps surprisingly, yes. The reason is because of momentum's tremendous diversification benefits when combined with value.[19]

Again, we use Kenneth French's data to run simple optimizations where we maximize the Sharpe ratio of a portfolio combining the market (RMRF), size (SMB), value (HML), and momentum (UMD). Exhibit 6 shows the optimal weight of momentum as a function of momentum returns, while holding constant the expected returns of the other factors and the correlations between factors at their long-term averages (1927–2013). Using the average momentum premium observed in the full sample, this simple optimization would place about 38% of a portfolio in UMD, which is not surprising given the evidence discussed above (this is the rightmost of the two vertical dashed lines). Moreover, the exhibit shows that even in the extreme case, where we assume a *zero* return for momentum, the optimal portfolio still places a significant positive weight on momentum. The diversification benefits are so great that even a zero expected return would be valuable to your portfolio. The logic is simple. Since value is a good strategy and momentum is –0.4 correlated with it, one

Exhibit 6 Optimal Weight Frontier for Momentum

should expect momentum to lose money based only on that information. Yet, the fact that it does not lose but in this assumed case breaks even makes it a valuable hedge.

Put simply, even if the expected return on momentum were to disappear to zero, the benefits of diversification would still push you to want a significant weight on momentum in your portfolio (though admittedly the above is a theoretical exercise maximizing Sharpe ratio on gross of costs, long-short portfolios). Although we believe in momentum as a stand-alone factor, we've always advocated combining it into a broader portfolio, particularly with value.[20,21] And, of course, we emphatically believe the going-for ward expected premium is positive, not zero.

Myth No. 8: Momentum Is Too Volatile to Rely On

Yep, they say it this way, too. Since this might just be the sporadic myth again, and since volatility is fully accounted for in momentum's reported Sharpe ratios (in fact, that's what a Sharpe ratio does), we'll interpret this often-repeated myth as a different sort of attack, which it probably is, because it's sometimes separately included by the same people at the same time attacking momentum. To the extent we are wrong and the mythmakers were just being repetitive, please consider this extra credit. But we think when people say something like this, they don't mean regular old volatility but the admitted empirical tendency for momentum to suffer some very bad short-term periods, in particular and most recently for a few months in 2009.

As with any factor, momentum does not make money all the time and occasionally suffers large losses, and historically this has been somewhat worse for stand-alone momentum than the other factors discussed here. Spring 2009 was one of these times. But although more extreme, this isn't unlike other factors. Every factor has its dark times. Witness value investing in the late 1990s (back then we vigorously defended value from its ubiquitous critics). Unfortunately, since this particular momentum episode was recent, it has prompted some of momentum's critics to overemphasize it. We think this is a gross overreaction and mischaracterization of the facts.

Recall that momentum has a much higher Sharpe ratio than the size or value strategy, despite including this episode (again to emphasize, all of the numbers we have looked at above include the dark periods observed in the history for all of these factors, including the recent one for momentum). So, on a risk-return basis momentum still comes out on top. If momentum had a superior return but a vastly inferior, and even unacceptable, Sharpe ratio due to very high volatility, then it might have made sense to criticize it this way, but that's just not close to true. In fact, it's why we use Sharpe ratios.

Some critics of momentum use 2009 as a glaring example to imply that you don't want to invest in something that can ever have a really bad period. One prominent value manager, and prominent momentum myth spreader, says specifically, with no further explanation, as if none is needed, "Momentum is also quite variable; in 2009, it was sharply negative for U.S. stocks." That is a fairly amazing thing for a student of these factors to say. No doubt 2009 was a terrible year to be a momentum investor, particularly if momentum was all you did. However, so was 1999 for a value investor, and so was 2008 for passive equity investors. And, for those interested, 1932 was also very ugly for momentum, and 1930 ugly for value, and in each case the other came through and the 60/40 value/momentum

portfolio results were reasonably calm (have we mentioned that we like value and momentum together, not as competitors?). We highlight this not to spend time analyzing Great Depression era long-short returns but to highlight the silliness of pointing to specific-period results for attractive but risky factors one is supposed to invest in for the long-term and as part of a diversified portfolio. Of course, any decent researcher knows far better than to point to one bad period for a factor with long-term success (success that, again, *includes* that bad period) and impugn it while letting other factors have a free pass regarding events in their own histories.[22]

Nevertheless, the fact that momentum can be volatile and experience large left tails, admittedly stand-alone larger left tails than some other factors, shouldn't be ignored and deserves study. Daniel and Moskowitz [2013] look specifically at momentum crashes to try to understand these rare but turbulent times. They find that these crashes typically occur after a long bear market (say, over the last two years) followed by an abrupt market upswing. This almost perfectly characterizes the spring of 2009 and also the other most extreme momentum crash in late summer of 1932. The authors dig into what happens to momentum at these times and find that conditional market exposure is the culprit. That is, by buying winners and selling losers, a momentum strategy following a bear market will be long low-beta stocks and short high-beta stocks. Then, when the market suddenly upswings, being short high-beta stocks will be a bad strategy. In fact, Daniel and Moskowitz [2013] show that all of the crashes to momentum are driven by being short the losers; the winners actually fare well. So, ironically given our discussion in myth No. 2, this is one circumstance where shorting losers distinguishes itself—it fully explains the crash episodes of momentum. Since these episodes are driven by market exposure, the authors further devise a way to hedge much of this risk and significantly reduce the crashes. Whether one uses this hedging strategy or not, it is important to note that these two crashes (the worst ones) for momentum come during very sharp market *upswings*, periods during which the portfolios of most investors are doing well otherwise, thus making the losses potentially more tolerable. Indeed, surprising to many is that momentum's long-term (1927–2013) average beta to the long-only stock market (RMRF in Kenneth French's data) is non-trivially negative, presumably influenced by these periods, and value's slightly positive, favoring momentum in a multi-factor portfolio including the all-important market factor over this longest test period.

But there's an even simpler and equally effective way to mitigate these crashes, as we mention repeatedly: combining momentum with value. This combination has effectively eliminated these crashes in our long-term sample evidence—and not just those for momentum but also the crashes that can occur for value investing. In other words, the diversification benefits of combining momentum with value don't just appear during normal times, but also during these extreme times, which makes their combination even more valuable. For example, Asness and Frazzini [2013] show that the combination of value and momentum did not suffer as badly in 2009. Going the other way, in 1999 momentum helped ameliorate value's pain. Both factors have worked well over the long-term, but neither has a Sharpe ratio of ten, meaning that both will have hard times occasionally, but when combined together they will have fewer hard times.

Using Kenneth French's data, we can show similarly that these very poor episodes for momentum and value are ameliorated.[23] The diversification benefits between momentum and value are evident, even during these extreme times. For example, the worst draw-down

over the full sample is −43% for value, −77% for momentum,[24] but only −30% for a 60/40 combination of value and momentum.

Pointing out one very bad period for momentum that doesn't substantially change the long-term results (and was quite survivable especially if one also included a meaningful value tilt, as we advocate), and then saying "see, look, it can be really rough sometimes, you'd better avoid it" is just not an intellectually defensible argument. There is a saying at the University of Chicago, "The plural of 'anecdote' is not 'data'." Neither 2009 for momentum nor 1999 for value are indicative of the overall health and strength of these strategies. Plus, again, their combination greatly mitigates these worst times.

Myth No. 9: Different Measures of Momentum Can Give Different Results over a Given Period

Okay, this isn't a myth, it's actually true, but it's tritely obvious and yet still often hurled as a critique of momentum so we've chosen to include it. The myth, more properly stated, would say, "Different measures of momentum give different results over a given time period and that's a terrible thing." But that just sounds too silly so people don't quite get that explicit (myth tellers often prefer their statements to be less obviously humorous). Yet, the statement is meant to imply that since different measures of momentum can give different results over a given period, momentum is not a stable process and possibly data mined. This is just false.

The notion that different measures can give different results is true with *any* strategy, because there are often several valid ways of measuring the same phenomenon. For instance, value measures usually contain some form of fundamental value-to-market value such as earnings-to-price, cash-flow-to-price, or book-to-market value. And, guess what, although all are effective over the long term, they give different results over any given period! In fact, Frazzini et al. [2013] show that combining multiple measures of value, instead of relying on just one, can lead to stronger results for the factor.

As with value, momentum can be measured in various ways. The idea is to capture relative past performance. Often Occam's razor applies: the simplest measure is the best. For momentum, the past 12-month return, skipping the most recent month's return (to avoid microstructure and liquidity biases) is the most frequently used measure and has been since Asness [1994]. Many other measures have been proposed, such as various return horizons ranging from 3–12 months,[25] consistency of the past returns,[26] or measures of fundamental momentum related to earnings announcement returns or analysts' revisions.[27] Although each of these other measures may add some incremental performance, the overall momentum effect over the long term is very similar across measures as shown in Chan et al. [1996].

To guard against data mining, choosing the simplest measure or taking an average of all reasonable measures tends to yield better portfolios, as shown by Frazzini et al. [2013]. As a case in point that has oddly been used to impugn momentum's stability, Novy-Marx [2012] argues that momentum in U.S. equities is better measured by past returns from seven to twelve months ago and that using the most recent six months of returns is not valuable. However, Goyal and Wahal [2013] replicate Novy-Marx's results in 36 international equity markets and find that in 35 out of 36 countries (the only exception being the U.S.),

his result does not hold up and the past 12-month return is a superior measure of momentum, with the most recent six months of returns contributing equally to performance as the more distant six months of returns. Both ways work but simplest is best.

The fact that different measures of momentum yield substantially similar results should rationally be taken as a sign of robustness, not as a critique.

Myth No. 10: There Is No Theory behind Momentum

One of the myths often repeated about momentum is that "it has no theory," as those, for instance, who dismiss it as a "hot potato" strategy imply. This is false. Like other robust return premia, such as size and value, there is much debate regarding the explanation behind momentum, and again, like size and value, none of the models are so compelling that a consensus exists on their explanation. Still, there are several reasonable theories.

Most theories fall into one of two categories: risk-based and behavioral. While the jury is still out on which of these explanations better fit the data, the same can also be said for the size and value premia.

The behavioral models typically explain momentum as either an underreaction or delayed overreaction phenomenon (it is of course possible that both occur, making it harder to empirically sort things out). In the case of underreaction, the idea is that information travels slowly into prices for a variety of reasons (for example, investors being too conservative, being inattentive, facing liquidity issues, or displaying the disposition effect—the tendency to sell winners too quickly and hold onto losers too long). In the case of overreaction, investors may chase returns, providing a feedback mechanism that drives prices even higher.[28]

The other possibility is that the momentum premium is compensation for risk. One set of models argues that economic risks that affect firm investment and growth rates can impact the long-term cash flows and dividends of the firm that generate momentum patterns. The idea is that high-momentum stocks face greater cash-flow risk because of their growth prospects or face greater discount-rate risk because of their investment opportunities, causing them to face a higher cost of capital.[29] In addition, others argue that the presence of a correlation structure across markets and asset classes of momentum strategies is indicative of a shared economic risk.[30]

Although academics debate whether risk or behavioral explanations matter more, for the practical investor the distinction is far less relevant. Why? Because both the risk and non-risk-based explanations provide an economic reason for the premium to exist and, what's important, persist. From a risk-based perspective, as long as risks and tastes for risks don't change, the premium will remain stable and long-lived. Likewise, under the behavioral explanations, as long as the biases, behaviors, and limits to arbitrage remain stable, the premium will as well. The evidence from more than 200 years of data, in dozens of financial markets, and in many different asset classes suggests that these phenomena are not short-lived.

And remember, some of momentum's biggest myth spreaders still want to use it in some capacity (as a screen or in an ancillary way). Although we've already discussed this in depth, it's important to again note this means they believe in momentum. Earlier we said you can't be a little pregnant, so one wonders, since these folks are clearly expecting, was the father behavioral or risk-based?

Despite all this, there are still some that say, "*The momentum premium is not large enough to trade profitably, because if it were, it would be an example of market mispricing*." This statement seems to be based mostly on religion rather than fact. The idea is that if the momentum premium is really as large and robust as we show it to be, then it must be due to a market inefficiency and therefore (and here's where the religion comes in) it can't be real, because markets are obviously perfectly efficient. This thinking implies that if markets are efficient, then the data on momentum must be wrong. Although we believe risk-based, efficient market explanations play an important part in all of these factors' returns, we also believe there is a role in each, perhaps at different degrees, for behavioral explanations. Some believe it's all one or the other.[31] But even if you believe that, the statement "What you're saying can't possibly be true despite the overwhelming evidence or my one-sided view of the world would be wrong" is not an argument but a tacit admission of defeat!

There are two alarming things with this myth. First, the data are undeniable, and (as history has shown repeatedly) rejecting data on the basis of theory can be dangerous (cf. Christopher Columbus 1492, Galileo Galilei 1615, and Salem Massachusetts 1692). Second, the statement denies any possible efficient markets stories for momentum, which, as discussed earlier, do indeed exist (and is ironic coming from the efficient-markets-only crowd).

Most important, although we can debate forever how efficient or inefficient markets are (indeed, the 2013 Nobel Prize committee couldn't decide and split that year's prize between the two camps), none of this debate should diminish momentum as a valuable investment tool. The point is not to confuse the theoretical debate (which is ongoing, not just for momentum, but for other premia, like value, as well) with the empirical consensus on the efficacy of momentum. We discovered the world wasn't flat before we understood and agreed why.

Conclusion

Now that you've seen the evidence and know where to find it, those repeating the myths above regarding momentum should have a harder time maintaining credibility. They never had the facts on their side, but although the myths have been around piecemeal, no one ever assembled a detailed refutation before nor tied some of the myths together (for example, you need to believe myths about momentum's lack of small-cap efficacy, short- versus long-side efficacy, and transactions costs, to believe the myth about screens). There has now been so much work done addressing and testing these myths that repeating them means ignoring the data. Given that most of these myths can be shattered by a quick visit to Kenneth French's website (like the infomercial says, "Don't just take our word for it"), they should stop being repeated by those who want to be considered informed consumers of the research.

If one wants to challenge the evidence, that is fine, too. For instance, doing your own research and/or building your own database to attempt to establish even the slightest truth behind these myths, and explaining why you find a different result than those found to date, or picking apart the papers referenced earlier to come up with a story for why you don't believe them is fine. Momentum, or any empirical regularity, should, of course, not be immune from criticism. Quite the opposite. But eventually you must confront the data. Barring new data or a new convincing interpretation you don't get to repeat specific

falsehoods pretending that there isn't an abundance of research and evidence refuting your statements. If someone discovers something challenging or enlightening versus what we have shown, we welcome it and wish to understand more. On the other hand, if someone creates new false myths to replace the old, we stand ready! At the very least, we hope that our thorough refutation finally puts a stop to momentum critics repeating these same old myths.

Endnotes

We thank Antti Ilmanen, Sarah Jiang, John Liew, Scott Richardson, Laura Serban, Rodney Sullivan, and Daniel Villalon for useful comments and suggestions. Tobias Moskowitz has an ongoing consulting relationship with AQR Capital Management. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the views of AQR Capital Management, its affiliates, or employees.

[1]The term *relative* is important. Momentum is sometimes confused with trend following—though related, they are not the same. The process behind momentum is to rank securities relative to their peers; in contrast, trend following typically focuses on absolute price changes. Unlike trends, which increase exposure during upswings and decrease exposure during downswings, momentum takes no explicit view on the market trend, but simply ranks securities relative to each other over the same time period (though in doing so some implicit, net directional market view may exist). Momentum's winners and losers are defined no matter how the market overall is doing. For example, during 2008 a winner would have been down only a few percent relative to other stocks that on average were down more than 30 percent. During market upswings, losers would similarly be defined as stocks that were up only a few percentage points.

[2] See Geczy and Samonov [2013] for evidence of momentum in U.S. stocks from 1801 to 2012 in what the authors call, with some justifiable pride, "the world's longest backtest."

[3]See Chabot et al. [2009].

[4]See Asness et al. [2013].

[5]See, for instance, Asness [1997]; Booth [2013]; Erb [2014]; Fama and French [2008]; Huij et al. [2014]; Larson [2013]; among others.

[6]Specifically, this is defined as the past 12-month return, skipping the most recent month's return (to avoid microstructure and liquidity biases), as defined by Asness [1994] and now generally used as the standard definition of momentum.

[7]SMB and HML are formed by first splitting the universe of stocks into two size categories (S and B) using NYSE market-cap medians and then splitting stocks into three groups based on book-to-market equity [highest 30% (H), middle 40% (M), and lowest 30% (L), using NYSE breakpoints]. The intersection of stocks across the six categories are value-weighed and used to form the portfolios SH (small, high BE/ME), SM (small, middle BE/ME), SL (small, low BE/ME), BH (big, high BE/ME), BM (big, middle BE/ME), and BL (big, low BE/ME), where SMB is the average of the three small stock portfolios (1/3SH+1/3SM+1/3SL) minus the average of the three big-stock portfolios (1/3BH+1/3BM+1/3BL) and HML is the average of the two high book-to-market portfolios (1/2SH+1/2BH) minus the average of the two low book-to-market portfolios (1/2SL+1/2BL). UMD is constructed similarly to HML, in which two size groups and three momentum groups [highest 30% (U), middle 40% (M), lowest 30% (D)] are used to form six portfolios and UMD is the average of the small and big winners minus the average of the small and big losers.

[8]A link to Kenneth French's data library can be found here if you want to use the data, update the series, or check the analysis yourself: http://mba.tuck.dartmouth.edu/pages/ faculty/ken.french/ data_library.html.

[9]Another of the never-ending attempts to knock down momentum is a vague comparison to some other effects that did not hold up out-of-sample (see, for example, http://www.dimensional.com/famafrench/questions-answers/qa-caninvestors-profit-from-momentum.aspx). While examining out-of-sample results is always crucial, comments like this give us even more motivation to see if the analogy is valid.

[10]If that's not a word, it should be.

[11]See Asness and Frazzini [2013] for the argument that value as defined by Fama and French, HML, was logical for its time, before momentum had been studied, but accidentally mixes about 80% of a pure value strategy with about 20% of a very odd (accidentally and thus poorly constructed) momentum strategy. Asness and Frazzini [2013] find that a small, logical change in how value is defined produces what they call "pure value." This definition of value is positive 81% of the five-year periods, below momentum's consistency.

[12]We chose the 60/40 weights deliberately in an attempt to build a balanced portfolio of the two. Part of the reason for choosing 60/40 and not 50/50, another possible balanced allocation, is found in Asness and Frazzini [2013] who show that the classic HML is best thought of as an approximately 80/20 combination of value and momentum. Thus, a 60/40 combination of these factors is actually closer to a real 50/50 combination of pure value and momentum. Results are not sensitive to this choice.

[13]The above are, again, long-short factors, which absolutely do indeed count as some trade this way directly. In fact, they were one of the original methodologies used by Fama and French [1993] to explore the value effect. But not all investors are able or willing to go long and short so we consider their situation as well.

[14]For this exercise, it is important to look at market-adjusted returns since we know that the market returns are generally positive and a stand-alone short momentum portfolio (losers) has a different market beta than a stand-alone long momentum portfolio (winners), which will average higher betas. Though, for completeness we also show the results without market-adjustment and they hold up as well (in fact, too well, because the lack of proper risk adjustment favors the long side).

[15]Note, as we are using Kenneth French's data, value is defined here as simply book-to-price, yet there are multiple ways of measuring value (see Israel and Moskowitz [2013a]), and some would argue that book-to-price is the wrong measure of value for relatively more mature, stable firms with lower expectations of subsequent earnings growth (see Penman et al. [2013]).

[16]We leave it to the reader as to why one would denounce a cake.

[17]Hitting below the Mendoza line has often been used in baseball to define incompetent hitting and is the threshold often used to claim that a player does not belong in major league baseball, regardless of his defensive prowess. It is named after Mario Mendoza, a good defensive shortstop who actually hit 0.215 in his career.

[18]Obviously, even in this scenario, there are still limits, because it requires investors remaining patiently invested in these strategies. We are not predicting this will happen, as we are actually actively betting it doesn't occur, but simply raising the analogous possibility here.

[19]The correlation between UMD and HML in Kenneth French's data is −0.4 over the full sample period (1927–2013). We have used the definition of HML as per Kenneth French's data for objectivity and to make it easy for the reader to replicate the results, but we note that using the definition of value in Asness and Frazzini [2013] dramatically increases the magnitude of this negative correlation (to −0.7) and the power of combining value and momentum. Following their methodology, the results of this section would be far stronger.

[20]By the way, we fully recognize and acknowledge that the past ten years have not been great for momentum, with the ten-year return for UMD falling in the seventh percentile of rolling ten-year returns (going back to 1927). At the same time, the past ten years have not been great for value, either, with the ten-year return for HML falling in the fifth percentile of rolling ten-year returns.

That, of course, makes the prior ten-year return of the 60/40 combination of the two low (second percentile), but still positive (12%). You know a strategy has a pretty great history when the second percentile return is still positive. As Exhibit 6 indicates, even with the lower-than-normal returns for UMD over the prior ten-year period, the optimal weight on momentum would still be high (that is, if you knew the returns on UMD would be as low as the past ten-years the ex post highest Sharpe ratio portfolio still wants a lot of it as it hedges value so well). Also, if investors are basing their concerns about momentum's returns disappearing on the prior ten years, despite the longer-term evidence, it would seem odd that they wouldn't similarly be concerned about value's returns. For both value and momentum we, obviously, think the longer-term evidence is most convincing.

[21]Asness [2011] shows a very similar thing in Japan. Momentum for choosing stocks within Japan is one of the few places we've seen a zero historical momentum premium (albeit over the much shorter international versus USA sample). Asness shows, among other arguments including the significant chance this was a random occurrence, that even with this result, a Sharpe ratio optimizing investor with perfect knowledge of this zero premium future would *still* put substantial weight on momentum for very similar diversification/hedging reasons to those discussed here in the hypothetical.

[22]By the way, for those students of the esoteric history of quant investing, in the period immediately before 2009 some quants (not the mythmakers we discuss in this piece but quantitative investors who had embraced both value and momentum) made the decision to overweight momentum versus value and other factors since before the painful 2009 episode the long- and in particular the short-term track record for momentum versus value was even stronger (this overweight was a misguided, in our opinion, attempt to use the momentum of momentum to time itself since momentum was stronger than value in the 2007–2008 financial crisis). We get plenty wrong, none of which we will volunteer here, but as in early 2000, when many argued for abandoning value and we yelled no, we also argued against this soon-to-be-disastrous overweight of momentum immediately prior to 2009. We are cheerleaders for giving momentum a balanced, significant weight in a process, not for trying to overweight it at the right time.

[23]Using value as defined by Asness and Frazzini [2013], which we do not do here since we stick with the normal formulation, this amelioration is dramatically more pronounced. Constructing value properly, and focusing on the returns of value and momentum together, something we always encourage as they form a system, shows 2009 to be only a modest event to the properly defined combination.

[24]This is a bigger drop than for value but not by as much because it appears as, again, momentum's natural volatility is higher and value, as defined on Kenneth French's website, contains about 20% momentum as we have discussed, which acts as a hedge.

[25]Jegadeesh and Titman [1993].

[26]Grinblatt and Moskowitz [2004].

[27]Chan et al. [1996].

[28]Some behavioral models that deliver momentum: DeLong et al. [1990]; Daniel et al. [1998]; Barberis et al. [1998]; Hong and Stein [1999]; Shefrin and Statman [1985]; Grinblatt and Han [2005]; and Frazzini [2006].

[29]Some risk-based models that deliver momentum: Berk et al. [1999]; Johnson [2002]; Sagi and Seasholes [2007]; and Zhang [2004].

[30]Asness et al. [2013].

[31]Ironically, the myth-spreading supporters of value come from both ends of the spectrum, some that believe value is purely a risk-based efficient markets effect, and some who believe value is instead only the product of noise in prices coming from inefficient markets (and, rather oddly, that they are the only ones to ever have advanced that possibility). Perhaps believing in extreme explanations is correlated with mythmaking?

References

Asness, C.S. "Variables That Explain Stock Returns." Ph.D. Dissertation, University of Chicago, 1994.

_____. "The Interaction of Value and Momentum Strategies." *Financial Analysts Journal*, Vol. 53, No. 2 (1997) pp. 29-36.

_____. "Momentum in Japan: The Exception that Proves the Rule." *The Journal of Portfolio Management*, Vol. 37, No. 4 (2011), pp. 67-75.

Asness, C.S., and A. Frazzini. "The Devil in HML's Details." *The Journal of Portfolio Management*, Vol. 39, No. 4 (2013), pp. 49-68.

Asness, C.S., A. Ilmanen, R. Israel, and T. Moskowitz. "Investing with Style." *Journal of Investment Management*, forthcoming, 2013.

Asness, C.S., T. Moskowitz, and L. Pedersen. "Value and Momentum Everywhere." *Journal of Finance*, Vol. 68, No. 3 (2013), pp. 929-985.

Barberis, N., A. Shleifer, and R. Vishny. "A Model of Investor Sentiment." *Journal of Financial Economics*, No. 49 (1998), pp. 307-334.

Bergstresser, D., and J. Pontiff. "Investment Taxation and Portfolio Performance." *Journal of Public Economics*, No. 97 (2013), pp. 245-257.

Berk, J., R. Green, and V. Naik. "Optimal Investment, Growth Options, and Security Returns." *Journal of Finance,* No. 54 (1999), pp. 1553-1607.

Booth, D. "The Evolution of Dimensional." Equius Partners, Inc., Asset Class Newsletter, 2013.

Carhart, M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, No. 53, 1997.

Chabot, B., Ghysels, E., and Jagannathan, R. "Price Momentum in Stocks: Insights from Victorian Age Data." National Bureau of Economic Research Working Paper 14500, 2009.

Chan, L., N. Jegadeesh, and J. Lakonishok. "Momentum Strategies." *Journal of Finance*, No. 51 (1996), pp. 1681-1713.

Daniel, K., and T. Moskowitz. "Momentum Crashes." Working paper, University of Chicago, 2013.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam. "A Theory of Overconfidence, Self-Attribution, and Security Market Underand Over-Reactions." *Journal of Finance*, No. 53 (1998), pp. 1839-1885.

DeLong, J.B., A. Shleifer, L.H. Summers, and R.J. Waldmann. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, Vol. 45, No. 2 (1990), pp. 379-395.

Erb, C.B. "Momentum Has Not Been 'Overgrazed': A Visual Overview in 10 Slides." Working paper, SSRN, 2014.

Fama, E.F., and K.R. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, No. 47 (1992), pp. 427-465.

_____. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, No. 33 (1993), pp. 3-56.

_____. "Dissecting Anomalies." *Journal of Finance*, No. 63 (2008).

_____. "Size, Value, and Momentum in International Stock Returns." *Journal of Financial Economics*, No. 105 (2012), pp. 457-472.

Frazzini, A. "The Disposition Effect and Underreaction to News." *Journal of Finance*, No. 61 (2006).

Frazzini, A., R. Israel, and T.J. Moskowitz. "Trading Costs of Asset Pricing Anomalies." Working paper, University of Chicago, 2013.

_____. "A New Core Equity Paradigm." White paper, AQR Capital Management, 2013.

Geczy, C., and M. Samonov. "212 Years of Price Momentum (the World's Longest Backtest: 1801–2012)." Working paper, University of Pennsylvania the Wharton School, 2013.

Goyal, A., and S. Wahal. "Is Momentum an Echo?" *Journal of Financial and Quantitative Analysis*, forthcoming, 2013.

Grinblatt, M., and B. Han. "Prospect Theory, Mental Accounting, and Momentum." *Journal of Financial Economics*, 2005.

Grinblatt, M., and T.J. Moskowitz. "Predicting Stock Price Movements from Past Returns: The Role of Consistency and Tax-Loss Selling." *Journal of Financial Economics*, No. 71 (2004), pp. 541-579.

Hong, H., and J. Stein. "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets." *Journal of Finance*, LIV, No. 6, 1999.

Hong, H., T. Lim, and J. Stein. "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies." *Journal of Finance*, No. 55 (2000), pp. 265-295.

Huij, H., S. Lansdorp, D. Blitz, and P. Stein. "Factor Investing: Long-Only versus Long-Short." Working paper, Erasmus School of Economics, 2014.

Israel, R., and T. Moskowitz. "The Role of Shorting, Firm Size, and Time on Market Anomalies." *Journal of Financial Economics*, Vol. 108, No. 2 (2013a), pp. 275-301.

_____. "How Tax Efficient Are Equity Styles?" Working paper, University of Chicago Booth School of Business, 2013b.

Jegadeesh, N., and S. Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance*, 48 (1993), pp. 65-91.

Johnson, T. "Rational Momentum Effects." *Journal of Finance*, No. 57 (2002), pp. 585-608.

Keim, D.B. "An Analysis of Mutual Fund Design: The Case of Investing in Small-Cap Stocks." *Journal of Financial Economics*, 51 (1999), pp. 173-194.

Korajczyk, R.A., and R. Sadka, "Are Momentum Profits Robust to Trading Costs?" *Journal of Finance*, 59 (2004), pp. 1039-1082.

Larson, R. "Hot Potato: Momentum as an Investment Strategy." Research Affiliates Fundamentals Newsletter, 2013.

Lesmond, D.A., M.J. Schill, and C. Zhou. "The Illusory Nature of Momentum Profits." *Journal of Financial Economics*, 71 (2003), pp. 349-380.

Novy-Marx, R. "Is Momentum Really Momentum?" *Journal of Financial Economics*, Vol. 103, No. 3 (2012), pp. 429-453.

Penman, S., F. Reggiani, S. Richardson, and A.I. Tuna. "An Accounting-Based Characteristic Model for Asset Pricing." Working paper, Columbia University, 2013.

Sagi, J., and M.S. Seasholes. "Firm-Specific Attributes and the Cross-Section of Momentum." *Journal of Financial Economics*, 84 (2007), pp. 389-434.

Shefrin, H., and M. Statman. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." *Journal of Finance*, 40 (1985), pp. 777-791.

Sialm, C., and H. Zhang. "Tax-Efficient Asset Management: Evidence from Equity Mutual Funds." Working paper, University of Texas at Austin, 2013.

Zhang, H. "Dynamic Beta, Time-Varying Risk Premium, and Momentum." Yale ICF Working Paper No. 04 –26, 2004.

This page intentionally left blank

Fact, Fiction, and Value Investing

Clifford Asness, Andrea Frazzini, Ronen Israel, and Tobias Moskowitz

While recently confronting the myths surrounding momentum investing (Asness et al. [2014]), we discovered two things. First, there is as much confusion about value investing as there is about momentum investing. Second, if we debunk the mythology around momentum investing, some will get the wrong impression: that defending momentum means denigrating value. Even experienced investors often seem to wrongly assume that one cannot simultaneously believe in both value and momentum investing.

Value is the phenomenon in which securities that appear cheap, on average, outperform securities that appear to be expensive. The value premium is the return achieved by buying (being long in an absolute sense or overweight relative to a benchmark) cheap assets and selling (shorting or underweighting) expensive ones. The existence of the value premium is a well-established empirical fact. It is evident in 87 years of U.S. equity data, in more than 30 years of out-of-sample evidence from original studies, in 40 other countries, in more than a dozen other asset classes (Asness, Moskowitz, and Pedersen [2013]), and even dating back to Victorian England.[1]

Importantly, our definition of value investing is the highly diversified, academic version of value (though many practitioners also follow it). We do not focus on concentrated value-based stock picking, which we discuss further in one of our sections. In addition, our starting point is pure value, meaning price relative to some fundamental, such as book value, based solely on quantifiable measures. We do not further adjust it for other qualities for which an investor might pay more, such as faster growth or more profitable firms. (Some call this "growth at a reasonable price," though obviously we are generalizing beyond just growth.) Later in this article we address the interplay between pure value and value that considers other qualities.

Reprinted from *The Journal of Portfolio Management*, Fall 2015, Vol. 42 (1), with permission.

Recipient of the 2015 Bernstein Fabozzi/Jacobs Levy Outstanding Article Award.

Value strategies have had a long and storied history in financial markets. They date back to the late 1920s and are often credited to Benjamin Graham and David Dodd, who advocated a form of value investing that involved buying profitable but undervalued assets—a double condition which, again, is an important distinction from what we call "pure value." Indeed, value investing has been an important part of the equity investment landscape for the better part of the last century, and likely for far longer albeit undocumented. (Somewhere there must have been a Roman saying: "I came, I saw, I purchased at a low multiple.")

Despite all of this, much confusion about value investing remains. Value's opponents propagate some of this confusion in an attempt to disparage the strategy. Intentionally or not, some of those who explicitly or implicitly advocate for value also spread confusion. One notable, pure-value investor even claims not to be!

We have organized this article by identifying a number of facts and fictions about value investing that need clarification. We show that value works best with other factors, which can still be consistent with risk-based explanations for value; that it is best measured by multiple value-related variables, rather than just a single variable such as book value relative to price; that it offers exactly what the recently popular investing approach called fundamental indexing does; and that it has a weak effect among large-cap stocks, especially relative to other factors that hold up in both large- and small-cap stocks. The fictions we attempt to clarify include the false notion that value investing is only effective in concentrated portfolios; that it is a passive strategy; that it is a redundant factor in the face of newly emergent academic factors (namely, Fama and French's new investment and profitability factors); and that it is only applicable in a stock-picking context. Finally, we will take on the commonly held belief that value is solely compensation for risk, yet somehow not a scary strategy, and that a continued future value premium can only be consistent with a risk-based, efficient-markets view of the world. We certainly do not reject efficient markets, but note that value's prior and continued success can occur in a world of efficient markets, inefficient markets, or the likely truth that lies somewhere in between. In each of these worlds, value is subject to time variation and hard to envision disappearing.[2]

As in our prior article on momentum, we address the facts and fictions of value investing using published and peer-reviewed academic papers, and conduct tests using the most well known and straightforward, publicly available data[3] in U.S. equity markets.[4]

Finally, the topics we address include value investing's positive and negative attributes. We consider ourselves among value investing's strongest proponents, particularly when value is used in combination with some other factors—such as momentum and, more recently, profitability. Our discussion is not at all meant to denigrate a strategy that we believe is a cornerstone of good investing. Instead, it is merely an attempt to see the pros, cons, and even ancillary beliefs more clearly.

Fiction: Value Investing Is an Idiosyncratic Skill That Can Only Be Successfully Implemented with a Concentrated Portfolio

We focus on the highly diversified, systematic version of "value investing," not concentrated, value-based, idiosyncratic stock picking. Yet some argue that a successful value investor

must apply value in a concentrated portfolio, deeply understanding each and every security in order to uniquely identify cheap stocks. Warren Buffett, often characterized as a value investor, certainly makes this claim. To quote Mr. Buffett, "Diversification is protection against ignorance. It makes little sense if you know what you are doing."

As Buffett states, his common investment theme is to find "discrepancies between the value of a business and the price of that business in the market." He applies this philosophy to a handful of stocks that he deeply investigates and understands, holding them in a concentrated portfolio for the long term. He's obviously done it incredibly well.

But Benjamin Graham, who Buffett credits as a mentor, actually believed in the long-term evidence in favor of a diversified portfolio, as opposed to a portfolio based on a few concentrated positions. In *The Intelligent Investor* (revised in 1973), he writes, "In the investor's list of common stocks there are bound to be some that prove disappointing… But the diversified list itself, based on the above principles of selection, plus whatever other sensible criteria the investor may wish to apply, should perform well enough across the years. At least, long experience tells us so."

But does Warren Buffett's lengthy, remarkable performance prove that an idiosyncratic value process dominates a systematic one? As the saying at the University of Chicago goes, "the plural of anecdote is not data." Warren Buffett is to value investing what comedian George Burns, who smoked 10 to 15 cigars a day for 70-plus years, is to the health effects of smoking.[5] Should Burns' experience mean that we ignore the evidence that smoking kills? Unless you own the magazine *Cigar Aficionado*,[6] the answer is clearly no. Strong, long-term evidence from a wide variety of assets shows that a systematic value strategy can deliver good long-term returns. That Warren Buffett was able to successfully pick individual cheap stocks should not derail that notion. Academic and practitioner evidence shows that diversified portfolios of cheap (on pure value measures) securities healthily outperform their more expensive brethren, all without the need to pick the handful of best ones—and take on the ex ante danger of doing so.

Of course, the systematic versus idiosyncratic value investing concepts are not mutually exclusive. An investment professional who is very good at persistently identifying idiosyncratic cheap positions deserves a lot of credit. But a manager who is able to systematically invest in a group of cheap securities can also capture a positive long-term source of returns. An investor looking to invest in one of these two approaches should consider both as a way to diversify the process that generates the returns to value, as long as that investor is confident that the returns generated by both processes are indeed persistent. In our view, the diversified, systematic process provides greater comfort in this regard, but that does not mean that a concentrated value process could not also add value over time.

The point is certainly not to denigrate Buffett's approach or his record, but to emphasize that value investing encompasses more than Buffett's version. In fact, we can think of these two things separately. The returns of a diversified portfolio of value stocks over their more expensive counterparts are available to any who choose to pursue them, and should be available at a reasonably low fee. Picking the exactly right small handful of value stocks may or may not be possible, but attempting to do so comes with both additional downside dangers and upside returns, and usually comes with a higher fee if purchased in the world of active management.

In this article, we extoll and also critically examine the diversified value-investing process, pointing out that it exists largely separately from the ability to build highly concentrated portfolios, but without necessarily dismissing that possibility. At the very least, we hope to convince the reader that these two investment philosophies are not mutually exclusive or indeed even in competition with one another.

Fiction: Value Is a Passive Strategy because It Is Rules-Based and Has Low Turnover

Although we differentiate between systematic value investing and Buffett-style concentrated active stock picking, we often hear that what we call systematic value is a passive strategy. In particular, some claim that value is passive, much like the strategy of buying and holding the equity market index. The implication is that a value strategy does not make active choices and is therefore not active management.

Our own view is that anything that deviates from the market portfolio, which weights assets in proportion to their market values, is active by definition, because the market portfolio is the only portfolio that everyone can hold simultaneously. A portfolio that deviates from market weights, on the other hand, must be balanced by other investors who are willing to take the other side of those bets. For every value investor who selects cheap value stocks, there must be an investor on the other side who is underweight value and overweight expensive growth stocks. Everyone cannot tilt toward value at the same time.

Some might argue that our definition of passive is too narrow, and that a better definition is this: a passive strategy is one that follows simple rules and has low turnover. However, simple counter-examples show that this is not an appropriate definition. For example, consider a single-stock portfolio that buys and holds that stock in perpetuity. Imagine a company employee putting all of her wealth in that stock. By this definition, we would consider this a passive portfolio, even though it is clearly a concentrated, idiosyncratic, and active bet in one firm. We would consider Buffett's portfolio passive under this definition, too, because his turnover is even lower than that of a typical diversified, systematic value strategy. When single-stock portfolios and Warren Buffett fall under the definition of "passive," the definition seems flawed.

The rules-based portion of this definition also fails. High-frequency traders who trade in microseconds are by definition rules-based, but we would be hard-pressed to call them passive investors.

Quite frankly, we think the debate between active and passive management is just semantics. The main issues facing investors are relatively simple: what they are buying (i.e., long-term expected returns with correlation properties that are valuable to their portfolio), at what price, and whether there are good reasons, either risk or behavioral, to believe these returns will persist. Whether products are active or passive is irrelevant, so long as they add value to a portfolio after cost.

We believe that a fairly priced, systematic, disciplined, rules-based, and low-turnover portfolio exposed to value (and to other factors as well) is a great investment, no matter what you call it—and we would call it active.

Fact: "Fundamental Indexing" Is Only Systematic Value Investing

Some acclaimed value investors, such as Buffett, are actually not pure value investors, because they also consider other quality measures. Before getting to that, we must first address the flip side of this issue: pure value investors who claim they are not pure value investors.

We have seen investment products with names such as "fundamental indexing" and others that are labeled "smart beta," but are just simple systematic tilts away from an index and toward value investing. That's great. We object only when people claim that their simple value tilts are something more or (even worse) a new discovery (Asness [2006] and Arnott [2006]).

Fundamental indexing's violations of this have been particularly acute, with claims that it is related to value investing, but somehow different and better.[7] We see various theories about why value (or fundamental indexing) works, many of which are reinventions that are treated as new, despite the fact that academics have long considered noise and mispricing as potential reasons for value's long-term success. They further obfuscate the subject, looking for points of distinction from value investing where none exist.

Fundamental indexing works by weighting stocks according to various fundamentals—book value, dividends, cash flows, sales, earnings, etc.—as opposed to market capitalization, which is the weight in traditional index funds. Fundamental indexing proponents correctly point out that if prices contain errors (and they do, of course) then by definition an index based on market capitalization overweights the too expensive and underweights the too cheap. Weighting by fundamentals creates an investment product that is less prone to this potential bias. Also, unlike some popular but impractical alternatives, such as equal-weight indices, it creates an investment product that is deep, liquid, and investable.

But when fundamental index proponents say that fundamental indexing is more than value investing, they add confusion and hide the truth. An equation illustrates the point. In a fundamental index (FI) based on one measure (e.g., book value), the weight of stock *i* held in the fundamental index is a function of its weight in a traditional market capitalization weighted index and its relative price-to-book ratio, as follows:

$$FI_i = MKT_i \times (P/B_{MKT}/P/B_i)$$

Here FI_i is the weight of the stock in the fundamental index, MKT_i is its weight in the traditional index weighted by market capitalization, P/B_{MKT} is the price-to-book ratio of the market-cap-weighted index, and P/B_i is the price-to-book ratio of firm *i*. Forming a fundamental index over muliple measures, which we find preferable to a single measure, complicates the mathematics but doesn't change the spirit.[8] The weight of firm *i* in the fundamental index is a direct function of relative valuations, as measured by the price-to-book ratio of firm *i* and the market itself.

Remember, value strategies have been around forever (e.g., Graham and Dodd [1934], Victorian England) and systematic value strategies have been studied in depth since at least the mid-1980s (e.g., Rosenberg, Reid, and Lanstein [1985] and Fama and French [1992]), and likely far longer. When posed this way, few investors would say that systematically

overweighting stocks with low price-to-book ratios and underweighting stocks with high price-to-book ratios, versus a capitalization-weighted index using a simple formula, is anything other than a pure value strategy.

To see what the data say about this, we run a regression using monthly data from Kenneth French's website and the backtest of fundamental index in large-capitalization U.S. stocks from 1962 through early 2014 (available on Bloomberg).[9] The left side measures by how much the fundamental index beats the cap-weighted market over this period, which we expect to be positive, by subtracting the returns of the cap-weighted market portfolio from the monthly FI backtest. The right side is the HML factor: the return spread between a diversified portfolio of stocks with low price-to-book ratios and stocks with high price-to-book ratios. To our knowledge, no one disputes that HML is pure value. The results, with *t*-statistics in parentheses: an intercept of –4 basis points per annum (–0.10), + 0.37 loading on HML (35.2), and a 66% R-squared.

The results show that the fundamental index delivers no additional average returns beyond its value returns, as measured by Fama and French's HML factor. The intercept is –4 basis points, showing that, after adjusting for pure value using Fama and French's HML factor, FI underperforms by 4 basis points. This is statistically no different from zero, as indicated by the –0.10 *t*-statistic. The enormous 35.2 *t*-statistic on HML and the very high R-squared testifies to the strong relationship between FI and value. Adding other factors or changing the stated factors (e.g., using Asness and Frazzini's [2013] version of HML, adding momentum and size factors, defining value more broadly than does HML using multiple measures, as fundamental indexing does, etc.) can change the results, moving the intercept in either direction, but the very tight relationship with value remains.[10]

For example, Asness and Frazzini [2013] find that the Fama-French value factor unnecessarily lags price in its formation and propose an alternative version of HML that uses up-to-date prices rebalanced monthly. FI rebalances annually and uses up-to-date prices when it does so; Fama and French's HML uses a six-month price lag in rebalancing. On this dimension, therefore, FI is somewhere between Fama and French and Asness and Frazzini. When we add the Asness and Frazzini version of HML (called HML-DEV, after that article's title) to the regression (*t*-statistics in parentheses), we see an intercept of –6 basis points per annum (–0.20), +0.21 loading on HML (14.7), +0.18 loading on HML-DEV (14.8), and a 75% R-squared.

As its design suggests, fundamental indexing comes out almost dead in the middle of these two HML constructs, with *t*-statistics of approximately 15 on both and an improved R-squared of 75%, up from 66% on just Fama and French's HML alone. (The R-squared is 75%, even though the fundamental index is formed on four value measures and HML and HML-DEV are formed on just on one.) The intercept also drops another 2 basis points to –6 basis points, but is still not reliably different from zero. (If we simply regressed the average of HML and HML-DEV, we get a *t*-statistic of 43.3—even more impressive than the earlier value of 35.2.)

Again, specific results will of course vary, based on specific regressors. But the core result—that FI loads gigantically on value—is very robust. A value strategy may be better or worse, as shown by a positive or negative alpha to a value index such as the Russell 1000 value, for example, but it is still a value strategy if the loading and R-squared are large. We can have reasonable arguments about whether a *t*-statistic of 2.0 or 3.0 is a standard of significance, but not about the significance of a *t*-statistic of 35 or 43. The R-squared is

certainly not 100%, because different choices were made in constructing the fundamental index versus building HML. Fundamental indexes use multiple measures of value, while both forms of HML use only a price-to-book ratio. Does that make one value and the other not? Of course not, though some have claimed otherwise.[11]

To be clear, there is room for many fairly priced value strategies. We believe in the value effect and do not consider it our private sandbox. Fundamental indexing does several things we like in a value strategy: it uses multiple measures, uses up-to-date prices when rebalancing, and does some implicit timing based on the size of valuation differences across stocks that we find intuitively appealing, even though it historically adds only modest benefit. It is a clear, simple, even clever way to explain and implement value investing.

Still, it is obvious that Fundamental indexing is only a systematic value strategy, and a simple one at that. It is not unrelated to value (the story 10-plus years ago) or related to, but still different from value (the story now). It is exactly value. Fundamental index proponents are free to argue why their version of value is better.

The arguments that fundamental indexing is not value, and pure value at that, should end. Fundamental indexing is literally a simple value tilt. It is time for us all to acknowledge that. Then we can get on with arguing over whose value portfolio is better.

Fact: Profitability Can Be Used to Improve Value Investing and Still Be Consistent with a Risk-Based Explanation for Value

Some have argued that using profitability, or other quality measures, to enhance a value strategy is inconsistent with a risk-based, efficient-markets view of the world. We don't believe that is necessarily true. The efficient-markets hypothesis (EMH) states that all information should be incorporated into prices, so that any return predictability must be about risk premia. Nowhere does the EMH state that all firms should have the same price or the same price multiple, such as book- to-price ratio.

In fact, the use of profitability to enhance value strategies can be consistent with an efficient- or inefficient-markets view of the world. By cleaning up valuation ratios to identify which firms have low (high) book-to-price ratios because they are more (less) profitable rather than less (more) risky, profitability helps identify the riskiest (highest expected return) assets from an efficient-markets perspective. From an inefficient-market scenario, profitability helps find the most underpriced assets with the best hopes for higher future returns. Put simply, under either hypothesis not all firms should have the same book-to-price ratios, and measures such as profitability can help remove the variation in book-to-price ratio that comes with variation in quality rather than in expected return. Both stories provide a role for profitability in making valuation ratios more informative.

Graham and Dodd advocated the use of profitability and other quality measures to clean up value. Although Graham (and perhaps Dodd) was more systematic than Buffett, neither were pure value investors as we use the term here. Witness the main criteria for security selection from *The Intelligent Investor* (revised in 1973): 1) adequate size; 2) a sufficiently strong financial condition; 3) continued dividends for at least the past 20 years; 4) no earnings deficit in the past 10 years; 5) 10-year growth of at least one-third in per-share

Exhibit 1 Combining Value with Momentum and Profitability

	Value	Profitability	Momentum	60/40 Val/Prof	60/40 Val/Mom	33/33/33 Val/Mom/Prof
Sharpe ratio	0.46	0.42	0.57	0.58	0.79	0.84

earnings; 6) stock price no more than 1.5 times net asset [book or balance sheet] value; and 7) price no more than 15 times average earnings over the past three years.

Most definitions would consider only the last two criteria valuation metrics. In our view, the others are useful in identifying growing, high-quality companies.[12] In fact, these criteria for security selection are consistent with Peter Lynch's (the portfolio manager of Fidelity's flagship Magellan fund) concept of growth at a reasonable price (GARP).

Not all stocks should necessarily sell at the same valuation ratios. That pure value investing ignores this truism and still works so well is a testament to its power as an investment tool. A very good strategy can survive a little noise.[13] However, an investor can do even better by recognizing that valuation ratios do not need to be treated exactly the same. Using measures of earnings quality or profitability can identify the cheap and profitable firms that can give a portfolio an even bigger boost.

Value and growth strategies such as profitability and momentum are not incompatible in theory. What do the data say? Exhibit 1 reports the annualized Sharpe ratios of value (HML), momentum (UMD), and profitability (Fama and French's RMW), as well as various combinations. As profit measures are only available beginning in July 1963, the results cover 1963 to 2014 and in each case use the long-short version of the factor.

As the exhibit shows, a simple 60/40 combination of value with profitability improves value's Sharpe ratio from 0.46 to 0.58 over this time period. Further, a 60/40 combination of value with momentum results in an even bigger improvement in Sharpe ratio, to 0.79. Importantly, as the last column of Exhibit 1 shows, with a one-third equal weighting of value, momentum, and profitability, the improvement in Sharpe ratio is even higher, at 0.84.[14] Hence, a value portfolio's Sharpe ratio almost doubles when we combine it with growth-like strategies, such as momentum and profitability. Interestingly, using a simple optimizer to choose positive weights and maximize the portfolio's Sharpe ratio gives us weights that are close to one-third in each category.[15]

Although adding measures such as profitability to improve a value strategy can still be consistent with an efficient-markets/risk-based view, that story works best if profitability is both negatively correlated with value and does not itself carry a positive premium. For example, all else being equal, a negative correlation with the value factor should imply a negative expected return, so if the return is zero or not as negative as expected, it can add significant diversification benefits. Profitability (and especially momentum) is strongly negatively correlated with value. However, its returns are not zero or less negative than expected; instead, they are strongly positive. This, of course, makes profitability (and momentum) an even more valuable factor to add to value, a fact that is difficult to reconcile from an efficient-markets point of view. If profitability is merely cleaning up value—something consistent with both risk and behavioral theories—then its solo efficacy should be flat or neutral. Yet profitability has a strong positive return premium. For those who take a risk-based view of the value premium, this presents a challenge, as expensive, high-quality stocks also enjoy a positive premium.[16] Our view is that most of these factors work

for a combination of reasons—risk and behavioral—and these results fit nicely into that paradigm.

For both systematic, diversified value managers and concentrated value managers, adding profitability measures as another bona fide factor can greatly improve a portfolio, which neither proves or disproves that value is consistent or inconsistent with efficient markets. Clearly, value does not work best alone. Far from it. Combining it with other economically intuitive and empirically strong factors, such as profitability and momentum, builds the best portfolio.

Fiction: Value Is "Redundant"

Fama and French [2014] advance a new five-factor model (FFM) that adds a profitability factor (RMW) and an investment factor (CMA) to their 1993 three-factor model. Similarly, Asness, Frazzini, and Pedersen [2014] advance a model that adds a composite quality factor that contains both profitability and investment-related factors that are commonly considered part of a firm's quality.

The FFM claims that HML, that stalwart value factor, is redundant, in the sense that it adds nothing beyond the other four factors in explaining returns. This caused enough of a stir that Fama and French decided to write about it, explaining "When we say that HML is redundant, what we mean is that its average return is fully captured by its exposures to the other factors of the five-factor model. This means HML has no information about average returns that is not in other factors, so we do not need HML to explain average returns."[17] Should we stop considering the facts and fictions related to value investing and concentrate instead on just the other four factors? Should we stop building investment products with value investing as a core feature? We argue no.

In principle there is nothing wrong with the implication that there are better ways to measure and capture the value effect, as proxied by HML, such as a combination of profitability and investment. It would not mean that value is useless, but that its efficacy could be fully captured in other ways. However, we do not believe value is redundant in even this way. Our reason has to do with two things we have already discussed: 1) the Fama-French HML factor's use of a highly lagged price and 2) momentum, the one missing factor that Fama and French have never embraced in their academic work.

Fama and French [2014] explicitly omit momentum, despite the overwhelming evidence that it contributes to explaining returns and is itself not captured by the other five factors, while simultaneously keeping value, despite their own evidence showing that their other factors drive out value. It's ironic that the momentum factor they left out resurrects the most famous factor associated with them—value—and is also consistent with a long-standing view that value and momentum are best viewed together, as a system. (See Asness et al. [2014], Asness, Moskowitz, and Pedersen [2013], Asness and Frazzini [2013], Asness [1997], and Asness et al. [2015].) Failure to do so can result in erroneous conclusions and poorer investment decisions. Value's apparent redundancy in the five-factor world is one example.

Two things that are related to viewing value and momentum together resurrect HML within the new Fama and French five-factor model. The first is the decision to explicitly include both a momentum factor and a value factor in the model. The second is the decision

Exhibit 2 Replicating the Fama and French Five-Factor Model (Table 6) and Resurrecting Value

	Intercept	RMRF	SMB	RMW	CMA	UMD	R-squared
HML	–0.48%	0.01	0.02	0.23	1.04		52%
	(–0.46)	(0.37)	(0.81)	(5.38)	(23.03)		
HML	0.52%	–0.01	0.03	0.24	1.03	–0.11	54%
	(0.51)	(–0.35)	(1.04)	(5.96)	(23.37)	(–5.92)	
HML-DEV	0.23%	0.06	0.00	–0.02	0.95		28%
	(0.15)	(2.04)	(0.11)	(–0.30)	(14.24)		
HML-DEV	4.87%	–0.01	0.03	0.07	0.89	–0.52	68%
	(4.74)	(–0.32)	(1.15)	(1.60)	(20.01)	(–27.31)	

to construct that value factor by using timely price measures, a seemingly small change that has sizable consequences.

Exhibit 2 replicates the Fama and French results, where other factors seemingly make HML redundant, as in the first panel of table 6 in Fama and French's [2014] working paper. The table reports results of regressing the monthly returns on each of their factors individually from July 1963 through December 2013. We ask whether a linear combination of the other factors can essentially replicate a factor. A factor so spanned is redundant.

The first row of Exhibit 2 replicates Fama and French, though we report the intercept in percent per annum.

Each row reports the regression coefficients, with *t*-statistics below in parentheses. The first column reports the intercept or alpha from the regressions; the last column reports the R-squared. If the intercept is reliably different from zero—e.g., statistically significant at a reasonable level of confidence, usually means having an absolute *t*-statistic value greater than 2—then the dependent variable factor in question contributes more to explaining returns than do the other four factors. If the intercept is statistically no different from zero, then the other four factors span or subsume the factor in question, making the factor redundant.

As the first row shows, HML is redundant in this particular model, which may be a surprise.

Digging more deeply into the HML regression results, we see that HML is not explained by or strongly related to RMRF or SMB, whose coefficients/betas/factor loadings are near zero. However, HML has large and significant exposure to RMW, the profitability factor, and an absolutely gigantic exposure to CMA, the investment factor (with a beta of 1.0 and a *t*-statistic of 23). This eliminates HML's alpha. After accounting for positive covariance with profitability (cheaper firms are more profitable firms, on average) and conservative investing (cheaper firms invest more conservatively, on average), there is no intercept left. In fact, it is even a tad negative.

This means that HML can be reconstructed and is better explained by a combination of RMW and CMA. But is the reverse true? Can some combination of HML and RMW explain CMA? Or a combination of CMA and HML explain RMW? The answer is no, as separate analysis shows.[18]

This does not mean that value is an ineffective solo strategy—far from it. It simply means that, after accounting for the two new factors, value does not add additional returns.

The next row of Exhibit 2 adds the momentum factor, UMD, to the regression. Because HML is negatively correlated with UMD (*t*-statistic of –5.92), and UMD is a positive return factor, HML's alpha increases by 1% annually. That is enough to flip its sign, but not enough for it to reach statistical significance. HML is not quite as redundant as before, but isn't resurrected, either.

To save value, we must change it. Fama and French's industry-standard HML construction uses annual June rebalancing, employing book-to-price ratio as the valuation measure to decide "H" and "L," with both book and price taken as of the prior December. Both book and price are six months old at portfolio formation and are 18 months old by the time the portfolio is next rebalanced. The initial six-month lag ensures that an investor has information in real time when forming a portfolio; as a result, the backtested results don't suffer from look-ahead bias.

As we discussed earlier, the researcher chooses which price to use. Fama and French used price and book numbers from the same date, so they match in time. Asness and Frazzini [2013] argue for a mismatch in book and price timing for two reasons. First, if you knew only that price had fallen dramatically (and vice versa, for all these examples) since you last had an accurately time-matched book-to-price ratio, you would guess that the book-to-price ratio went up, because book does not tend to move as much as price. Second, a properly constructed value strategy is naturally negatively correlated with momentum. A six-month price lag that grows to 18 months before the next rebalance throws away much of this natural, elegant, and intuitive negative correlation.

After making these arguments, Asness and Frazzini [2013] construct an alternative to Fama and French's HML by preserving all aspects of their methodology, but letting the portfolio rebalance monthly, using last month's price to scale book values. In the third row of Exhibit 2 we show the results of re-running the regression by replacing Fama and French's HML with HML-DEV. For the most part, the result is a non-event. HML-DEV experiences an intercept that is not radically different from that found by using Fama-French HML and is still no different from zero. It is now insignificantly positive instead of insignificantly negative, so it is still redundant.

Timely value, or HML-DEV, now has an economically and statistically large intercept, even with a very large loading on the positive CMA factor. The negative correlation with the successful momentum factor is that powerful. Again, value and momentum are best thought of as a system. They are both strong alone, but are much stronger together due to their negative correlation, which shows up most clearly when value is defined with timely prices.

However, when we add momentum back in the fourth row of Exhibit 2, we see meaningfully different results, all from doing two simple things.

Fama and French's latest five-factor model may be a useful way to summarize the known playing field of factors, and it brings some very good things to the table. However, for reasons we do not find compelling, it leaves out momentum. With no change necessary to the value factor, it is absolutely compelling to add momentum back, creating a better six-factor model. But as we have argued for some time, the value factor can be made timelier. Doing so makes momentum even stronger, and the value factor, rendered distressingly redundant by the five-factor model, is suddenly and clearly resurrected.

Strong proponents of pure value may rejoice at this news, but at the same time they must face the irony that it was momentum that rescued them.

Fact: Value Investing Is Applicable to More Than Just Choosing Stocks to Own or Avoid

Can value really be applied outside of equities? To many, value is a concept that applies exclusively to stocks, in part because most of the academic literature and evidence has focused on stocks and, relatedly, because the most common way people measure value is by some ratio of accounting value to market value, such as the book-to-market equity ratio. Because accounting values are nonexistent in other asset classes, such as bonds, commodities, currencies, etc., value is often thought not to apply to those assets.

We can think more broadly about value investing's goal: to identify cheap versus expensive assets. If we can measure cheapness and expensiveness in other asset classes, we can form a value portfolio in other asset classes. There are many reasonable ways to measure value, including the reversal of long-term past returns. And, as every asset has a measureable return, this is at least one measure of value we can use for any asset class.

In other asset classes, we can also form direct fundamental value measures. For instance, in bonds a measure of value is the real bond yield, or yield on a bond minus expected inflation. For currencies, deviations from purchasing power parity (PPP), as proxied by the price of a basket of goods in two countries relative to their exchange rate, might also indicate cheapness and expensiveness, as in the long run exchange rates should converge to PPP across countries.

Using some of these value measures, Asness, Moskowitz, and Pedersen [2013] document significant value return premia in bonds, country equity index futures, commodities, currencies, and equities globally from 1972 to 2011. They find not only a reliable value premium in each asset class, but a positive correlation of value strategies across asset classes. Interestingly, the value strategy correlations are higher than passive exposures to the asset classes themselves, indicating that value strategies in different asset classes using different measures, but tied together by the same theme of identifying cheap versus expensive assets, are capturing a similar phenomenon. In other words, cheap assets in one asset class move with cheap assets in other asset classes, bonded by an overall value effect that pervades all of these markets.

Value is more than a narrowly defined, equity-only concept and can be applied more broadly to any asset class. The implication is that we can create more robust and diversified value strategies that deliver better, more stable performance. Asness, Moskowitz, and Pedersen [2013] show that a diversified value strategy applied across all asset classes more than doubles the Sharpe ratio of a U.S. equity-only value strategy, such as HML. Moreover, combining value with momentum across all asset classes improves an overall portfolio by substantially more.

Regardless of whether we take advantage of these global (across geography and asset class) results, they are out-of-sample tests of the original U.S. results, which makes us more confident that both value and momentum are effective alone and even more effective together, a robust feature of the data and not an artifact of data mining.

Fact: Value Can Be Measured in Many Ways, and Is Best Measured by a Composite of Variables

Simple intuition tells us that this should be true. The opposite idea—that a single measure of anything is optimal, given estimation error, data mining concerns, and absent any strong theory—seems at best remote and most likely false.

Nevertheless, the data put this statement to the test. In academia, the predominant way to measure value is to use the book value of a firm's equity relative to its market value, referred to as the book-to-market ratio, or expressed per share as the book-to-price ratio. Fama and French [1992, 1993, 1996, 2008, and 2012] have used this particular measure of value in a series of articles. However, we know of no theoretical justification for it as the true measure of value, versus other reasonable competitors. In fact, Fama and French [1996] use a variety of fundamental-to-price ratios, such as the earnings-to-price ratio and cash flow-to-price ratio, and other measures of value, such as dividend yield, sales growth, and even reversal of the past five-year returns.

The results are consistent across measures, and the portfolios constructed from different value measures yield highly correlated returns. Exhibit 3 reports summary statistics on HML-style portfolios formed using different value measures to rank stocks. These portfolios are taken from Kenneth French's website and pertain to the top 30% of stocks (value stocks) minus the bottom 30% of stocks (growth stocks) based on book-to-market equity ratio (BE/ME), earnings-to-price ratio (E/P), cash flow-to-price ratio (CF/P), dividend yield (D/P), and negative past five-year returns.

Though there is some variation in returns across the different measures, all of the HML-style portfolios based on different value measures produce positive returns and are highly correlated.

The last column of the exhibit reports an HML-style portfolio using a composite (simple equally weighted average) of all value measures. As the table indicates, using a composite of value measures results in a more stable portfolio, as indicated by the lowest volatility relative to all of the individual measures. Comparing the traditional HML portfolio, based on BE/ME only, to the HML-style portfolio, based on the composite of all five value measures, the average returns remain about the same, but the composite HML portfolio's volatility is 20% lower, resulting in a modestly higher Sharpe ratio, even though the correlation between them is 0.9.[19]

Exhibit 3 Single vs. Multiple Measures of Value

	HML Based on					
1951–2014	**BE/ME**	**E/P**	**CF/P**	**D/P**	**5-Year Return**	**Composite**
Average	3.6%	5.3%	4.5%	1.8%	2.5%	3.5%
Stdev	9.9%	9.8%	9.9%	11.5%	8.2%	8.1%
Sharpe ratio	0.36	0.54	0.45	0.15	0.30	0.43
Correlation to BE/ME	1.0	0.8	0.8	0.6	0.5	0.9

It is worth noting that the valuation ratios, such as BE/ME, E/P, and CF/P, deliver better and more robust results than more tenuous measures, such as negative past five-year returns and dividend yield. This makes sense, as past returns do not contain any information about a firm's fundamentals and because many firms (see Fama and French [2001]) increasingly do not pay dividends. Hence, we would expect both of these measures to perform worse than the other valuation ratios. A simple composite of just the three other valuation ratios would generate a strategy that produces an average 4.5% return per year with an annualized Sharpe ratio of 0.48.

Not only is the book-to-market ratio not the only measure of value, but an average of multiple measures results in a somewhat better and more stable portfolio. This is intuitive. Each individual measure has error in it (due to accounting mis-measurement, missing accounting items for some firms, and random errors), so an average of measures helps reduce noise. Frazzini et al. [2013] and Israel and Moskowitz [2013] also show that multiple measures of value produce more stable value portfolios that deliver higher Sharpe ratios, higher information ratios, and more robust returns. As with any systematic process, unless theory dictates preferring one metric to all others, an average of sensible measures is generally the best, most robust approach.

There is an additional advantage to using multiple measures, which relates to the ability to reduce errors. A strategy's out-of-sample performance is usually better (i.e., more closely matched to the backtest) when we use an average of multiple measures. As with any specific data sample, we will always find some measures that work particularly well in sample and some that do not (e.g., E/P versus D/P in Exhibit 3). However, without theory telling us a priori why one measure should out-perform another, this is largely due to chance. As a consequence, we would not expect that same measure to outperform out of sample. Taking an average of multiple measures guards against picking one particular measure over others because it happened to work well in one particular sample. In other words, it helps prevent data mining by extracting more of the signal and avoiding overfitting errors.

In Exhibit 4 we report the Sharpe ratios of each value measure separately by decade, from 1951 to 2014. The highest Sharpe ratio in each decade is highlighted in grey, while the lowest is highlighted in black. Since we know ex post that E/P produced the highest Sharpe ratio and D/P produced the lowest, over the full period E/P should have stronger

Exhibit 4 Sharpe Ratios of Different Value Measures by Decade

	B/M	E/P	CF/P	D/P	5-Year Return	Composite
1951–1960	0.28	1.16	0.82	0.08	–0.83	0.51
1961–1970	0.78	1.02	0.86	0.34	0.77	0.85
1971–1980	0.51	0.33	0.54	–0.11	0.37	0.35
1981–1990	0.44	0.29	0.26	0.48	0.04	0.37
1991–2000	0.01	0.39	0.03	–0.15	0.74	0.20
2001–2014	0.28	0.39	0.45	0.30	0.29	0.45

Highest | Lowest

decade-by-decade performance than D/P, which it does. However, as the table shows, E/P produced the highest Sharpe ratio in only two decades: the first two (1951 to 1960 and 1961 to 1970). Over the last four decades, a different value metric produced the highest Sharpe ratio each time, including D/P from 1981 to 1990, which is the lowest Sharpe ratio value strategy over the full period. This is just an informal (though informative) way to see that attempting to choose the best single measure from theoretically similar ones is both dangerous and unproductive over the long term.

On the other hand, looking across all measures of value in each decade, we see times when all value measures do better or worse. For instance, 1961 to 1970 was a great time for value, no matter how it was measured. The years from 1991 to 2000 were not very good for value in general. Consequently, the composite index of all five value measures, reported in the last column, is better able to capture the true variation in value's returns by averaging out the errors and idiosyncrasies associated with any singular measure.

However you identify value, cheap assets outperform expensive ones. No single measure of value is demonstrably better than another. An average of multiple measures is typically best.

Fact: By Itself, Value Is Surprisingly Weak among Large-Cap Stocks

Many academic studies show that return predictability is stronger among smaller stocks, which is true for value, too, and for some potentially good reasons.[20] However, value's return predictability as a stand-alone factor is fairly ineffective among large stocks.

From French's data we can see how different large- and small-cap value are, and how weak the large-cap results are. Exhibit 5 looks at "HML small," which goes long cheap and short expensive only among small stocks, "HML large," which does the same among large caps, and repeats the results for regular HML, which is simply the average of the portfolio returns of HML small and HML large. We report the portfolios' average market-adjusted returns over four sample periods: 1) the longest period for which French

Exhibit 5 Large- and Small-Cap Returns of Value

	Value Premia (Market-Adjusted Returns)		
Sample	**HML Small**	**HML Large**	**HML**
1926–2014	5.5	1.7	3.6
(*t*-stat)	(4.07)	(1.16)	(2.81)
1926–1962*	2.5	0.1	1.3
(*t*-stat)	(1.14)	(0.06)	(0.63)
1963–1981	6.6	6.0	6.3
(*t*-stat)	(3.15)	(2.52)	(3.14)
1982–2014*	9.2	1.7	5.5
(*t*-stat)	(4.78)	(0.93)	(3.15)

**Out-of-sample periods from the original academic studies.*

provides data on date of many of the original academic studies on value, including Fama and French's seminal papers [1992, 1993] on the three-factor model, 3) the period from January 1963 to December 1981 largely covering the in-sample period of the academic studies of Rosenberg, Reid, and Lanstein [1985] and Fama and French [1992, 1993], and 4) the period from January 1982 to July 2014 covering the out-of-sample period after the original value studies. We also include *t*-statistics on the significance of the average returns to formally test whether they are reliably different from zero.

Over the entire sample period, the market-adjusted return to value within small-cap stocks is a significant 5.5% per annum, but within large cap it is an insignificant 1.7% per annum (in other words, not reliably different from zero).

Looking at the sub-period results, the only period in which there seems to be a significantly positive HML premium among large-cap stocks is the in-sample period from 1963 to 1981, when the bulk of the original academic work on value took place. Over both out-of-sample periods—prior to these studies, from 1926 to 1962, and after these studies, from 1982 to 2014—there is no evidence for a healthy value premium among large-cap stocks. Even over the entire 88-year sample period that includes the in-sample evidence, HML among large-cap stocks does not yield significantly positive returns.[21]

This may come as a surprise to some readers who remember the original academic studies using data from 1963 to the early 1980s, where large-cap value does seem to work. However, after further review, revisiting the data, and updating the analysis, there is no strong stand-alone value premium among large caps. Perhaps there never was.

Pushing this a bit further also reveals something not generally appreciated about the construction of HML, the benchmark by which most researchers measure value. HML, which is an equally weighted combination of HML small and HML large, by construction gives much more weight to small than a simple passive cap-weighted value portfolio would, leading to better-looking results. The last column of Exhibit 5 shows that HML looks alive and well in every period except 1926 to 1962, even though small-cap HML performed well. Although HML is touted and used as a benchmark and often thought of as a passive portfolio, it is actually a monthly rebalanced, equally weighted portfolio of small-cap HML and large-cap HML, in which giving 50% weight to small-cap HML significantly overweights exposure to small-cap value relative to a cap-weighted value benchmark. Furthermore, because the risk of small-cap stocks is higher than that of large caps, the exposure to small caps is even greater and more imbalanced from a risk perspective. A purely cap-weighted value portfolio with market-cap weights would look much like HML large, as the large stocks would dominate in cap-weighting. This does not reveal a significant return premium over the full sample period. It would likely look slightly better, as it would have very small, but positive, exposure to smaller stocks.

Despite the weakness of the large-cap results, we are still big proponents of value investing, even among large-cap stocks. Large-cap value's weak stand-alone evidence should not be confused with value's very valuable contribution to a portfolio, particularly one with momentum or profitability in it, as we showed earlier.

Exhibit 6 separately looks at small- and large-cap value in combination with momentum. Even though the small-cap value strategy has the higher Sharpe ratio, it is still greatly improved by combining it with momentum, raising the Sharpe ratio from 0.48 to 0.82. Large-cap value, which by itself only generates a 0.25 Sharpe ratio, combined with large-cap momentum produces a robust 0.65 Sharpe ratio, which is not so far off the combination

Exhibit 6 Sharpe Ratios of Value and Momentum Combinations in Small and Large Caps (January 1927–July 2014)

Small Cap			Large Cap		
HML	UMD	60/40 HML/UMD	HML	UMD	60/40 HML/UMD
0.48	0.49	0.82	0.25	0.49	0.65

of value and momentum for small caps.[22] In other words, combining value with momentum and viewing them as a system, gives similar results for small- and large-cap stocks. More importantly, the results are quite strong for large cap. Once again, momentum rides to value's rescue!

Fiction: Value's Efficacy Is the Result of a Risk Premium, Not a Behavioral Anomaly, and Is Therefore in No Danger of Ebbing

There are two parts to this. First is the assertion that value is a risk premium, meaning that a value strategy delivers attractive long-term returns by taking a compensated risk in a rational market. Though we certainly don't argue against this as part of the story, we argue that the evidence is far from conclusive. The academic community continues to debate this notion, and our best guess is that both risk and behavioral causes are at work.

The second part is that value (and presumably any risk premia) will not disappear in the future. Even if value is a risk premium, that does not mean it could not disappear in the future. Conversely, behavioral anomalies do not have to disappear. The second part of the statement does not necessarily follow, regardless of whether we consider value a risk premium or a behavioral anomaly.

It is important to acknowledge the lively and healthy debate regarding value's economic explanation. No model of value is so compelling that a consensus exists for its explanation. Risk-based stories center on the value premium as compensation for bearing some type of systematic risk and going through periods, often prolonged, of underperformance. Value suffering from 1998 to 2000 during the technology run-up, the Great Depression, and the global financial crisis could support the risk-based story, especially during the last two examples, as these were particularly painful times.[23] Fama and French have suggested that distress risk may be related to value's risk premium, where value stocks have a higher beta on some marketwide distress factor. Evidence for this theory is somewhat mixed, as Campbell, Hilscher, and Szilagyi [2011] point out. Furthermore, Novy-Marx [2012] argues that the results on firm profitability and its interaction with value provide a real challenge to the distress risk story. We have hit on this earlier. If value is purely a distress premium, a chance to be paid to mitigate that distress by buying stronger, more profitable companies would be odd. Typically, you must pay to alleviate risk, not get paid. If profitability were a negative return that hedged value, that would be a more consistent story. In addition, the

notion of distress is hard to reconcile with the evidence on value effects in commodities, for example, where it is difficult to think of what "distress" means.

The behavioral theories focus on investor mis-reaction to information that causes temporary mispricing. A leading story originally suggested by DeBondt and Thaler [1985] and Lakonishok, Shleifer, and Vishny [1994] and later formalized by Daniel, Hirshleifer, and Subrahmanyam [1997] is that investor overreaction drives the value premium. The idea is that value stocks are neglected stocks that investors have fled from and now shun, while growth stocks are glamour stocks toward which investors have irrationally stampeded, causing value stocks to be underpriced and growth stocks to be overpriced. This story, though it was formulated previously, is often seen to be buoyed by the technology boom and bust of the late 1990s to early 2000s and the corresponding bust and boom in value.

Academics from the camp that emphasizes rational, risk-based, efficient markets continue to wage war with academics from the camp emphasizing behavioral, irrational, and inefficient markets over what drives the value premium. The jury is still out on which of these explanations better fit the data; indeed, the 2013 Nobel Prize committee split the prize between the two camps.[24] Even so, almost all participants agree that the data are undeniable: value offers a robust return premium that's highly unlikely to be the random result of data mining. In our view, both theories have some truth, as is likely the case with most other premia. The world is rarely so bright-lined that one theory is correct and the other completely wrong. Elements of both risk and behavior are likely present.

To make things more interesting (or more complicated, depending on your perspective), if both explanations contain important elements of the truth, nothing says that the mix is constant through time. For instance, inefficiency-based behavioral reasons may have driven much of the period from 1999 to 2000, when value suffered and then soared, both at historic levels. That could be the exception that proves the rule. The important point, however, is that both theories offer good reasons to expect the value premium to persist in the future.

If value is a risk premium, however, would that imply that we do not expect it to disappear? (Consider the questions's subtext as well: if not for risk, then would it disappear?) Just because something is related to risk—and we reiterate that this is far from certain for value—does not guarantee it will not be greatly reduced. If risks or the compensation for risk changes, then so will the expected returns to that risk. Critics of efficient markets have long mistaken time-varying risk premia for a refutation of the idea of risk premia based on efficient markets. As long as risks and tastes for risks do not change, then and only then will the premium remain stable and long-lived. It might be perfectly reasonable to believe that these risks will not change, or will change slowly, but then we should state that this is our assumption. It is false to claim that something will not change, diminish greatly, or (in the extreme) disappear just because it is risk.

Conversely, if value is due to mispricing rather than to risk, it does not follow that the returns to value would necessarily and eventually disappear. For mispricing to disappear, either investor biases would have to disappear, or enough capital willing to take the other side would have to dominate trading. Efficient markets proponents often claim the latter will naturally happen, but there are impediments, such as limits of arbitrage and the more general arguments in Fama and French [2007] that prevent it from happening, and there is no reason to think that irrational investors will not stay irrational enough

to keep the mispricing going. (In fact, mispricing can get worse!) As long as the biases, behaviors, and limits to arbitrage remain relatively stable, the premium will also be stable. During the 1999–2000 technology episode,[25] while most investors either bought into tech prices hook, line, and sinker or thought them irrational, still others thought we were witnessing the rational arbitraging away of the formerly effective value strategy. Of course, rarely has an observation been so backwards. Value was not being arbitraged away; it was being ignored! And this episode happened many years after academics discovered value and eons after Graham and Dodd, with more and better information available to the whole market.

To be clear, our point is that both risk-based and behavioral stories provide plausible reasons to expect a continued value premium in the future. Evidence from more than a century of data in plenty of out-of-sample periods, in dozens of financial markets and different asset classes, and with no signs of getting weaker, despite investor knowledge of value investing going back at least three decades (see Israel and Moskowitz [2013], who show no degradation in value's returns), is further testament that the value premium is not likely to disappear soon.

Finally, what if value did disappear?[26] Despite all of the theory and evidence to the contrary, suppose that value had a zero expected return going forward. It would still be a valuable investment tool, so long as the other major risk premia or anomalies remain viable, because of value's tremendous diversification benefits when combined with other factors, such as momentum or profitability.

Following Asness et al. [2014], but applied to value instead of momentum, we run simple optimizations that maximize the Sharpe ratio of a portfolio combining market (RMRF), size (SMB), value (HML), and momentum (UMD). Exhibit 7 shows value's optimal weight as a function of value returns, while holding constant the other factors' expected returns and the correlations between factors at their long-term averages from 1927 to 2014. Using the average value premium observed in the full sample, the optimization would place about 33% of a portfolio in HML. Moreover, the exhibit shows that even in the extreme case, where we assume a zero return for value, the optimal portfolio still places a significant positive weight on value (about 13%). The logic is simple. Because momentum is a good strategy and value is significantly negatively correlated with it, we would expect value to lose money, and the fact that it breaks even makes it a valuable hedge. The diversification benefits are so great that even a zero expected return would be valuable to a portfolio. Even assuming that the value premium goes away but the momentum premium remains, we think this is strong testimony to the power of diversification across negatively correlated strategies.

Another point that is often missed is that not every investor can (or should, if it's a risk) hold or tilt toward value. Remember, all investment must by definition aggregate to the market. For every value investor, there must be a growth investor willing to take the other side. If everyone decided to tilt toward value stocks and away from growth stocks, then the value premium would cease to exist.

For both explanations of the value premium, therefore, the other side of the trade is a key component. In the behavioral stories, it is clear that those who suffer from behavioral biases chase the glamorous growth stocks and neglect the fallen value stocks. As long as enough of them survive, the value premium remains intact. For the risk-based explanation, the answer lies in investors who do not want to bear the risk associated with value, or

Exhibit 7 Optimal Weight Frontier for Value

(more accurately) those who would pay not to bear it. That is, investors who are long value get a risk premium, like an insurance premium, from those who are willing to pay for this insurance because they are naturally short or underweight value. Of course, to fully understand this insurance-based explanation, we must fully understand the catastrophe that some investors are rationally insuring against.

This then begs the question: if a manager offering value to its clients believes that value is related to risk, why not also offer a growth fund for those wishing to bear less of this risk? In other words, if HML provides a risk premium to investors willing to bear value's risk, should they not also offer HML to the other set of investors who are willing to pay to eliminate this risk? To our knowledge, no one explicitly does.

The jury is still out as to whether the value premium exists because of risk or behaviorally based explanations, and we believe the truth is likely a combination of the two. Both theories give a plausible reason to expect a value premium to persist.

Conclusion

It's been around practically forever and formally studied for at least 30 years, but there is still a lot of confusion surrounding value investing. Now that you have seen the evidence, know where to find it, and can replicate it yourself, we hope the truth behind value investing will be clearer.

As we said before, if one wants to challenge the evidence, that is fine, too. As always, if someone discovers something challenging or enlightening versus what we have shown, we welcome it and wish to understand it.

In dispelling the myths about momentum in our earlier article, and in detailing the facts and fictions about value in this one, we end up even stronger believers in both factors, and in particular their efficacy when used together.

Endnotes

We thank Phil DeMuth, Antti Ilmanen, Sarah Jiang, Johnny Kang, Samuel Lee, Dori Levanoni, John Liew, Lasse Pedersen, Scott Richardson, Rodney Sullivan, Laura Serban, and Daniel Villalon for useful comments and suggestions.

The views and opinions expressed herein are those of the authors and do not necessarily reflect the views of AQR Capital Management, its affiliates or employees.

[1]See Chabot, Ghysels, and Jagannathan [2015], who show evidence of a value effect using dividend yield in U.K. stocks going back to the 1860s.

[2]Wondering what didn't make the list? We have two honorable mentions: 1) value is better for taxable investments than are other factors (such as momentum) and 2) value has a lot more evidence behind it than do other factors. Both of these are fiction and covered extensively in Asness et al. [2014], who write from the perspective of momentum investing.

[3]Kenneth French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) provides returns for market (RMRF), small (SMB), value (HML), momentum (UMD), profitability (RMW) and investment (CMA) factors, including separate returns for both long and short sides and for large- and small-capitalization securities, all of which we use in this article. AQR's data library (https://www.aqr.com/library/data-sets) provides returns for an improved value factor (HML-DEV) using timelier price data, which we use in this article.

[4]Our results are extremely robust when ported to other countries. We invite others to verify this claim.

[5]We attribute a similar analogy to John Cochrane, from a speech about Warren Buffett and market efficiency in honor of Eugene Fama's 2013 Nobel Prize. It also appears in print in Cochrane and Moskowitz [2015].

[6]*Cigar Aficionado* published an article in 1994 that made quite a claim: "Comedian George Burns is not only a living legend, he's living proof that smoking between 10 and 15 cigars a day for 70 years contributes to one's longevity." Or, as Mr. Burns put it, "If I'd taken my doctor's advice and quit smoking when he advised me to, I wouldn't have lived to go to his funeral." Of course, there is a good chance that they were joking. We certainly hope so…

[7]For example, Arnott, Hsu, and Moore [2005] say "A Fama-French three-factor regression shows that the Fundamental Indexes have factor exposure to the value factor." Of course, we would argue that "have factor exposure" is a giant understatement!

[8]This formula originally appeared in Asness [2006]. Arnott, Hsu, and West [2008] also present this formula. While they concede it shows that "in any snapshot in time" fundamental indexing is "pure value," they go on and try to explain why it is not pure value because it is a different amount of pure value at different times.

[9]"FTSE RAFI US 1000 Total Return Index" with Bloomberg code "FR10XTR Index."

[10]As Arnott, Hsu, and West [2008] noted, this type of regression compares a long-only portfolio to a long-short portfolio, potentially and probably biasing it against the long-only portfolio. That's precisely why we subtract the market return from the left side of the regression. However, Arnott, Hsu, and West make similar comparisons when they say, "It's striking to note that the 0.3 percent Fama-French alpha on Table 10.3 *soars* to 1.1 percent in the Fama-French-Carhart analysis. This is 1.1 percent of return, which is *utterly* unexplained in a Fama-French-Carhart model!" Ignoring the hyperbole and the fact that the intercept is by definition "utterly unexplained" by the model, we think this increase in intercept is real and happens for precisely the reason Asness and Frazzini [2013] documented. They are right in that the intercept is positive, but they imply that it comes from

something to do with FI, not the more mundane improvement found in any value strategy that rebalances with current prices.

[11]Some claim that FI is not "value" as it apparently beats the Russell 1000 Value index, which is the straw man here standing in as the only pure "value" strategy. The Russell 1000 Value is an odd beast, as it is not purely based on price; it also considers specific measures of growth. We don't know anyone who considers it the only definitional standard for value. Even if it were the value standard, beating it only implies that you may have a *better* value strategy, not something different from value.

[12]Frazzini, Kabiller, and Pedersen [2013] found that Buffett bought cheap stocks, as defined by pure value measures, but he also bought low-risk and high-quality (i.e., profitable, stable, growing, and with high payout ratios) stocks. Accounting for these factors helps explain a large part of Buffett's performance.

[13]When you sort stocks only on a measure such as price-to-book ratio and prefer the cheaper ones, you implicitly state that firms should all sell for the same price-to-book ratio, so the cheaper ones must be more attractive. Can you identify the systematic characteristics that should lead us to rationally pay more or less for some firms? We argue that profitability and others pass this test. We also argue that pure value's effectiveness says that, though you can improve by accounting for quality measures, the cross-sectional differences in expected returns that pure value identifies are large enough to survive the sub-optimal act of ignoring them.

[14]We note again that these findings are robust to testing in other countries, and for use in other asset classes where appropriate. (Momentum always has an analogy for other asset classes; profitability sometimes does.)

[15]Of course, these returns are to long-short portfolios and before trading costs, taxes, and other practical costs an investor might face. However, the diversification benefits of combining value with momentum and profitability also extend to trading costs and tax considerations, and to long-only portfolios, particularly when viewed versus a neutral benchmark (see Frazzini, Israel, and Moskwitz [2012] and Israel and Moskowitz [2012]). After taxes and trading costs, the historically optimal portfolio has still been far from pure value, including sizable weights to momentum and profitability.

[16]Perhaps for this reason, some use profitability (and momentum) as a screen rather than as a separate factor in their investment process. We covered this ground in our last article on momentum, so we will only rehash the highlights here. Either you believe in these other factors and so want to use them effectively, or you do not believe in them and do not want to use them at all. Adding only a little bit of a factor via screens is inconsistent with both positions and only optimal under restrictive conditions that are wildly inconsistent with the data. Even if you only believe in a factor "a little," the proper action would be to give it a little weight in creating your desired portfolio, not to use it as a screen.

[17]See the Fama/French Forum hosted by Dimensional: "Q&A: What Does It Mean To Say HML Is Redundant?"(http://www.dimensional.com/famafrench/questions-answers/qa-what-does-it-mean-to-say-hml-is-redundant.aspx.)

[18]For more detail, see the post at https://www.aqr.com/ cliffs-perspective/our-model-goes-to-six-and-saves-value-from-redundancy-along-the-way.

[19]Regressing each measure on the market, size, and momentum factors, again from Ken French's website, yields a similar superiority for the composite with the *t*-statistic of its alpha, or intercept from the regression, being higher than those of all single-value-measure portfolios. In addition, all of the value portfolios exhibit similarly negative correlations with the momentum factor, highlighting the fact that different value measures also yield similar results in terms of their relationship to other factors, such as momentum.

[20]Penman et al. [2013] argue that the accounting system encourages firms to reflect risky activities through deferrals and accruals that depress current earnings during risky times and create a wedge

between earnings and book multiples, where book values help identify risky stocks. The importance of book values increases where earnings growth is higher and more uncertain, as in small stocks.

[21]For a more detailed study on this topic, see Israel and Moskowitz [2013], who find that the value premium is virtually non-existent among large stocks.

[22]In fact, resurrecting large cap value is quite easy. Merely adding 20% of momentum to large cap value bumps the Sharpe ratio up from 0.25 to 0.40.

[23]Those selling value strategies as a risk-based factor should emphasize these bad periods. Put differently, it would be a heroic marketing technique to call something "risk" while presenting no evidence of risk. That would be like telling your investors they get to earn the insurance premium, but will never have to pay. That is a fairy tale. If we says something is entirely risk-based, we should also explain why it is sometimes excruciating, even life threatening.

[24]Asness and Liew [2014] give a thorough discussion of this from both academic and practitioner perspectives.

[25]Some of the paper's authors are old enough to have been systematic value investors during this period!

[26]If the returns to value did disappear over time, a value investor would benefit as the returns converged to zero. Unless you believe the returns have already gone away (which, of course, we do not), you should not be concerned about this long-term possibility. In fact, you might welcome it.

References

Arnott, R. "An Overwrought Orthodoxy." *Institutional Investor*, December 2006: 36–41.

Arnott, R., J. Hsu, and P. Moore. "Fundamental Indexation." *Financial Analysts Journal*, (2005), pp. 83–99.

Arnott, R., J. Hsu, and J. West [2008]. *The Fundamental Index*. New Jersey: John Wiley & Sons, 2008.

Asness, C. "The Interaction of Value and Momentum Strategies." *Financial Analysts Journal*, Vol. 53, No. 2 (1997), pp. 29–36.

_____. "The Value of Fundamental Indexing." *Institutional Investor*, October 2006, pp. 94–99.

Asness, C., and A. Frazzini. "The Devil in HML's Details." *The Journal of Portfolio Management*, 39 (2013), pp. 49–68.

Asness, C., A. Frazzini, and L. Pedersen. "Quality Minus Junk." Working paper, AQR Capital Management, 2014.

Asness, C., and J. Liew. "The Great Divide over Market Efficiency." *Institutional Investor*, March 2014.

Asness, C., T. Moskowitz, and L. Pedersen. ''Value and Momentum Everywhere.'' *Journal of Finance*, Vol. 68, No. 3 (2013), pp. 929–985.

Asness, C., A. Ilmanen, R. Israel, and T. Moskowitz. "Investing with Style." *Journal of Investment Management*, Vol. 13, No. 1 (2015), pp. 27–63.

Barberis, N., A. Shleifer, and R. Vishny. "A Model of Investor Sentiment." *Journal of Financial Economics*, 49 (1998), pp. 307–34.

Campbell, J., J. Hilscher, and J. Szilagyi. "Predicting Financial Distress and the Performance of Distressed Stocks." *Journal of Investment Management*, Vol. 9, No. 2 (2011), pp. 1–21.

Chabot, B., E. Ghysels, and R. Jagannathan. "Momentum Trading, Return Chasing, and Predictable Crashes." Working paper, Northwestern University, 2015.

Cochrane, J., and T. Moskowitz, eds. *The Fama Portfolio*. University of Chicago Press, 2015.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam. "A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-Reactions." *Journal of Finance*, 53 (1997), pp. 1839–1885.

DeBondt, W., and R. Thaler. "Does the Stock Market Overreact?" *Journal of Finance*, Vol. 40, No. 3 (1985), pp. 793–805.

Fama, E., and K. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), pp. 427–465.

_____. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), pp. 3–56.

_____. "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity To Pay?" *Journal of Financial Economics*, Vol. 60, No. 1 (2001), pp. 3–43.

_____. "Disagreement, Tastes, And Asset Prices." *Journal of Financial Economics*, Vol. 83, No. 3 (2007), pp. 667–689.

_____. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance*, Vol. 51, No. 1 (1996), pp. 55–84.

_____. "A Five-Factor Asset Pricing Model." Working paper, University of Chicago, 2014.

Frazzini, A., R. Israel, and T. Moskowitz. "Trading Costs of Asset Pricing Anomalies." Working paper, University of Chicago, 2012.

Frazzini, A., R. Israel, T. Moskowitz and R. Novy-Marx. "A New Core Equity Paradigm." White paper, AQR Capital Management, 2013.

Frazzini A., D. Kabiller, and L. Pedersen. "Buffett's Alpha." Working paper, AQR Capital Management, 2013.

Graham, B. *The Intelligent Investor*, 4th ed. New York: Harper, 1973.

Graham, B., and D. Dodd. *Security Analysis*, New York: McGraw Hill, 1934.

Israel, R., and T. Moskowitz. "How Tax Efficient Are Equity Styles?" Working paper, University of Chicago Booth School of Business, 2012.

_____. "The Role of Shorting, Firm Size, and Time on Market Anomalies." *Journal of Financial Economics*, Vol. 108, No. 2 (2013), pp. 275–301.

Lakonishok, J., A. Shleifer, and R. Vishny. "Contrarian Investment, Extrapolation, and Risk." *Journal of Financ*e, 49 (1994), pp. 1541–1578.

Novy-Marx, R. "Is Momentum Really Momentum?" *Journal of Financial Economics*, Vol. 103, No. 3 (2012), pp. 429–453.

Penman, S., F. Reggiani, S. Richardson, and A. I. Tuna. "An Accounting-Based Characteristic Model for Asset Pricing." Working paper, Columbia University, 2013.

Rosenberg, R., K. Reid, and R. Lanstein. "Persuasive Evidence of Market Inefficiency." *The Journal of Portfolio Management*, Vol. 11, No. 3 (1985), pp. 9–16.

Craftsmanship Alpha

An Application to Style Investing

Ronen Israel, Sarah Jiang, and Adrienne Ross

Today, many investors are looking to take advantage of alternative sources of return, specifically those that are uncorrelated with traditional assets. In doing so, they have turned their attention to *style premia*—a set of systematic sources of returns that are well researched and have been shown to deliver long-run returns that are uncorrelated with traditional assets.[1] Styles have been most widely studied in U.S. equity markets, but they have been shown to work across other geographies and asset classes (such as bonds, commodities, and currencies). There is a logical, economic rationale for why they work and are likely to continue to do so.

The growing popularity of style investing has also given rise to a variety of investment options for investors. There are variations not only in the types of style portfolios but also—importantly—in how different managers choose to build those portfolios. Although practitioners might define styles with similar "labels," actual portfolios can differ significantly from one another. Even a single style such as value has variations that an investor should consider. For example, it can be applied as a tilt to a long-only equity portfolio, or it can be applied in a "purer" form through long/short strategies across multiple asset classes; it can be based on multiple measures of value or a single measure; it can be done in isolation or combined with other styles in a synergistic way. Investors should consider these differences and also recognize that there is potentially value added (or subtracted) in every step of the investment process: signal choice, portfolio construction, risk management, and cost-effective trading.

Our article focuses on the craftsmanship required to build effective style portfolios. That is, the kind of decisions that happen *after* we have already agreed on the type of style portfolio that we want to build. We start with a brief discussion of the types of style portfolios an investor may choose and then go into more detail on design decisions related to building style portfolios. Finally, we address other considerations for style investing, such as trading and risk management.

Reprinted from *The Journal of Portfolio Management Multi-Asset Strategies Special Issue*, December 2017, with permission.

Although our article focuses on equities, the takeaways apply across all asset classes. We share our thoughts on a number of enhancements that can be made without deviating from the main thesis. While many of these enhancements reflect our opinions on better ways to build portfolios, our main point is that we need to make these choices consciously. Certain design choices may improve the risk/return characteristics of the overall portfolio by enhancing returns, reducing risk, or a combination of both. We call the sources of alpha that involve implementation choices "craftsmanship alpha."

What Kind of Style Portfolio?

Style premia investing has certainly grown in popularity, but there is still considerable disagreement over which styles drive returns—which styles one should believe in. Generally, the most widely accepted and utilized styles relate to value (the tendency for cheap assets to outperform expensive ones), momentum (the tendency for outperformers to continue to outperform), defensive (the tendency for low-risk, high-quality assets to outperform high-risk, low-quality assets on a risk-adjusted basis), and carry (the tendency for higher-yielding assets to outperform lower-yielding assets). However, there are also other styles such as size or liquidity. Although the efficacy of the latter two styles has generally been more challenged,[2] it is outside the scope of this article to present the merit of each style. Instead, we aim to highlight that even general agreement on *which* style to include does not map as directly to a portfolio as many think.

To understand this point, consider, just as a starting point, all the different ways to take advantage of a single style such as value. Arguably, the most popular expression is the long-only (or "smart beta") approach, which applies tilts within equities to overweight stocks that are relatively cheap.[3] This approach results in deviations from market capitalization weights that in practice imply certain systematic style tilts.[4] Relative to other forms of value investing, long-only style tilts are typically easier for many institutions to adopt because they involve less peer risk (they have lower tracking error to conventional portfolios and benchmarks), have greater capacity, and do not require the use of leverage, shorting, or derivatives. But for investors who are more comfortable with less-traditional implementations, a "purer" expression of value would be a long/short approach, which seeks to capture the entire style premium and none of the traditional beta.[5] Such an approach may be valuable for investors who want to add uncorrelated sources of return.[6] Both long-only and long/short approaches can have merit, sometimes even for the same investor.

While single-style (long-only or long/short) investing may be beneficial on its own, we believe that investors may do better by combining styles in a multi-style portfolio. Relative to a single-style approach, multistyle approaches may produce more robust portfolios.[7] Investing in styles that are lowly correlated with each other can have attractive diversification benefits because the styles tend to pay off at different times.[8] In particular, combining value with momentum, for example, allows investors to take advantage of two different potential sources of returns. Importantly, however, how they are combined matters (we will come back to this point later).

It's also worth mentioning that if investors are comfortable with a long/short approach, they may also apply styles across a broader range of asset classes. Combining styles may offer greater diversification in a similar vein to combining asset class portfolios.[9]

Regardless of the implementation choice, we believe investing in styles will continue to produce positive long-run excess returns, but it is critical to implement them efficiently. We now turn to a discussion of portfolio construction.

How to Build Style Portfolios

Once investors have decided on the type of style portfolio they are most comfortable with (e.g., long/short value), there are a number of choices that can be made in actually building that portfolio. Put differently, two long/short strategies that rely on the same style may have different exposures and performance over time, likely a result of different design decisions in portfolio construction. While certain design decisions may result in better investment outcomes, these choices may not materialize in every period (or even over a 5- to 10-year period). However, to the extent they are based on sound economic logic and empirical evaluation over longer histories, we do expect them to pay off over the long run. We now focus on a number of the portfolio construction decisions that may improve the targeting and capturing of style premia; we focus mostly on long/short value portfolios in U.S. equities but, where relevant, may use a long-only approach or a different style to illustrate our point.

Smarter Style Measures

Typically, value portfolios sort stocks based on some measure of fundamental value relative to price, such as book value relative to price or earnings relative to price. Although these concepts might seem simple, practitioners may vary the inputs they use or the adjustments they make in defining each measure. For example, the decision whether to include intangible assets and/or nonoperating assets can differ among managers in calculating a company's book value.[10] For earnings calculations, managers may also treat unusual or infrequently occurring items differently.

Surprisingly, even the choice of market price can also vary by manager: The price used can be the latest figure or a lagged one. The standard academic approach, HML (high minus low), uses the price that existed contemporaneously with the book value, which may be lagged by 6 to 18 months because of financial reporting. To understand why this decision is important and why using lagged price might differ from using the latest price, consider a company that looked expensive based on its book value and price from six months ago but whose stock price has fallen over the past six months. If we hold book value constant,[11] this stock should then look better from a valuation perspective (because the price is lower). Yet, in a traditional definition (using lagged prices), the stock is viewed the same way irrespective of the price move. As a result, HML can be viewed as an incidental bet on both value and momentum.

To correct for this "noisy" combination of value and momentum, Asness and Frazzini [2013] suggest replacing the 6- to 18-month lagged market price with the current market price to compute valuation ratios that use more updated information. Measuring HML using current price (referred to as "HML Devil") seeks to eliminate any incidental exposure to momentum, resulting in a better proxy for true "value," while still using information available at the time of investing. Although such an adjustment may seem inconsequential,

it can actually make a big difference; ultimately, the devil is in the details (as we will show again and again throughout this article!).

Multiple Style Measures

Although stocks selected using the traditional academic measure of value (discussed previously) perform well in empirical studies,[12] there is no theory that says B/P is the best measure for value. In fact, we believe other measures can be used and applied simultaneously to form a more robust and reliable view of a stock's value[13]. In addition to the fact that no theory explains why investors would use only one measure, there is a strong theoretical argument that explains why investors would want to use more than one: Regardless of whether one believes that style returns come from capturing a risk premia or a mispricing, a multiple-measure approach can reduce the measurement noise associated with any one measure. Utilizing multiple measures can help isolate the common component (the "true" measure of value), which is the core of what we're trying to capture. As such, investors can relate prices to a variety of other reasonable fundamentals, including, but not limited to, earnings, cash flows, and sales. To illustrate the potential benefits of a multiple-measure approach, Exhibit 1 looks at rolling five-year Sharpe ratios for two value portfolios: one based on just B/P and another based on a composite of five different value indicators.[14]

The results show that both B/P and multiple value measures provide positive risk-adjusted returns, on average (0.07 and 0.26, respectively),[15] and are highly correlated with each other (roughly 0.9 correlated based on monthly returns). However, the multiple-measure value portfolio outperforms B/P in almost every five-year period since 1990.[16] The same multiple-measure approach can be applied to other styles too; for example, momentum portfolios that include both earnings momentum and price momentum may be more robust. Note that using multiple measures is not a form of factor proliferation, which can lead to concerns about data mining; instead, using additional measures leads to a more robust version of the ideas behind the factors because there isn't a single perfect definition of each style.

Stock Selection and Weighting Schemes

Another portfolio construction design choice that can result in different exposure to a given style, and thus different performance, is in selecting and weighting stocks within a portfolio. A common approach among academics is to set a cutoff and weight stocks in a style portfolio via their market capitalization. The standard academic approach—in Fama–French's HML, for example—is to rank stocks based on B/P, go long the top 33% of stocks with the highest B/P and short the bottom 33% with the lowest B/P, and then weight those stocks based on their market capitalization. However, the 33% cutoff is just one choice a manager can make, as is the choice of weighting by market capitalization.

To illustrate this point, we go back to our hypothetical value portfolio based on multiple measures from Exhibit 1. Here, we are creating a portfolio based on the top 50% of stocks with the highest "composite" value rank across five different value measures. Every month we rank stocks based on each value measure and go long the top 50% of stocks with the highest aggregate score and short the bottom 50% with the lowest score; we then

Exhibit 1 Rolling Five-Year Sharpe Ratios of B/P vs. Multiple Value Measures: U.S. Stocks Long/Short, January 1990–December 2015

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR manages. For all measures, we use current prices. The multiple measures include B/P, cash flow/enterprise value, earnings/price, forecasted earnings/price, and sales/enterprise value. Value portfolios are formed every month by ranking all U.S. stocks in the Russell 1000 universe on these metrics. Portfolios are formed by going long the top half (cheap) and short the bottom half (expensive) of stocks; stocks are weighted by market capitalization. Hypothetical returns are gross of estimated transaction costs. Hypothetical performance data has certain inherent limitations, some of which are discussed in disclosures. Please read important disclosures at the end of this article.

Sources: AQR, Russell.

weight the stocks in the resulting portfolio by their market capitalization (i.e., market-cap weight). The 50% cutoff is one choice we made; if, however, we had decided to restrict the portfolio to the top 33% (as per the standard academic approach), the portfolio would end up having more concentrated value exposure and slightly better performance, but also higher risk, and therefore virtually the same Sharpe ratio.

An alternate weighting choice that can result in even more value exposure is to account for relative cheapness within stocks held long and short. This approach assigns larger positive (negative) weights to the stocks that rank most (least) favorably on its value rank (i.e., signal weighting). While weighting stocks in a value portfolio based on signal strength results in the highest exposure to the desired style (and therefore potentially higher returns and Sharpe ratios), it tends to result in exposure to smaller, less liquid stocks, effectively foregoing some of the potential liquidity benefits of a market-capitalization weighting scheme. As such, an approach that blends market capitalization and signal strength may provide a good balance between liquidity and higher expected gross return.[17]

The results for all these weighting choices are shown in Exhibit 2. The point here, again, is not to argue for specific choices, though we of course have and will share our favorites. The broader point is that even after substantial agreement on the basic styles, there are many important choices to be made.

Exhibit 2 Selecting and Weighting Stocks in a Value Portfolio: U.S. Stocks Long/Short, January 1990–December 2015

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR manages. The portfolio is constructed every month by first sorting stocks on multiple measures including B/P, cash flow/enterprise value, earnings/price, forecasted earnings/price, and sales/enterprise value. The long/short portfolio is formed by going long (short) the cheap (expensive) stocks. Value portfolios are formed every month by ranking all U.S. stocks in the Russell 1000 universe. We test two cutoffs: the top third minus the bottom third and the top half minus the bottom half. The portfolio weights stocks by the market capitalization of the stocks, the relative ranking of the stocks, or a blend of the two measures based on a 50/50 weight. Throughout, we use current prices. Hypothetical returns are gross of estimated transaction costs. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures. Please read important disclosures at the end of this article.

Sources: AQR, Russell.

Unintended Risks

There are also specific choices managers may make to help mitigate poorly rewarded or unintended risk exposures. Even if the manager has the intention of capturing the purest form of a risk premium, certain design choices can result in other embedded risks. These risks might be unintentional, or even worse, unwanted and perhaps uncompensated.

To illustrate this point, we examine the Fama–French HML portfolio, but we construct it using current prices, as done in HML Devil.[18] Recall that this portfolio follows an academic methodology: ranking stocks based on B/P and selecting the top 33% of stocks to go long and the bottom 33% to short. This simple ranking results in a portfolio that implicitly takes style bets both within industries and across industries, without any explicit risk controls on the relative contributions of each. In addition, even though this portfolio is constructed with $1 long and $1 short, it can still have varying market beta over time.[19] The average risk exposures of this portfolio are shown in the pie chart in Exhibit 3 where we use portfolio holdings to break down risk exposures. The results show that the majority of risk is coming from market exposure and industry selection (value across industries), with only 32% of the risk coming from stock selection (value within industries).

To understand how a long/short value portfolio can have such significant market (and industry) risk, consider what happened during the technology bubble: Technology stocks were expensive based on B/P ratios and were generally higher risk as well, which meant that they exhibited higher market beta. Because these stocks fell on the short side of a naive $1 long/$1 short value portfolio, that portfolio would have been net short the technology sector and effectively the market as well (i.e., higher beta on the short side than the long side). A key question here is whether timing the market and technology sector is intentional. In general, we believe that if an exposure is intentional and believed to be compensated, it's better to separately and explicitly gain said exposure, rather than let it fall out of a naive implementation. If exposure is neither intentional nor compensated, we believe that risk should be eliminated.

So how should managers deal with these risks? One way may be to "hedge" out these risks and build a value portfolio that isolates style exposure. For market risk, that might mean constructing a portfolio with a beta of zero (i.e., the long-side beta is equal to the short-side beta), effectively hedging out market risk.[20] For industry risk, one might apply style measures (book-to-price, for example) on a within-industry basis (i.e., focus on comparisons relative to industry peers).[21] Managers may then avoid unintended industry bets but also make explicit interindustry bets where deemed beneficial. However, for styles in general, we believe it makes sense to assign more risk to within-industry comparisons given the greater breadth (i.e., more securities within industries than industries to compare). In addition, at least when it comes to value, industry selection may not be as well compensated over the long run.[22]

To understand the potential benefits of constructing a "pure play" value portfolio, observe the right-hand side of Exhibit 3, which compares the risk-adjusted performance for a B/P portfolio with significant market and industry risk to one that eliminates any incidental industry and market risk.[23] The B/P portfolio that isolates value exposure has delivered higher risk-adjusted returns. Overlaying an industry neutrality constraint on value strategies often reduces volatility while keeping long-run returns broadly unaffected, which implies higher portfolio Sharpe ratios.[24] But the tradeoff is that this adjustment typically results in higher turn-over.[25] We are not saying these choices and trade-offs are easy or obvious, just that they're real and should be made consciously.

It's also important to note that these risks are not just present in value portfolios. A momentum portfolio that simply ranks stocks based on past 12-month price returns and overweights the relative winners may result in a portfolio that implicitly makes market bets. For example, during a bull market, the stocks that tend to outperform are typically higher beta names; so by going long these stocks, the momentum portfolio will have significant market beta. Put differently, the momentum portfolio construction implicitly times the market—generally overweighting the market after an up market and underweighting after a down market. This may or may not be desired, but it should be consciously chosen!

Lastly, while we have focused on market and industry risk, there are also other unintended risks that may be hedged from portfolios. For example, in international equity portfolios, a manager can have stock, industry, country, and currency risks.[26] For instance, in creating a portfolio, if one just ranked all global stocks on valuation, part of the risk would come from buying cheap stocks within each country, part from country tilts, and perhaps

Exhibit 3 Comparing Different B/P Portfolios: U.S. Stocks Long/Short, January 1990–December 2015

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR manages. We form our value portfolio every month by ranking all U.S. stocks in the Russell 1000 universe on B/P. The portfolio is formed by going long the top third (cheap) and short the bottom third (expensive) of stocks; stocks are weighted based on market capitalization. The risk decomposition (on the left) is computed by analyzing the holdings of the hypothetical long/short value portfolio. Using the BARRA USE3L risk model, we decompose holdings into three components: 1) a market portfolio based on the value portfolio's market beta multiplied by the Russell 1000 index, 2) an industry portfolio, based on the value portfolio's beta to each industry multiplied by the respective cap-weighted industry portfolio, and 3) the residual portfolio that is free of market and industry biases. The Sharpe ratios (on the right) use the long/short value portfolio compared to the residual portfolio that is free of market and industry biases. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures. Please read important disclosures at the end of this article.

Sources: AQR, Russell, MSCI BARRA

even odder, part from currency bets. Investors may or may not desire all those bets, but they certainly may desire them at different levels of risk/confidence than what falls out of this simple ranking approach. We believe a better approach to building style portfolios separates and manages these risks independently. This approach may improve the comparability of stocks by controlling for differences across industries and countries and allows the manager to better allocate and control risk to each source of return (i.e., capture and target risk to each return source *independently*).

Volatility Targeting

Typically, investors understand that they must periodically rebalance their portfolios to maintain a strategic asset allocation. Yet, many investors commonly let their portfolio volatility fluctuate with market volatility, which can result in large time-varying portfolio risk in which returns tend to be dominated by certain periods of higher volatility. Volatility (or risk) targeting is an approach that seeks to yield more consistent risk-taking over time, by adjusting nominal position sizes dynamically in response to these portfolio volatility changes. Such an approach may lead to more stable portfolios and better diversification

across time periods. Similarly, volatility targeting each style within a multistyle portfolio can lead to better diversification at all times. As a result, volatility targeting is a technique that practitioners may use when constructing long/short style portfolios. In long-only portfolios, the analogous concept is targeting tracking error.

Of course, targeting a more consistent level of volatility only makes sense if investors believe that it is forecastable (one needs forecastable volatility in order to target it) and that the trading costs of doing so do not eliminate the benefits. To examine whether this approach is realistic and beneficial, Exhibit 4 looks at the realized volatility over time for two hypothetical value portfolios—a long/short value portfolio in which the volatility is not controlled and one in which the portfolio is scaled to target a fixed volatility. The volatility-targeted portfolio has realized volatility that is more stable and closer to the target volatility over time.

It is also worth noting that there could be a case for allowing volatility to fluctuate, for example, if higher volatility is associated with higher Sharpe ratios. But without a view on attractiveness, taking a consistent level of risk (i.e., being diversified across time) seems to be a defensible approach.

Exhibit 4 Rolling Three-Year Volatilities of Different B/P Portfolios: U.S. Stocks Long/Short, January 1990–December 2015

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR manages. Volatilities above are computed on the three-year rolling returns of long/short B/P portfolios. Value portfolios are formed every month by ranking all U.S. stocks in the Russell 1000 universe on B/P. The long/short portfolio is formed by going long the top half (cheap) and short the bottom half (expensive) of stocks; stocks are weighted by value signal strength. Both volatility-adjusted and nonvolatility-adjusted portfolios are designed to be market- and industry-neutral ex ante. The volatility-adjusted portfolio targets 5% volatility at each rebalance; this level was chosen because it was the average volatility of the B/P portfolio. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures. Please read important disclosures at the end of this article.

Sources: AQR, Russell.

Integrating Styles in a Multistyle Portfolio

Similar to weighting stocks in a single-style portfolio, there are also different ways to weight styles in a multistyle portfolio. There are two popular approaches that are often considered as potential starting points for investors: the "portfolio mix" that builds a multistyle portfolio by investing in separate, stand-alone style portfolios (effectively an "a la carte" approach to style investing) and the "integrated approach" that endogenously integrates styles directly in the portfolio construction process. The portfolio mix approach may be used by individual investors building their own portfolios (perhaps diversifying across managers) or by managers building a multistyle portfolio for them.

To understand how these approaches can differ, consider two portfolios that aim to take advantage of two historically lowly correlated styles—value and momentum: The portfolio mix first picks the "best" value stocks and the "best" momentum stocks and then combines them to form a portfolio; in contrast, the integrated approach first blends each stock's value and momentum scores and then forms a portfolio with the stocks that look best in combination (i.e., it considers the exposure to both styles before investing). The two approaches can actually end up holding very different stocks (especially when investing in lowly correlated styles, such as value and momentum). To see why this may be the case, the top half of Exhibit 5 shows a simple example of how 16 representative stocks (represented by the boxes) map to different portfolios based on their value and momentum scores or "grades."

Focusing first on the top left long-only example, we see that stocks that score highly based on value and momentum make it into both the mix and the integrated portfolios, but those that look excellent on one style (e.g., "A" on value) and terrible on the other (e.g., "F" on momentum) are those that make it into the portfolio mix only. Notably, the integrated approach avoids these stocks (e.g., "A/F" value/momentum combinations) and instead focuses on stocks that look "good" on *both* value and momentum (e.g., "B/B" combinations).

Now, consider how the same 16 stocks map to different long/short portfolios (top right of Exhibit 5). The ability to go short means that we can now take advantage of stocks that are unattractive based on value and momentum.[27] For both the portfolio mix and integrated portfolios, stocks that look excellent (terrible) on value and momentum make it into the long (short) side, while the integrated portfolio also focuses on stocks that look good (bad) in combination for the long (short) sides. There is one subtle nuance with the long/short approach that is worth highlighting: The A/F stock does not actually make it into the mix portfolio (as it did for long-only); this is due to the offsetting nature between the A stock long and the F stock short, which can be expressed in the long/short example.[28]

It turns out that these kinds of distinctions have implications when it comes to returns. Stocks that look excellent on one style but horrible on the other (e.g., A/F stocks) have mediocre expected returns, while those that look good on each style (e.g., B/B stocks) have higher expected returns. The long-only integrated portfolio correctly avoids the former types of stocks and instead focuses on the latter, underscoring the importance of an integrated approach for long-only investing (a point that is discussed at length in Fitzgibbons et al. [2016]). However, integration is also important for long/short portfolios because the types of stocks that make it into the mix still tend to underperform those in the integrated. The bottom half of Exhibit 5 examines the performance differences between various groupings of stocks. It shows that the stocks that are unique to the integrated portfolio tend to outperform those that are unique to the portfolio mix—for *both* long-only and long/short portfolios.

Exhibit 5 Comparing Portfolio Mix and Integrated Approaches: Long-Only and Long/Short

Comparing Portfolio Mix and Integrated Approaches: Long-Only and Long/Short

Long-Only Portfolio Mapping
Hypothetical Example of 16 Stocks

Momentum "Grade" \ Value "Grade"	A	B	C	F
A	Both	Both	Portfolio Mix Only	Portfolio Mix Only
B	Both	Integrated Only		
C	Portfolio Mix Only			
F	Portfolio Mix Only			

Long/Short Portfolio Mapping
Hypothetical Example of 16 Stocks

Momentum "Grade" \ Value "Grade"	A	B	C	F
A	Both	Both	Portfolio Mix Only	
B	Both	Integrated Only		Portfolio Mix Only
C	Portfolio Mix Only		Integrated Only	Both
F		Portfolio Mix Only	Bothgv	Both

Stocks not in any Style Portfolio

Hypothetical Performance of Various Groups of Stocks
U.S. Stocks, January 1990–December 2015

	Long-Only			Long/Short		
	Stocks in Both	**Stocks in Integrated Only**	**Stocks in Portfolio Mix Only**	**Stocks in Both**	**Stocks in Integrated Only**	**Stocks in Portfolio Mix Only**
Alpha*	1.9%	1.6%	0.7%	6.5%	5.4%	4.2%
t-statistic	2.88	2.11	0.73	3.63	2.94	2.15
Example Stocks Long: Val/Mom "Grade"	A/A	B/B	A/C & A/F	A/A	B/B	A/C
Example Stocks Short: Val/Mom "Grade"	-	-	-	F/F	C/C	B/F

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR manages. The long-only "portfolio mix" is constructed by equally weighting stand-alone value (B/P) and momentum (past 12-month price return, excluding the recent month), each with a 15% cutoff based on the Russell 1000 universe. The long-only "integrated" portfolio is constructed by first combining value and momentum scores and then selecting a cutoff so that it has the same tracking error as the portfolio mix. The long/short portfolio is constructed using a similar approach, with the portfolio mix using a 15% cutoff, and the cutoff for the integrated portfolio determined so that it has the same leverage and volatility as the portfolio mix. These simple portfolios are neither market- nor industry-neutral. Returns are gross of transaction costs. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures. Please read important disclosures at the end of this article.

*Alpha is measured relative to an equally weighted composite of Russell 1000 stocks. The stocks in each respective group are also equally weighted.

Sources: AQR, Russell.

Although we have focused on performance differences, an additional benefit of an integrated approach is the potential reduction in transaction and other costs. By forming a portfolio of multiple styles in an integrated way, a manager can net positions and trades, which may allow for a reduction in the portfolio's turnover.[29] However, the trade-off is that allocation and attribution may not be as straightforward.[30]

Strategic or Tactical

A bigger-picture question regarding investing in styles is whether to tactically time or to simply maintain strategic allocations. Tactical timing means owning more (less) of a particular style when its expected return is higher (lower) than normal. Practitioners use a variety of ways to determine whether a style is conditionally attractive; for instance, timing may be based on valuation metrics (increasing/decreasing the weight of a style when the spreads in valuation between the long and short sides are wide/tight), momentum measures (increasing/decreasing the weight of a style when recent performance has been better/worse), or macroeconomic conditions (identifying the best market environments for each style). Although several studies have looked at the efficacy of tactical timing (whether for markets or styles, based on a variety of measures), the evidence shows that timing is very difficult in practice.[31] Valuation timing tends to have slightly more predictive power for long-horizon factor returns, particularly for slower-turnover factors (such as the market itself and, to a lesser extent, the value factor) versus faster ones.[32] Still, the evidence shows that the benefit of timing strategies has been weak historically, and some tests of the long-term power of timing may even be exaggerated and/or inapplicable.[33]

Specifically, for factors the evidence on timing is further complicated by the fact that timing styles based on valuation is highly correlated with having direct value factor exposure in the portfolio. As such, value timing may not add much to a portfolio that already has a strategic allocation to value. These results are addressed in Asness et al. [2017], who show that the bar is raised further for contrarian style rotation when the relevant benchmark is a strategically diversified multistyle strategy with direct value exposure. Thus, contrarian tilting across style premia is potentially even harder to do successfully than contrarian market timing.

To further understand the high bar to timing styles, we ask the question of how good (or skilled) an investor would need to be at timing to justify tactical tilts. We do so relative to a diversified baseline portfolio—an equally weighted 50/50 combination of value and momentum[34]—and for a timing strategy that is based on valuation metrics (i.e., one that is correlated with strategic value exposure).[35] For simplicity, we ignore the (weak) empirical evidence on the efficacy of value timing and instead define "skill" by assumed stand-alone timing strategy Sharpe ratios. The results in Exhibit 6 show that the amount of timing (or how large the tactical tilt should be) varies as a function of skill: Modest timing skills merit modest tilts. Importantly, timing skills need to be better than a 0.1 Sharpe strategy to merit any weight at all; that is, when the timing strategy has a 0.1 Sharpe ratio, it may be better to stick with the strategic value/momentum weights. But even when timing skills are very large (the standalone timing strategy Sharpe ratio is 0.5), the results show that

Exhibit 6 Size of Optimal Tactical Tilt as a Function of Skill: Given a Strategic Baseline Portfolio with a Sharpe Ratio of 1.0

Notes: This exhibit is for illustrative purposes only. The baseline portfolio is a simple equal-weighted 50/50 combination of value and momentum, assuming the two styles are lowly correlated (−0.5 correlation). For value and momentum, we assume arithmetic Sharpe ratios of 0.5 and volatilities of 10%, such that the combination of value and momentum (the strategic baseline portfolio) has a Sharpe of 1.0. We then evaluate the inclusion of a timing strategy with a correlation of +0.7 with value and −0.6 with momentum, and solve for optimal risk allocation to the timing strategy at different Sharpe ratios. That is, given a Sharpe assumption for a timing strategy, how much risk would you want to take in it; we then translate that risk to a capital tilt. Transaction costs will likely further penalize tactical strategies. Please read important disclosures in the endnotes section. The Sharpe ratios and volatility assumptions used are meant to be illustrative and are not based on any actual underlying data. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures.

Sources: AQR.

the size of the tilt should *still* be modest, at +/−12%. More extreme tactical tilts require higher levels of skill to justify the concentration. Timing skills must be good enough to overcome not only the hurdle of a good starting point (such as a well-diversified portfolio of multiple styles) but also the transaction costs of doing so as tactical allocators often face higher turnover than their strategic counterparts. Finally, it's important to note that if the timing strategy were lowly correlated with the baseline portfolio,[36] there could be a case for higher tilts. Similarly, a lower baseline strategic Sharpe ratio could also merit higher tilts.[37]

The decision of whether to time styles (and by how much!) should ultimately depend on how skilled an investor is at timing, combined with what is already in his or her portfolio. We believe that strategic diversification generally beats extreme tactical tilts (i.e., tactical concentration) and that if investors try to time factors (or the market), they should do so on the margin (i.e., "sin a little").[38] The positive craftsmanship choice here is resisting the temptation to sin a lot.

In this section, we have focused on design decisions and choices that managers can make when constructing style portfolios; each choice has the opportunity to add value or subtract if done poorly, but it may also require sophisticated portfolio construction techniques. These individual choices may result in more modest Sharpe ratio improvements than, say, the large gains to diversifying across factors. Although these improvements may seem incremental, they are still worth pursuing because every little bit counts. At the end of the day, it is up to the manager to decide how to construct her style portfolio; we aim

to highlight how various craftsmanship choices can affect exposures and performance of style portfolios.

How to Execute Style Portfolios

While portfolio construction decisions may allow managers to more efficiently harvest the underlying risk premia through enhanced signals or better weighting schemes, it is important to ensure that trading and risk management decisions do not erode the value added from other steps of the process.

Portfolio Implementation

The previous section outlined some choices managers need to make when designing their theoretical style portfolios. So far, the discussion has focused on portfolio construction in a world without real-world costs, such as the cost to trade toward that theoretical, ideal portfolio. Effective portfolio implementation is about achieving high returns net of those costs. A simplistic implementation is to build theoretical portfolios and trade directly to them, without weighing the expected benefit of trading to the ideal portfolio against the expected trading cost of doing so. Style investing requires active portfolio rebalancing; what was considered cheap yesterday may no longer be trading cheap today. Because of this dynamic nature of style investing, it's very important to think about execution and implementation.

To illustrate this point: consider a style portfolio that rebalanced at different frequencies. A portfolio that is rebalanced daily may result in the "freshest" portfolio (with the highest correlation with the ideal portfolio) and the highest gross-of-transaction-costs performance; while a portfolio that is rebalanced annually will likely be "stale" (further away from the ideal portfolio) and therefore have worse gross performance. But these comparisons don't take transaction costs into account.[39] A better point of comparison is to focus on net-of-transaction-costs performance. Obviously the daily rebalanced portfolio will incur high transaction costs, thereby reducing the gross performance benefits. And the annual rebalanced portfolio will have much lower transaction costs but will be applied against lower gross returns.

Ultimately, the transaction cost savings of rebalancing at a lower frequency needs to be weighed against the performance degradation of trading a more stale portfolio (i.e., there is a trade-off between the cost of trading versus the opportunity cost of not trading). Exhibit 7 looks at a simple long/short momentum portfolio and shows the impact that different rebalancing frequencies can have on gross and net performance.

We focus on momentum because of its higher natural turnover, relative to other styles. As a result of its higher turnover, rebalance decisions may have a larger impact for momentum portfolios.

Another way to potentially reduce transaction costs is to allow for some deviation from the ideal (or "freshest") portfolio through an optimization. For example, if a manager has decided to rebalance daily, he may still reduce trading costs by allowing some deviation to the ideal portfolio. It turns out that the effect of varying the deviation is similar to that of changing the rebalance frequency: A higher deviation induces greater style drift and therefore greater performance degradation, but also less turnover and, thus, lower transaction costs.[40]

Exhibit 7 Performance at Different Rebalance Frequencies for Hypothetical Price Momentum Portfolios: U.S. Stocks Long/Short, January 1990–December 2015

Notes: This exhibit is for illustrative purposes only and is not representative of an actual portfolio AQR managed. We vary the rebalance frequency on a long/short momentum portfolio. Returns are shown gross of cash. The momentum portfolio is formed at various frequencies, as shown along the x-axis, by ranking all U.S. stocks in the Russell 1000 universe on past 12-month price returns, excluding the most recent month. A long/short portfolio is formed by going long the top half (outperformers) and short the bottom half (underperformers); stocks are weighted by momentum-signal strength. Reported are returns gross and net of transaction costs. All numbers are also net of cash but gross of financing costs. Transaction costs are calculated using a propriety trading cost model. Turnover is annualized and quoted as total buys and sells over total short-side notional. Please read important disclosures in the endnotes section. Hypothetical performance data have certain inherent limitations, some of which are discussed in disclosures.

*The "ideal" portfolio reflects the model based on past 12-month price returns, excluding the most recent month, rebalanced daily.

Sources: AQR, Russell.

This exercise is not meant to suggest that there is one magic turnover level that works for all style portfolios. The behavior of style portfolios constructed by different managers can vary greatly (we've already shown several design decisions in this article that can lead to vastly different portfolios); as a result, it may not be informative to compare turnover or transaction costs in isolation across different products or managers. Low, or high, turnover by itself should not be seen as a virtue. Ultimately, net returns, judged over a long-enough sample period, provide a better way to compare different strategies.

Cost-effective Execution

We have discussed the trade-off between expected returns and rebalance frequency and how transaction costs play a crucial role in that determination. But it's also important that once

the rebalance frequency or turnover is set, managers utilize smart trading techniques to minimize costs incurred per dollar traded. A big portion of trading costs for large investors is price impact[41]—that is, how different the execution price is relative to the price when the manager enters the market. Price impact often dwarfs explicit costs like commissions. Interestingly, however, many managers and investors focus on the explicit, rather than the implicit, trading costs.

One way that managers may reduce their market impact is by being patient. Using real-world execution data, Frazzini et al. [2012] observe that the greater the participation rate (shares traded/market trading volume), the costlier it is to trade. Exhibit 8, extracted from their paper, shows that trading 2% of a stock's daily volume results in roughly 17 basis points (bps) of market impact; in contrast, trading 6% costs roughly 28 bps per dollar traded, on average. One implication is that if managers want to trade 6% of a stock's daily volume, they need to decide how best to do so. In the case of style investing, trades are generally based on slower signals (rather than higher-frequency ones), so managers may choose a patient trading approach. Rather than completing the trades in one day, they may spread them out across three days (utilizing roughly 2% of trading volume per day). A patient trading approach such as this may reduce trading costs.[42] A similar logic can apply to intraday trading: Spreading out trades over the course of a day, rather than trading in a short period (e.g., around market close), will likely lead to lower market impact.

Exhibit 8 Market Impact by Fraction of Trading Volume, August 1998– September 2013

Notes: This exhibit is for illustrative purposes only. The average market impact is plotted for actual live trades from AQR's proprietary database, as in Frazzini et al. [2012]. The authors sort all trades in their dataset into 30 bins based on their fraction of daily volume and compute average market impact for each bucket. This includes all available developed market equity transactions (cash equities and equity swaps) in the authors' data between August 1998 and September 2013. See Exhibit A1 in the appendix for more detail. Market impact is in basis points (annualized) and is defined as the difference between trade-weighted average execution price and the price when the manager enters the market. Please read important disclosures in the endnotes section

Sources: Frazzini et al. [2012]; the data through 2013 are an extension of data from 2011 using the same methodology.

Risk Management

Another important aspect of running a style portfolio is being intentional in the type of risks that the portfolio takes on and managing those risks through time. We have already touched on hedging out unintended risks (such as market or even industry risk) and volatility targeting, but there are other dimensions of risk that investors need to be mindful of—including leverage, illiquidity, solvency (i.e., adequate free cash levels), left-tail risk, and correlation risk (i.e., styles becoming more correlated). For long/short portfolios, leverage is an important risk dimension that can be managed by varying exposures as volatilities move around (e.g., reducing leverage when volatility increases), limiting leverage at some absolute level, trading liquid instruments, and maintaining comfortable levels of cash to support that leverage.

Even with thoughtful risk management, there can be painful times for style portfolios. Having a prespecified plan for how to handle a crisis is of paramount importance, especially for long/short levered portfolios. One alternative is doing nothing, but it is unlikely a manager can actually stick to that in all circumstances. The reality is that in a crisis, risk aversion tends to increase; eventually, risk appetite and risk levels diverge enough such that they have to be brought back in line, which is typically accomplished via deleveraging. This behavior often means that investors capitulate at the bottom or worst possible time and then are averse to putting risk back on.

To help avoid this type of situation, investors can benefit from having a systematic plan in place for handling a crisis. This type of plan begins cutting risk when the portfolio experiences a drawdown, or if the portfolio's short-term tail risk goes up. It then systematically adds back risk as the returns improve and the left-tail risk subsides.[43] In other words, mechanical drawdown rules may dial down risk if prespecified loss levels are reached (and dial risk exposures back up when performance recovers). A prespecified, planned drawdown control system may be beneficial in times of panic—underscoring the importance of having a plan before you need one. Such an approach may allow investors to maintain diversification and stay invested in tough times.

Conclusion

Throughout this article, we discussed many of the craftsmanship choices that can be made when constructing style portfolios. Although there may be broad agreement on the major styles that drive asset returns, we have shown that when it comes to style investing, many details matter—from how to transform signals into portfolio weights, to risk control, optimization, and trading.

While there may not always be a clear right or wrong with some of these design decisions, managers should be able to defend their choices and understand their implications. We believe design choices that are based on economic principles and empirical evidence found across different markets, time periods, and asset classes should lead to better investment outcomes—even if not in every period (over the short-term, randomness can trump an ex ante edge!). Ultimately, what may seem like inconsequential decisions can lead to a meaningful edge over time.

Appendix

Exhibit A1 Trade Execution Data from Exhibit 8, Summary Statistics

Amount Traded (Billion USD)							
Year	**Total**	**By Region**		**By Size**		**By Portfolio Type**	
		U.S.	**International**	**Large Cap**	**Small Cap**	**Long/Short**	**Long-Only**
1998*	2.96	1.29	1.67	2.96		2.96	
1999	5.29	1.99	3.30	5.29		5.29	
2000	1.99	0.76	1.23	1.99		1.86	0.13
2001	1.08	0.55	0.53	1.08		1.00	0.08
2002	4.21	0.71	3.50	4.21	0.00	1.40	2.81
2003	5.43	2.69	2.75	5.43	0.00	4.17	1.26
2004	10.00	2.95	7.05	9.99	0.01	6.38	3.62
2005	16.16	8.06	8.10	15.75	0.41	11.45	4.71
2006	67.01	34.79	32.22	64.23	2.78	44.69	22.31
2007	129.46	50.70	78.76	125.21	4.25	96.65	32.81
2008	108.29	25.06	83.24	104.27	4.02	69.30	38.99
2009	111.12	18.58	92.54	108.12	2.99	85.50	25.62
2010	117.17	29.15	88.02	113.78	3.38	91.94	25.23
2011	146.50	56.62	89.88	141.93	4.58	115.69	30.81
2012	179.09	121.39	57.70	173.41	5.68	141.97	37.13
2013**	141.18	92.87	48.31	136.04	5.14	95.21	45.98
Total	**1,046.94**	**448.15**	**598.79**	**1,013.69**	**33.25**	**775.46**	**271.48**

Notes: This exhibit is for illustrative purposes only. This table shows summary statistics of the trade execution database used in Exhibit 8. See Frazzini et al. [2012] for more detail

*Indicates partial year from August 31, 1998.

**Indicates partial year up to September 30, 2013.

Source: Frazzini et al. [2012]; the data through 2013 are an extension of data from 2011 using the same methodology.

Endnotes

The views and opinions expressed are those of the authors and do not necessarily reflect the views of AQR Capital Management, its affiliates, or its employees; do not constitute an offer or solicitation of an offer, or any advice or recommendation, to purchase any securities or other financial instruments; and may not be construed as such.

We thank Ryan Wei for providing research assistance and Gregor Andrade, Cliff Asness, Andrea Frazzini, April Frieda, Antti Ilmanen, Tobias Moskowitz, and Lasse Pedersen for helpful comments and suggestions.

[1]See Fama and French [1993], Ilmanen [2011], Asness et al. [2013], and Asness, Ilmanen, Israel, and Moskowitz [2015]. Past performance is not a guarantee of future performance.

[2]Some of the challenges related to the size anomaly relate to its lack of pervasiveness (only applicable in equities), the limited economic intuition behind its efficacy, and its weak in-sample evidence. Even though Asness, Frazzini, Israel, Moskowitz, and Pedersen [2015] show that the size premium may be resurrected when you control for quality, we believe the size anomaly may be

better thought of as a subset of illiquidity, which is more robust, more pervasive, and accompanied by better theory (but comes at a cost—it is less liquid, so more costly to trade).

[3]In the case of value in equities, the portfolio is typically constructed by tilting a market portfolio according to some fundamental measure. For example, a manager might overweight high book-to-price (B/P) stocks and underweight low B/P stocks. The resulting portfolio will have both market and "pure" style exposure. It is the long-only equivalent of Fama and French's famous HML factor, the return spread between a diversified portfolio of high B/P and low B/P stocks.

[4]See Arnott et al. [2004]; Asness and Liew [2014]; Ilmanen et al. [2014]; and Asness, Frazzini, Israel, and Moskowitz [2015] for more on this.

[5]For more on how both the long and short sides contribute to sytle premia returns, see Israel and Moskowitz [2012].

[6]For more on long-only versus long/short style investing, see Ilmanen et al. [2014]. Although we do not get into "130/30" (also called relaxed constraint or active extension) implementation in any detail here, it's worth mentioning that it can be considered a hybrid between long-only and long/short approaches.

[7]We do see a stand-alone role for some risk-reducing styles (such as defensive equity or trend-following) in which investors focus on downside protection for their total portfolio.

[8]See Novy-Marx [2012, 2013]; Asness, Moskowitz, and Pedersen [2013]; Asness, Ilmanen, Israel, and Moskowitz [2015]; and Frazzini et al. [2013]. Diversification does not eliminate the risk of experiencing investment losses.

[9]To fully embrace this multi-asset approach, however, investors need to adopt the long/short framework. A long-only style portfolio wouldn't apply to certain asset classes, such as currencies, without a real benchmark. For more on multi-asset style investing, see Asness, Ilmanen, Israel, and Moskowitz [2015].

[10]See Penman et al. [2006] and Duncombe et al. [2016].

[11]Although such an assumption may be unrealistic, Asness and Frazzini [2013] show that moves in book value are generally smaller than market moves.

[12]See Fama and French [1992].

[13]See Frazzini et al. [2013] and Asness, Frazzini, Israel, and Moskowitz. [2015].

[14]Here we use B/P, earnings/price, forecasted earnings/price, cash flow/enterprise value (an adjusted measure of price), and sales/enterprise value.

[15]B/P was the worst performer within the multiple-measure value composite; however, the composite still outperformed three out of the five stand-alone measures on a risk-adjusted basis over this period.

[16]Hypothetical performance data have certain inherent limitations, some of which are discussed in the disclosures.

[17]Note that these results are gross of transaction costs.

[18]Fama and French utilize lagged prices and also construct their portfolios over the entire CRSP universe (large- and small-cap stocks, including many micro-cap stocks). For the remainder of this article, we will utilize current (rather than lagged) prices, as done under the HML Devil approach, as discussed earlier. We will also use the investable universe of U.S. large-cap stocks, rather than the entire CRSP universe. For simplicity, we will refer to this portfolio as B/P.

[19]See Israel and Ross [2017] for more on how HML's market exposure varies over time.

[20]For a long-only portfolio, the analogous concept is to target a beta of 1 to the benchmark; such an approach ensures that active returns are not driven by the market.

[21]Note that industries are just one way to define peers; an additional craftsmanship refinement may be to define peer groups through additional economic and statistical linkages, for example.

[22]For momentum, there is more evidence that it is effective across industries. See Moskowitz and Grinblatt [1999] and Asness, Porter, and Stevens [2000].

[23]Managers can also seek to mitigate these risks by explicitly hedging them, separating comparisons of stocks within industries from comparisons across industries or diversifying across multistyle portfolios. For instance, in the case of industry exposure, combining value and momentum may provide offsetting industry exposure. Think of the tech bubble—during this period, a value portfolio would be underweight "expensive" tech stocks, while a momentum portfolio would be overweight "increasing-in-price" tech stocks. The combination of these two styles may help mitigate industry exposure.

[24]Specifically, portfolios that are long high within-industry B/P stocks and short low within-industry B/P stocks should have about the same to slightly higher expected return, but less variance, than portfolios long high market-wide B/P stocks and short low market-wide B/P stocks (Asness, Porter, and Stevens, [2000]).

[25]In general, when portfolio adjustments result in higher turnover, investors need to consider transaction costs and taxes. The higher costs associated with additional turnover may be mitigated through efficient trading (more on this to come). The tax efficiency of a portfolio is a complicated function of turnover, short- and long-term gain and loss realizations, and dividend income. See Israel and Moskowitz [2012] and Sialm and Sosner [2017] for more information on the after-tax performance, tax exposure, and tax efficiency of equity styles.

[26]See Aghassi et al. [2011].

[27]Note that a long-only portfolio can take advantage of unattractive stocks by holding them at a weight less than the benchmark (i.e., underweight these stocks). However, for simplicity, we choose to represent pure longs and shorts here.

[28]Such a stock actually "makes it" into both the long and short side, but it would not appear in the implemented portfolio because of position netting. If the long/short value and momentum portfolios were run separately, it would indeed appear in the long portfolio of one and the short portfolio of the other.

[29]There are also potential tax benefits to an integrated approach. See Israel and Moskowitz [2012] for more information on the tax efficiency of stand-alone and integrated equity styles.

[30]For an integrated approach, allocating to styles is done in the portfolio formation process, and therefore the attribution process is not as straightforward as allocating to stand-alone style portfolios. See Fitzgibbons et al. [2016] for more on the benefits of an integrated approach.

[31]Asness, Ilmanen, and Maloney [2017] focus on valuation-based market timing and show that correlation evidence (in sample and over long horizons) makes contrarian market timing look promising, yet when an actual contrarian trading rule is applied, the performance improvement is weak.

[32]There may also be a case for timing at extremes: when the spread between "cheap" and "expensive" assets is particularly wide. One period in which value timing was at extremes was during the tech bubble (1999–2000). While a value investor was ultimately rewarded handily after the tech bubble burst, she would have lost a lot before making anything, highlighting another example of how value timing may be difficult in practice.

[33]Timing strategy backtests can be vulnerable to look-ahead biases. In particular, the use of in-sample spreads may overfit the past and underdeliver in the future (see Asness, Ilmanen, and Maloney [2017]); and attempts to increase the sample size by using overlapping data or performing multiple regressions over different horizons do not improve inference about long-horizon return predictability (see Boudoukh et al. [2017]).

[34]We assume equal (50/50) capital allocation to each style, where each style has stand-alone volatilities of 10% and expected Sharpe ratios of 0.5. The equally weighted combination has a Sharpe ratio of around 1.0, due to diversification. These styles are diversifying to each other as they are –0.5 correlated. The 1.0 strategic baseline Sharpe ratio is consistent with long-run empirical evidence on an equally weighted value (HML) and momentum (UMD) portfolio.

[35]Historically, the contemporaneous correlations between the returns to value timing (difference between the value-timed multistyle portfolio and the nontimed multi-style portfolio) and returns to value and momentum are +0.7 and –0.6, respectively (see Asness et al. [2017]).

[36]For example, if an investor does not have value exposure in their strategic portfolio, then value timing may be potentially additive. However, compared to an explicit risk-targeted strategic allocation to value, value timing provides an intermittent and suboptimal amount of value exposure; see Asness et al. [2017].

[37]In our example, we have assumed a baseline portfolio that has a Sharpe of 1.0. If, however, the baseline portfolio had a lower Sharpe ratio, the results would imply higher tactical tilts.

[38]See Asness et al. [2017] and Asness, Ilmanen, and Maloney [2017].

[39]See Israel and Moskowitz [2012] for more information on taxes and style investing and Frazzini, Israel, and Moskowitz [2012] for more on trading costs.

[40]Keim [1999] looks at a passive small-cap equity strategy and shows how design choices that allow for some deviation relative to an underlying index and smarter trading techniques can improve net-of-transaction-cost returns.

[41]See Frazzini et al. [2012] for additional information on explicit (commissions, bid–ask spread) and implicit (price impact) trading costs.

[42]In our example, the trading cost savings could be 11 bps (28 bps – 17 bps), but this is likely an overestimate because we cannot assume the cost functions over three days are independent. That is, trading 2% on the last day might be costlier than trading 2% on the first day, because you've been trading in the same direction for consecutive days. However, this estimate of reduced trading costs should be directionally correct.

[43]Our view is that such a plan is likely unnecessary in long-only implementations but becomes more helpful in long/short portfolios with leverage.

References

Aghassi, M., C. Asness, O. Kurbanov, and L. Nielsen. "Avoiding Unintended Country Bets in Global Equity Portfolios." AQR White Paper, 2011.

Arnott, R., J. Hsu, and P. Moore. "Fundamental Indexation." *Financial Analysts Journal,* Vol. 61, No. 2 (2004), pp. 83–99.

Asness, C., S. Chandra, A. Ilmanen, and R. Israel. "Contrarian Factor Timing Is Deceptively Difficult." *The Journal of Portfolio Management,* Vol. 43, No. 5 (2017), pp. 72–87.

Asness, C., and A. Frazzini. "The Devil in HML's Details." *The Journal of Portfolio Management,* Vol. 39, No. 4 (2013), pp. 49–68.

Asness, C., A. Frazzini, R. Israel, and T. Moskowitz. "Fact, Fiction, and Value Investing." *The Journal of Portfolio Management,* Vol. 42, No. 1 (2015), pp. 34–52.

Asness, C., A. Frazzini, R. Israel, T. Moskowitz, and L. Pedersen. "Size Matters, If You Control Your Junk." Forthcoming *Journal of Financial Economics,* 2018.

Asness, C., A. Ilmanen, R. Israel, and T. Moskowitz. "Investing with Style." *Journal of Investment Management,* Vol. 13, No. 1 (2015), pp. 27–63.

Asness, C., A. Ilmanen, and T. Maloney. "Market Timing: Sin a Little." *Journal of Investment Management,* Vol. 15, No. 3 (2017), pp. 23–40.

Asness, C., and J. Liew. "Smart Beta: Not New, Not Beta, Still Awesome." *Institutional Investor,* 2014.

Asness, C., T. Moskowitz, and L. Pedersen. "Value and Momentum Everywhere." *The Journal of Finance,* Vol. 68, No. 3 (2013), pp. 929–985.

Asness, C., R. Porter, and R. Stevens. "Predicting Stock Returns Using Industry-Relative Firm Characteristics." AQR White Paper, 2000.

Boudoukh, J., R. Israel, and M. Richardson. "Long-Horizon Predictability: A Cautionary Note." AQR White Paper, 2017.

Duncombe, G., J. Saret, and L. Sholder. "Risk Factors Are Not Generic." Two Sigma White Paper, 2016.

Fama, E.F., and K.R. French. "The Cross-Section of Expected Stock Returns." *The Journal of Finance,* Vol. 47, No. 2 (1992), pp. 427–465.

_____. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics,* Vol. 33, No. 1 (1993), pp. 3–56.

Fitzgibbons, S., J. Friedman, L. Pomorski, and L. Serban. "Long-Only Style Investing: Don't Just Mix, Integrate." AQR White Paper, 2016.

Frazzini, A., R. Israel, and T. Moskowitz. "Trading Costs of Asset Pricing Anomalies." Working paper, 2012, SSRN 2294498.

Frazzini, A., R. Israel, T. Moskowitz, and R. Novy-Marx. "A New Core Equity Paradigm." AQR White Paper, 2013.

Ilmanen, A. *Expected Returns.* Chichester, UK: Wiley, 2011.

Ilmanen, A., R. Israel, and D. Villalon. "Style Investing: The Long and Long/Short of It." *Investments and Pensions Europe,* 2014.

Israel, R., and T. Moskowitz. "How Tax Efficient Are Equity Styles?" Working paper, 2012, SSRN 2089459.

_____. "The Role of Shorting, Firm Size, and Time on Market Anomalies." *Journal of Financial Economics,* Vol. 108, No. 2 (2013), pp. 275–301.

Israel, R., and A. Ross. "Measuring Factor Exposures: Uses and Abuses." *The Journal of Alternative Investments,* Vol. 20, No. 1 (2017), pp. 10–25.

Keim, D.B. "An Analysis of Mutual Fund Design: The Case of Investing in Small-Cap Stocks." *Journal of Financial Economics,* Vol. 51, No. 2 (1999), pp. 173–194.

Moskowitz, T., and M. Grinblatt. "Do Industries Explain Momentum?" *The Journal of Finance,* Vol. 54, No. 4 (1999), pp. 1249–1290.

Novy-Marx, R. "The Quality Dimension of Value Investing." Working paper, 2012.

_____. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics,* Vol. 108, No. 1 (2013), pp. 1–28.

Penman, S., S. Richardson, and I. Tuna. "The Book-to-Price Effect in Stock Returns: Accounting for Leverage." *Journal of Accounting Research,* Vol. 45, No. 2 (2006), pp. 427–468.

Sialm, C., and N. Sosner. "Taxes, Shorting, and Active Management." *Financial Analysts Journal,* Vol. 74, No. 1 (2017).

Contributors

Clifford S. Asness, PhD
Managing and Founding Principal, AQR Capital Management

John C. Bogle
Founder, The Vanguard Group, Inc.
President of the Bogle Financial Markets Research Center

Andrea Frazzini, PhD
Principal, AQR Capital Management

Antti Ilmanen, PhD
Principal, AQR Capital Management

Ronen Israel
Principal, AQR Capital Management

Sarah Jiang
Managing Director, AQR Capital Management

David G. Kabiller, CFA
Founding Principal, AQR Capital Management

Robert Krail, PhD
Founding Principal, AQR Capital Management (retired)

John M. Liew, PhD
Founding Principal, AQR Capital Management

Mark L. Mitchell, PhD
Co-Founder and Principal, CNH Partners

Tobias J. Moskowitz, PhD
Principal, AQR Capital Management
Dean Takahashi Professor of Finance, Yale University

Yao Hua Ooi
Principal, AQR Capital Management

Diogo Palhares
Vice President, AQR Capital Management

Lasse H. Pedersen, PhD
Principal, AQR Capital Management
Professor of Finance, Copenhagen Business School

Todd C. Pulvino, PhD
Co-Founder and Principal, CNH Partners

Scott Richardson, PhD
Principal, AQR Capital Management
Professor of Management Practice in Accounting, London Business School

Adrienne Ross
Vice President, AQR Capital Management

This page intentionally left blank

Over the past 20 years, AQR founders Cliff Asness, David Kabiller, and John Liew have built an investment powerhouse that has published over 200 scholarly articles in peer-reviewed finance and economics journals. Its clients include some of the largest and most sophisticated investors around the globe.

The firm's story began at the University of Chicago's PhD program, where Asness, Liew, and Robert Krail (another AQR co-founder) met and the foundation of AQR's approach to systematic investing was established. Those roots in academia remain at the core of what AQR does today. The firm's culture of intellectual curiosity compels its people to challenge the status quo, disrupt long-held beliefs, and uncover new insights.

Praise for

20 for TWENTY

"Read every chapter, and you'll find acute analysis, fascinating insight, and some of the best writing on finance and investing that you'll ever see."

—**John C. Bogle**
Founder, Vanguard Group

"The top professionals at AQR are mostly former first-rate PhD students and faculty of the Booth School of Business at the University of Chicago. I know them all quite well and respect what they bring to the money management world. This book is an excellent compilation of their data-driven research, applied and academic."

—**Eugene F. Fama**
2013 Nobel Laureate in Economic Sciences and Robert R. McCormick Distinguished Service Professor of Finance, The University of Chicago Booth School of Business

"There is likely no better endorsement than a foreword by legend, Jack Bogle. Still, only a few firms have had as great an impact on the study of investing and portfolio management as AQR."

—**Frank J. Fabozzi**
Editor, *Journal of Portfolio Management* and Senior Professor and Scientific Advisor, EDHEC-Risk Institute

"This is a valuable collection of papers at the forefront of investment science. AQR's academic engagement, through the Insight Award, the impressive publication record by AQR-affiliated academics, and simply the application of academic research to cost-efficient strategies is pushing the investment community to take a more analytic approach to investing, to the benefit of investors."

—**Kent Daniel**
William von Mueffling Professor of Business, Columbia Business School

ISBN 978-1-5323-8538-4
90000>
9 781532 385384