Journal of Financial Economics 138 (2020) 95–117

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

On the performance of volatility-managed portfolios☆

Scott Cederburg[a], Michael S. O'Doherty[b,*], Feifei Wang[c], Xuemin (Sterling) Yan[d]

[a] Eller College of Management, University of Arizona, McClelland Hall, Room 315R, Tucson, AZ 85721, USA
[b] University of Missouri, Trulaske College of Business, 401C Cornell Hall, Columbia, MO 65211, United States
[c] Farmer School of Business, Miami University, 2064 FSB, Oxford, OH 45056, USA
[d] College of Business, Lehigh University, 471 Rauch Business Center, Bethlehem, PA 18015, USA

ARTICLE INFO

Article history:
Received 19 November 2018
Revised 15 March 2019
Accepted 17 March 2019
Available online 1 May 2020

JEL classification:
G10
G11
G12

Keywords:
Volatility-managed portfolios
Portfolio choice

ABSTRACT

Using a comprehensive set of 103 equity strategies, we analyze the value of volatility-managed portfolios for real-time investors. Volatility-managed portfolios do not systematically outperform their corresponding unmanaged portfolios in direct comparisons. Consistent with Moreira and Muir (2017), volatility-managed portfolios tend to exhibit significantly positive alphas in spanning regressions. However, the trading strategies implied by these regressions are not implementable in real time, and reasonable out-of-sample versions generally earn lower certainty equivalent returns and Sharpe ratios than do simple investments in the original, unmanaged portfolios. This poor out-of-sample performance for volatility-managed portfolios stems primarily from structural instability in the underlying spanning regressions.

1. Introduction

Recent studies show strong empirical performance for volatility-managed versions of popular trading strategies, including the market (Ang, 2014; Moreira and Muir, 2019), momentum (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016), betting-against-beta (Barroso and Maio, 2018), and financial distress (Eisdorfer and Misirli, 2020) factors.[1] These portfolios are characterized by conservative positions in the underlying factors when volatility was recently high and more aggressively levered positions following periods of low volatility. Although each of the papers noted above examines an individual volatility-scaled strategy, Moreira and Muir (2017) find that the empirical success of volatility management is a pervasive phenomenon. They show that volatility-scaled strategies earn systematically positive alphas across a wide range of asset pricing factors, and these alphas imply pronounced increases in Sharpe ratios and large utility gains for mean-variance investors. Taken as a whole, existing studies leave readers with the impression that volatility-managed equity strategies routinely lead to improved performance. The findings have important implications for investors, have

☆ We thank Jessica Wachter (the referee), Andrew Detzel, Chris Lamoureux, Jonathan Lewellen, Hai Lin, Rick Sias, Neal Stoughton, Mitch Towner, and seminar participants at the 2017 Wellington Finance Summit, the 2018 Northern Finance Association, Iowa State University, North Carolina State University, and the University of Arizona for helpful comments and suggestions.

* Corresponding author.
E-mail addresses: cederburg@email.arizona.edu (S. Cederburg), odohertym@missouri.edu (M.S. O'Doherty), wangf10@miamioh.edu (F. Wang), xuy219@lehigh.edu (X. Yan).

[1] Also see related work on volatility-managed currency strategies by Daniel et al. (2017) and Maurer et al. (2018). Grobys et al. (2018) examine the properties of risk-managed industry momentum, and volatility scaling is also used in the time-series momentum literature (e.g., Moskowitz et al., 2012; Baltas and Kosowski, 2017). Kirby and Ostdiek (2012) consider volatility-timing strategies for portfolio allocations across multiple stock portfolios. Much of this recent work follows from Fleming et al. (2001, 2003), who demonstrate large economic benefits from volatility timing for short-horizon investors allocating across several asset classes.

https://doi.org/10.1016/j.jfineco.2020.04.015
0304-405X/

received considerable attention in the financial press, and have also had an impact on industry applications.[2]

In this paper, we assess whether volatility management is systematically advantageous for investors and place specific emphasis on real-time implementation. We contribute to the literature in three primary ways. First, based on a substantially broader sample of 103 equity trading strategies, we find no statistical or economic evidence that volatility-managed portfolios systematically earn higher Sharpe ratios than their unmanaged counterparts do. Second, despite this mixed evidence from direct performance comparisons, we confirm that Moreira and Muir's (2017)) finding of systematically positive spanning regressions alphas for volatility-managed portfolios also holds in our extended sample. The trading strategies implied by the spanning regressions are not implementable in real time, however, as they require investors to combine the volatility-scaled and unscaled versions of a given portfolio using ex post optimal weights (e.g., Gibbons et al., 1989). We explore an array of reasonable out-of-sample versions of these "combination" strategies and find that they typically underperform simple investments in the original, unscaled portfolios. Third, we examine why the in-sample alphas for volatility-managed portfolios do not readily translate into out-of-sample gains for investors. We provide evidence that this result is driven by substantial structural instability in the underlying spanning regressions for these strategies. Overall, our findings suggest a more tempered interpretation of the practical value of volatility-managed portfolios relative to prior literature.

Our empirical tests evaluate the performance of volatility-managed versions of various trading strategies. Each of the managed portfolios is constructed as

$$f_{\sigma,t} = \frac{c^*}{\hat{\sigma}^2_{t-1}} f_t, \tag{1}$$

where f_t is the original, unscaled portfolio's excess return in month t; $\hat{\sigma}^2_{t-1}$ is the realized variance of daily portfolio returns in month $t-1$; and c^* is a constant chosen such that $f_{\sigma,t}$ and f_t have the same full-sample variance. We consider the nine volatility-managed equity factors from Moreira and Muir (2017) and report detailed results for these factors in the paper.[3] Although this set of strategies provides a reasonable representation of factors from leading asset pricing models, recent studies suggest that a much larger number of characteristics is needed to summarize both covariance risk (e.g., Kelly et al., 2019) and cross-sectional variation in expected stock returns (e.g., Kozak et al., 2020). We therefore augment Moreira and Muir's (2017) sample with a set of volatility-scaled portfolios formed on 94 anomaly variables from Hou et al. (2015) and McLean and Pontiff (2016). Our combined sample of 103 trading strategies allows for a substantially broader assessment of the merits of volatility management.

Most prior studies (e.g. Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016; Barroso and Maio, 2018; Eisdorfer and Misirli, 2020) assess the value of volatility management by directly comparing the Sharpe ratios earned by scaled strategies similar to those in Eq. (1) with the Sharpe ratios earned by the corresponding unscaled strategies. We follow this approach and find no systematic evidence that volatility-managed portfolios outperform their unmanaged versions. Volatility scaling generates a higher Sharpe ratio for five of the nine equity factors examined by Moreira and Muir (2017). In the more comprehensive sample of 103 equity portfolios, the volatility-managed versions outperform in 53 cases, whereas the original versions outperform in 50 cases. We also find that only eight strategies in the broad sample yield statistically significant Sharpe ratio differences in favor of volatility management. These cases are concentrated among momentum-related strategies, in accord with the conclusions from Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016).

These initial tests suggest that stand-alone investments in volatility-managed portfolios do not systematically improve investment outcomes. Moreira and Muir's (2017) broad evidence on the value of volatility management, however, is not based on direct performance comparisons of scaled and unscaled factors. For each factor, they instead estimate the following spanning regression using time-series data on monthly strategy excess returns:

$$f_{\sigma,t} = \alpha + \beta f_t + \varepsilon_t. \tag{2}$$

Moreira and Muir (2017) find that, with the exception of the size factor, each of their volatility-managed portfolios yields a positive regression intercept, and almost all of the estimates are statistically and economically significant. They also confirm that the volatility-managed portfolios continue to exhibit large alphas when the Fama and French (1993) factors are included in Eq. (2) as additional controls. They interpret the results as strong evidence in favor of volatility management and highlight that the positive alphas are synonymous with increased Sharpe ratios and utility gains for mean-variance investors.

We reproduce Moreira and Muir's (2017) in-sample spanning regression results for the nine equity factors and show that they extend to our broader sample of equity strategies. In particular, 77 out of the 103 volatility-scaled portfolios earn positive alphas in spanning tests, with 23 significantly positive estimates compared with just 3 significantly negative ones. Spanning regressions that control for exposure to the market, size, and value factors produce 70 positive intercepts. These findings offer a confirmation of the potential economic gains from volatility-managed portfolios.

The investment implication of a positive spanning regression intercept is that the optimal ex post combination of scaled and unscaled portfolios (with positive

[2] Representative examples of recent press coverage include "Reassessing the classic risk-return tradeoff," The Financial Times, March 9, 2016 and "When markets get scary, panicking is smart," CNBC, March 23, 2016. For an example of volatility management in practice, BlackRock offers the following description of the investment strategy for its Managed Volatility V.I. Fund: "In periods of heightened volatility, the portfolio will de-risk into less volatile assets like fixed income and cash and re-risk when market turbulence subsides."

[3] The equity factors in Moreira and Muir (2017) include the market, size, and value factors from the Fama and French (1993) three-factor model, the momentum factor from the Carhart (1997) four-factor model, the profitability and investment factors from both the Fama and French (1993) five-factor model and the Hou et al. (2015) q-factor model, and Frazzini and Pedersen's (2014) betting-against-beta factor.

weight on the scaled portfolio) expands the mean-variance frontier relative to the unscaled portfolio (e.g., Gibbons et al., 1989). That is, the increased Sharpe ratios and utility gains referenced in Moreira and Muir (2017) are earned by combinations of two strategies rather than by the volatility-managed portfolios themselves. This interpretation highlights a concern with Moreira and Muir's (2017) spanning regression evidence in terms of its implications for real-time investors. Moreira and Muir (2017) acknowledge the importance of out-of-sample applications by emphasizing that their volatility-managed strategies are "easy to implement in real time." Although a volatility-managed portfolio constructed according to Eq. (1) is straightforward to construct in real time, the investment strategy implied by Eq. (2) is not. Because the optimal weighting of scaled and unscaled portfolios depends on in-sample return moments, the required strategy is not known prior to the end of the sample.

A natural question is whether real-time investors can capture the economic gains implied by the spanning regressions. We conduct an extensive analysis of the out-of-sample performance of combinations of volatility-managed and original portfolios. We adopt the standard approach of using a training sample of historical data to estimate optimal allocations to the scaled and unscaled versions of a given strategy. Prior literature suggests that estimation risk is a key issue in the out-of-sample, mean-variance portfolio choice problem associated with Eq. (2) (e.g., Black and Litterman, 1992; Green and Hollifield, 1992; Jagannathan and Ma, 2003; DeMiguel et al., 2009a; 2009b). In practice, estimated portfolio weights are often unstable, and real-time portfolios often underperform considerably relative to their in-sample optimal counterparts. There is ex ante reason for optimism in the current context, however, given the strong in-sample results and DeMiguel et al.'s (2009b) evidence that out-of-sample performance degradation is less severe when the number of test assets is small.

Our out-of-sample tests focus on quantifying the economic impact of including a volatility-managed portfolio in the investment opportunity set. We compare the Sharpe ratio and certainty equivalent return (CER) for two real-time strategies: (i) a strategy that allocates between a given volatility-managed portfolio, its corresponding original portfolio, and a risk-free security and (ii) a strategy constrained to invest only in the original portfolio and the risk-free asset. The baseline results correspond to a mean-variance investor with a risk aversion coefficient of five and also feature a leverage constraint of five on portfolio positions. Our design choices to include a risk-free security and a leverage constraint are favorable to real-time investors, as these features reduce estimation risk associated with extreme positions (e.g., Kirby and Ostdiek, 2012).

In contrast to the impressive in-sample results for the nine equity factors studied by Moreira and Muir (2017), volatility management often harms real-time performance. The out-of-sample strategy combining the volatility-managed market portfolio and the unmanaged market portfolio, for example, earns an annualized Sharpe ratio of 0.42 compared with 0.46 for the strategy that limits its risky investment set to the unmanaged market portfolio. This combination strategy also leads to a reduction in CER. There are some positive findings on the out-of-sample value of volatility management, as scaled versions of the momentum (*MOM*), profitability (*ROE*), and betting-against-beta (*BAB*) factors contribute to large utility gains for mean-variance investors. Nevertheless, there is little statistical or economic evidence for the remaining six factors that incorporating volatility management improves real-time portfolio outcomes.

The extended sample of 103 trading strategies provides our most comprehensive and convincing evidence on the poor out-of-sample performance of combination strategies. In our base case design, the real-time combination of volatility-managed and original portfolios earns a lower CER relative to the original portfolio in 72 of 103 cases. We consider a wide range of robustness checks by iterating through alternative design features, including minimum window lengths of historical data used to form portfolio weights, expanding versus rolling training samples, investor risk aversion parameters, and leverage constraints on portfolio positions. We also study the impact of augmenting the investment opportunity set with the three Fama and French (1993) factors. None of these reasonable modifications tilts the results in favor of the combination strategies.

Finally, we explore the economic underpinnings of our results. We do so by comparing the out-of-sample performance of the combination strategies that incorporate volatility-managed portfolios with the out-of-sample performance of traditional anomaly strategies. We find that translating alpha into real-time performance is challenging in general, but out-of-sample performance degradation is noticeably more severe in the volatility-managed portfolios setting. We also provide statistical evidence that the more tenuous link between in-sample alpha and out-of-sample performance in the volatility-managed portfolios setting is attributable to a propensity for structural breaks in the spanning regressions for the volatility-managed portfolios. Simply put, the spanning regression parameters that real-time investors estimate from past data often fail to provide a reliable indication of the future performance of volatility-managed portfolios relative to their unscaled versions.

We contribute to the literature on volatility management by offering a complementary viewpoint to the one presented in Moreira and Muir (2017). Their study provides important insights on the dynamics of the conditional risk-return relation, and their in-sample analysis is valuable to researchers attempting to map the data to an asset pricing model in the rational expectations framework (i.e., Inoue and Kilian, 2004).[4] Such an approach inherently

[4] An important contribution of Moreira and Muir (2017) is to demonstrate that several leading macrofinance models are unable to generate the large, positive estimate of the spanning regression alpha for the market portfolio. In the Internet Appendix, we complement their diagnostic by introducing a decomposition of spanning regression alpha into a component that measures the relation between lagged volatility and return, $(1+\bar{f}_t^2/\hat{\sigma}_f^2)\text{cov}(c^*/\hat{\sigma}_{t-1}^2, f_t)$, and a component that measures the relation between lagged and current volatility, $-(\bar{f}_t/\hat{\sigma}_f^2)\text{cov}(c^*/\hat{\sigma}_{t-1}^2, f_t^2)$. We decompose the market factor's annualized spanning regression alpha of 4.63% and find that the first and second components contribute −0.24% and 4.87%, respectively. This diagnostic indicates that an asset

assumes that investors know more than the econometrician does about the true data generating process. From a practical investment perspective, their results speak to the potential economic benefits of volatility-managed portfolios. Our findings suggest that the in-sample alphas and utility gains do not readily translate into enhanced portfolio outcomes for investors who must learn about volatility-managed portfolios from prior data.

The remainder of the paper is organized as follows. Section 2 describes the data and introduces volatility-managed portfolios. Section 3 compares volatility-managed and original strategies. Section 4 contains our empirical tests on real-time strategies that combine volatility-managed portfolios with their unscaled versions, and Section 5 concludes. The Appendix presents additional detail on data construction, and the Internet Appendix reports supplementary results.

2. Data

Section 2.1 introduces the data on trading strategies used in our empirical tests, and Section 2.2 discusses the construction of volatility-managed portfolios.

2.1. Data description

We first consider the nine equity factors examined by Moreira and Muir (2017). We collect daily and monthly data on factor excess returns for the market (*MKT*), size (*SMB*), and value (*HML*) factors from the Fama and French (1993) three-factor model, a momentum factor (*MOM*), the profitability (*RMW*) and investment (*CMA*) factors from the Fama and French (1993) five-factor model, the profitability (*ROE*) and investment (*IA*) factors from the Hou et al. (2015) *q*-factor model, and Frazzini and Pedersen's (2014) betting-against-beta factor (*BAB*).[5,6]

We augment the first group of test portfolios with a second group covering a broader set of trading strategies based on established asset pricing anomalies. We start with the lists of anomaly variables reported in Hou et al. (2015) and McLean and Pontiff (2016). We then restrict our analysis to strategies that are based on a single, continuous sorting variable and can be constructed using the Center for Research in Security Prices (CRSP) Monthly and Daily Stock Files, the Compustat Fundamentals Annual and Quarterly Files, and the Institutional Brokers Estimate System (IBES) database. This process identifies 94 anomaly variables, which we summarize in Table A1. For each anomaly, we construct a value-weighted hedge portfolio that takes a long (short) position in the decile of stocks predicted to outperform (underperform) based on prior literature. We use the CRSP daily (monthly) file to construct the daily (monthly) return series for each anomaly. The Appendix provides additional detail on our sample selection and portfolio formation procedures.

The combined sample consists of 103 equity trading strategies (i.e., 9 factors and 94 anomaly portfolios). Many of the strategies are based on related characteristics, and we group them into the following eight categories based on the classification scheme in Hou et al. (2015): accruals, intangibles, investment, market, momentum, profitability, trading, and value.

2.2. Construction of volatility-managed portfolios

For a given asset pricing factor or anomaly portfolio, let f_t be the buy-and-hold excess portfolio return in month t. We follow Moreira and Muir (2017) and construct the corresponding volatility-managed portfolio return as

$$f_{\sigma,t} = \frac{c^*}{\hat{\sigma}^2_{t-1}} f_t, \tag{3}$$

where c^* is a constant and $\hat{\sigma}^2_{t-1}$ is the realized variance of daily returns during the month preceding the portfolio formation date. The managed portfolio is a scaled version of the original strategy, with investment positions proportional to the inverse of lagged variance. Let $j = 1, \ldots, J_t$ index days in month t, and let f_t^j be the excess return for a given portfolio on day j of month t. We compute realized variance in month t as

$$\hat{\sigma}^2_t = \frac{22}{J_t} \sum_{j=1}^{J_t} \left(f_t^j\right)^2. \tag{4}$$

Given the full time series of f_t and $\hat{\sigma}^2_{t-1}$, we select the scaling parameter, c^*, such that f_t and $f_{\sigma,t}$ have the same unconditional volatility. The scaling parameter is not known to an investor in real time, but we note that some performance measures (e.g., Sharpe ratios and appraisal ratios) are invariant to the choice of c^*.[7] In our applications, f_t always represents the excess return for a zero-cost portfolio. As such, the dynamic investment position in the underlying portfolio, $c^*/\hat{\sigma}^2_{t-1}$, is a measure of the leverage required to invest in the volatility-managed portfolio in month t.

3. Direct comparisons

Several prior studies argue that volatility-managed versions of popular trading strategies exhibit impressive

pricing model must produce much more volatility in market return variance compared with traditional macrofinance models to match the sample moments from the data. Our decomposition complements the spanning regression alpha proposed by Moreira and Muir (2017) by providing additional guidance on model features that could reproduce this alpha.

5 Data on *MKT*, *SMB*, *HML*, *MOM*, *RMW*, and *CMA* are from Kenneth French's website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. Data on *BAB* are from Andrea Frazzini's website at http://people.stern.nyu.edu/afrazzin/. We thank Kenneth French and Andrea Frazzini for making these data available. We thank Lu Zhang for sharing the data on *ROE* and *IA*.

6 Moreira and Muir (2017) also examine a currency carry strategy. We omit this portfolio from our set of test assets given our focus on equity strategies.

7 Several studies differ from Eqs. (3) and (4) in their construction of volatility-managed trading strategies. These differences include scaling by realized standard deviation rather than by realized variance (e.g., Barroso and Santa-Clara, 2015; Barroso and Maio, 2018), using a parametric model to estimate volatility (e.g., Daniel and Moskowitz, 2016; Moreira and Muir, 2017), using longer intervals to estimate lagged realized volatility (e.g., Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016), scaling by market volatility rather than by factor-specific volatility (e.g., Eisdorfer and Misirli, 2020), and selecting c^* to achieve a target level of strategy risk rather than to match the standard deviation of the original factor. Daniel and Moskowitz (2016) also incorporate information on the conditional factor mean in specifying their dynamic momentum strategy.

Table 1
Volatility-managed and original factors.
The table compares the performance of volatility-managed and original versions of nine equity factors. The factors are the market (*MKT*), size (*SMB*), and value (*HML*) factors from the Fama and French (1993) three-factor model, a momentum factor (*MOM*), the profitability (*RMW*) and investment (*CMA*) factors from the Fama and French (2015) five-factor model, the profitability (*ROE*) and investment (*IA*) factors from the Hou et al. (2015) *q*-factor model, and Frazzini and Pedersen's (2014) betting-against-beta factor (*BAB*). For a given factor, the volatility-managed factor return in month t is $f_{\sigma,t} = (c^*/\hat{\sigma}^2_{t-1})f_t$, where f_t is the monthly return for the original factor, $\hat{\sigma}^2_{t-1}$ is the realized variance of daily factor returns in month $t-1$, and c^* is a constant chosen so that f_t and $f_{\sigma,t}$ have the same unconditional standard deviation over the full sample period. We present the mean return, standard deviation, and annualized Sharpe ratio for each original (volatility-managed) factor in Panel A (Panel B). The means and standard deviations are reported in percentage per year. Panel C shows the difference between the Sharpe ratio of the volatility-managed factor and that of the original factor, and the figures in brackets are *p*-values from Jobson and Korkie (1981) tests. Panel D reports the correlation between each original factor and the corresponding volatility-managed factor and the 1st, 50th, and 99th percentiles of the time-series distribution of the scaled factor's implied weight in the original factor. The sample period starts in August 1926 for *MKT, SMB*, and *HML*; January 1927 for *MOM*; August 1963 for *RMW* and *CMA*; February 1967 for *ROE* and *IA*; and February 1931 for *BAB*. The sample periods end in December 2016.

	Factor								
	MKT (1)	*SMB* (2)	*HML* (3)	*MOM* (4)	*RMW* (5)	*CMA* (6)	*ROE* (7)	*IA* (8)	*BAB* (9)
Panel A: Performance measures for original factors									
Mean	7.80	2.57	4.84	7.94	2.92	3.72	6.52	4.99	8.23
Standard deviation	18.61	11.12	12.14	16.39	7.71	6.97	8.83	6.48	10.71
Sharpe ratio	0.42	0.23	0.40	0.48	0.38	0.53	0.74	0.77	0.77
Panel B: Performance measures for volatility-managed factors									
Mean	9.55	0.86	4.64	16.17	3.94	2.79	9.39	4.69	10.81
Standard deviation	18.61	11.12	12.14	16.39	7.71	6.97	8.83	6.48	10.71
Sharpe ratio	0.51	0.08	0.38	0.99	0.51	0.40	1.06	0.72	1.01
Panel C: Performance comparisons									
Sharpe ratio difference	0.09	−0.15	−0.02	0.50	0.13	−0.13	0.32	−0.05	0.24
	[0.30]	[0.09]	[0.86]	[0.00]	[0.29]	[0.23]	[0.01]	[0.68]	[0.01]
Panel D: Properties of volatility-managed factors									
Correlation with original factor	0.63	0.63	0.57	0.48	0.59	0.68	0.68	0.70	0.62
$P_{01}(c^*/\hat{\sigma}^2_{t-1})$	0.04	0.03	0.04	0.04	0.04	0.06	0.06	0.06	0.04
$P_{50}(c^*/\hat{\sigma}^2_{t-1})$	0.96	0.81	1.02	1.01	1.11	0.97	1.08	0.96	1.00
$P_{99}(c^*/\hat{\sigma}^2_{t-1})$	6.47	5.07	5.89	8.64	5.02	4.56	4.73	4.45	5.09

performance. The empirical tests in these studies typically highlight that the volatility-managed version of a given portfolio directly outperforms its corresponding unmanaged portfolio. For example, Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) compare Sharpe ratios and cumulative returns for scaled and unscaled versions of the momentum factor. Barroso and Maio (2018) and Eisdorfer and Misirli (2020) present similar evidence for the betting-against-beta and financial distress strategies, respectively. We adopt this focus on direct performance comparisons in this section and contribute to the literature by assessing a much broader set of volatility-managed portfolios.

We begin our analysis by presenting detailed results on direct performance comparisons for the nine equity factors in Table 1. Panel A (Panel B) reports mean excess returns, standard deviations, and Sharpe ratios for the original (volatility-managed) factors, and Panel C shows the Sharpe ratio differences between the volatility-managed and original factors. To determine whether each difference is statistically significant, we follow the approach proposed by Jobson and Korkie (1981).[8] Both the volatility-managed and original versions of each factor earn positive average returns, but neither version yields systematically superior performance across the factors. In five cases the volatility-managed factor earns a higher average return and Sharpe ratio than the original strategy does, whereas the original factor outperforms in the remaining four cases. Three of the nine differences are significantly positive, as the volatility-managed versions of *MOM, ROE*, and *BAB* achieve Sharpe ratio gains by outperforming the original factors by 8.23%, 2.86%, and 2.58% per year, respectively (recall that the volatility-managed and original versions of each strategy have the same standard deviation by construction). The findings are consistent with prior literature on the benefits of volatility management for the momentum (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2016) and betting-against-beta (Barroso and Maio, 2018) strategies. Based on these prior studies, the significant result for volatility-managed *ROE* is perhaps also unsurprising, given the original factor's relatively high

[8] Let $\hat{\mu}_i$ and $\hat{\sigma}_i$ be the mean and standard deviation of excess returns for portfolio i over a period of length T. Similarly, $\hat{\mu}_j$ and $\hat{\sigma}_j$ are the mean and standard deviation for portfolio j, and $\hat{\sigma}_{i,j}$ is the covariance between excess returns for the two portfolios. To test the null hypothesis of equal Sharpe ratios for portfolios i and j, we compute the following Jobson and Korkie (1981) test statistic, which is asymptotically distributed as a standard normal: $\hat{z}_{JK} = \frac{\hat{\sigma}_j\hat{\mu}_i - \hat{\sigma}_i\hat{\mu}_j}{\sqrt{\hat{\theta}}}$, where $\hat{\theta} = \frac{1}{T}\left(2\hat{\sigma}_i^2\hat{\sigma}_j^2 - 2\hat{\sigma}_i\hat{\sigma}_j\hat{\sigma}_{i,j} + \frac{1}{2}\hat{\mu}_i^2\hat{\sigma}_j^2 + \frac{1}{2}\hat{\mu}_j^2\hat{\sigma}_i^2 - \frac{\hat{\mu}_i\hat{\mu}_j}{\hat{\sigma}_i\hat{\sigma}_j}\hat{\sigma}_{i,j}^2\right)$. The test incorporates the correction noted by Memmel (2003).

Table 2
Summary of volatility-managed and original portfolios: broad sample.
The table summarizes results for direct comparisons of volatility-managed and original versions of 103 trading strategies. For a given factor or anomaly portfolio, the volatility-managed strategy return in month t is $f_{\sigma,t} = (c^*/\hat{\sigma}^2_{t-1})f_t$, where f_t is the monthly return for the original portfolio, $\hat{\sigma}^2_{t-1}$ is the realized variance of daily portfolio returns in month $t-1$, and c^* is a constant chosen so that f_t and $f_{\sigma,t}$ have the same unconditional standard deviation over the full sample period. In each case, we compute the difference between the Sharpe ratio of the volatility-managed portfolio and that of the original portfolio. Panel A reports results for the full set of 103 trading strategies. Panel B presents separate results for the 9 factors and the 94 anomaly portfolios. Panel C breaks the results down by trading strategy type. For each set of comparisons, the table presents the number of Sharpe ratio differences that are positive, positive and significant at the 5% level, negative, and negative and significant at the 5% level. We assess statistical significance of the Sharpe ratio differences using the Jobson and Korkie (1981) approach.

		Sharpe ratio difference	
Sample (1)	Total (2)	$\Delta SR > 0$ [Signif.] (3)	$\Delta SR < 0$ [Signif.] (4)
Panel A: Combined sample			
All trading strategies	103	53 [8]	50 [4]
Panel B: By category			
Factors	9	5 [3]	4 [0]
Anomaly portfolios	94	48 [5]	46 [4]
Panel C: By trading strategy type			
Accruals	10	4 [0]	6 [0]
Intangibles	10	3 [0]	7 [0]
Investment	11	3 [0]	8 [1]
Market	1	1 [0]	0 [0]
Momentum	9	9 [5]	0 [0]
Profitability	22	15 [1]	7 [1]
Trading	21	11 [1]	10 [1]
Value	19	7 [1]	12 [1]

correlation with *MOM* and *BAB*.[9] Aside from *MOM*, *ROE*, and *BAB*, the remaining six factors exhibit differences in Sharpe ratios that are insignificant at the 5% level and average return differentials between −1.71% and 1.74% per year.

Panel D of Table 1 shows that the correlation coefficients for excess returns for the scaled and unscaled strategies range from 0.48 to 0.70. Panel D also highlights that investing in volatility-scaled portfolios requires aggressively altering exposures to the underlying factors over time. Although the median investment position for each of the dynamic portfolios is around one, the 99th percentile of required leverage exceeds 400% in each case and reaches as high as 864% for the momentum strategy.

To offer a more comprehensive view on the performance of volatility-managed portfolios, we examine the expanded set of 103 equity trading strategies. Table 2 provides a summary of the Sharpe ratio differences between the volatility-managed and original strategies. Many of the portfolios are formed on related characteristics, so we classify them into strategy types relating to accruals, intangibles, investment, market, momentum, profitability, trading, and value. Across all 103 strategies, Panel A of Table 2 reports the number of Sharpe ratio differences that are positive or negative and the number of these differences that are statistically significant at the 5% level.[10] Panel B separates the results for the 9 factors and 94 anomaly portfolios, and Panel C displays the corresponding figures for each strategy type. As in Table 1, positive (negative) differences indicate outperformance (underperformance) for the volatility-scaled versions.

The results in Table 2 suggest that volatility-managed portfolios do not systematically outperform their original counterparts. In Panel A, volatility management leads to improved and worsened performance at roughly the same frequency. The performance differences across the 103 trading strategies include 53 positive and 50 negative values, and few of the differences are statistically significant.

Panel C reveals that the majority of the significantly positive Sharpe ratio differences are attributable to the nine momentum strategies. Volatility management improves performance for every momentum strategy, and five of the nine performance differences are statistically significant at the 5% level. The findings are consistent with the impressive performance of volatility-managed momentum portfolios demonstrated by Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016). Barroso and Santa-Clara (2015) notably find that managing the risk of momentum is robust to using alternative windows (i.e., one, three, or six months) to estimate realized volatility. We complement this result by showing that the performance of volatility-managed momentum is also robust to several alternative definitions of the momentum strategy itself.

Outside of the momentum group, the volatility-managed versions of just three strategies exhibit statistically significant outperformance: *ROE*, *BAB*, and Loughran and Wellman's (2012) enterprise multiple. We also find that volatility scaling results in significantly lower Sharpe ratios for four portfolios. This group includes growth in book equity (Lockwood and Prombutr, 2010), change in sales less change in inventory (Abarbanell and Bushee, 1998), 1/share price (Miller and Scholes, 1982), and long-term reversal (De Bondt and Thaler, 1985).

To interpret the broad-based results in Table 2, we compare the total number of positive Sharpe ratio differences in the data with the number that would be expected under reasonable assumptions about the data generating process. The performance of a given volatility-managed portfolio is driven by two factors: (i) the relation between lagged volatility and future volatility and (ii) the relation between lagged volatility and future expected return. Volatility management is likely to be successful if volatility is persistent and the risk-return relation is flat. In this scenario, a portfolio's conditional Sharpe ratio is negatively associated with its lagged volatility, and investors can capitalize on these dynamics in the conditional risk-return trade-off by taking more aggressive investment positions following low-volatility periods. If lagged volatility is negatively related to average return for a given strategy, volatility management becomes even more attractive. A

[9] The correlation coefficient between *ROE* and *MOM* (*BAB*) is 0.50 (0.26).

[10] We present detailed results on performance comparisons for the individual factors and anomaly portfolios in the Internet Appendix.

positive risk-return trade-off, in contrast, makes volatility management less effective.

We examine the performance of volatility-managed portfolios relative to the null hypothesis that each of the 103 original portfolios exhibits persistence in conditional volatility but zero correlation between lagged volatility and future expected return. This null reflects both the overwhelming evidence of volatility clustering in asset returns (e.g., Engle, 2004) and the lack of conclusive evidence on the relation between conditional variance and average return for the strategies of interest.[11] The test of this hypothesis allows us to use the performance of volatility-managed portfolios to assess the risk-return trade-off in our broad sample.

To evaluate the performance of volatility-managed portfolios relative to the expectation under this null hypothesis, we develop a bootstrap analysis that is described in detail in the Internet Appendix. This procedure generates bootstrap samples that (i) preserve for each strategy the empirical relation between the inverse of lagged variance and the variance of the realized strategy return and (ii) have, on average, no predictive relation between lagged variance and return. We generate 100,000 bootstrap samples under the null hypothesis, run 103 direct performance comparisons in each bootstrap sample, and count the number of positive performance differences. We then compare the number of positive differences in the data to the bootstrap distribution under the null hypothesis to assess statistical significance.

Fig. 1 shows the bootstrap distribution of the number of positive Sharpe ratio differences. As expected, volatility management is an attractive strategy under the null hypothesis of persistent volatility but no risk-return relation. Across the bootstrap samples, the average number of positive Sharpe ratio differences is 66 out of 103. In contrast, the 53 observed positive Sharpe ratio differences in the data are considerably fewer than what would be expected under this null. The two-tailed bootstrap p-value of 0.01 indicates that the null hypothesis of persistent volatility but no risk-return relation is rejected. As noted above, a positive risk-return relation for a given strategy works to degrade the performance of a given volatility-managed portfolio such that the data indicate that the risk-return relations tend to be positive across the broad set of 103 strategies.

We also consider a second null hypothesis that the expected Sharpe ratios for the volatility-managed and original versions of each strategy are equal. The corresponding test provides insight on the practical question of whether or not investors should generally favor volatility-managed portfolios over original portfolios. The frequencies of positive and negative performance differences in Table 2 are in line with this second null hypothesis. We compute the two-tailed p-value for the number of positive Sharpe ratio differences relative to the binomial distribution under the null that the volatility-managed and original versions of each of the 103 portfolios are equally likely to outperform the other. The p-value of 0.84 indicates that a null hypothesis of equal performance is not rejected. These findings are in accord with the general conclusion that volatility-managed strategies do not systematically outperform the corresponding original strategies.

4. Combination strategies

Whereas the results in Section 3 provide evidence that volatility-managed portfolios do not systematically outperform original portfolios, Moreira and Muir's (2017) spanning regression tests suggest that volatility-scaled portfolios are potentially more valuable when used in combination with their original counterparts rather than as stand-alone investments. In Sections 4.1 and 4.2, we note the differences between direct performance comparisons and spanning tests and highlight the portfolio implications of Moreira and Muir's (2017) spanning regression approach to evaluating volatility-managed portfolios. These sections also show that the trading strategies implied by spanning regressions are not available to real-time investors. Given this limitation, we turn to a comprehensive analysis of the out-of-sample performance of combination strategies that incorporate volatility management in Section 4.3.

4.1. Spanning regressions

Moreira and Muir's (2017) evidence on the success of volatility-managed portfolios follows from the spanning regression approach. They evaluate volatility-managed factors by estimating time-series regressions of the form

$$f_{\sigma,t} = \alpha + \beta f_t + \varepsilon_t. \tag{5}$$

Their tests focus on α, which they estimate to be positive, economically large, and statistically significant for a wide range of popular asset pricing factors.[12] Moreira and Muir (2017) further emphasize that positive alphas are synonymous with increased Sharpe ratios relative to the original factors and pronounced utility gains for mean-variance investors. In this section, we consider the portfolio properties associated with the spanning regressions and develop intuition for why the in-sample results do not readily extend to real-time investment settings.

We note that a positive alpha in regression (5) is a lower bar for declaring success of a given managed strategy relative to a positive Sharpe ratio difference in a direct comparison (e.g., Section 3). We demonstrate this point formally in the Internet Appendix. In particular, consider the case in which both the managed and unmanaged versions of a given strategy earn positive average returns (i.e., $\bar{f}_{\sigma,t} > 0$ and $\bar{f}_t > 0$). Because $f_{\sigma,t}$ and f_t have identical

---

[11] Several studies examine the relation between the expected return and conditional variance of the market portfolio. In early work, for example, Campbell (1987) finds a negative relation, whereas French et al. (1987) find an insignificant or positive relation. Subsequent papers also produce mixed evidence on the sign of the risk-return trade-off for the market, with the results being sensitive to the sample period, the method for estimating conditional volatility, and the time-series approach to relating conditional variance and return. There is considerably less published evidence on the nature of the risk-return trade-off for other factors and anomaly portfolios.

[12] Moreira and Muir also demonstrate that volatility-managed strategies earn positive alphas when unscaled versions of the Fama and French (1993) three factors are included in Eq. (5) as additional controls.

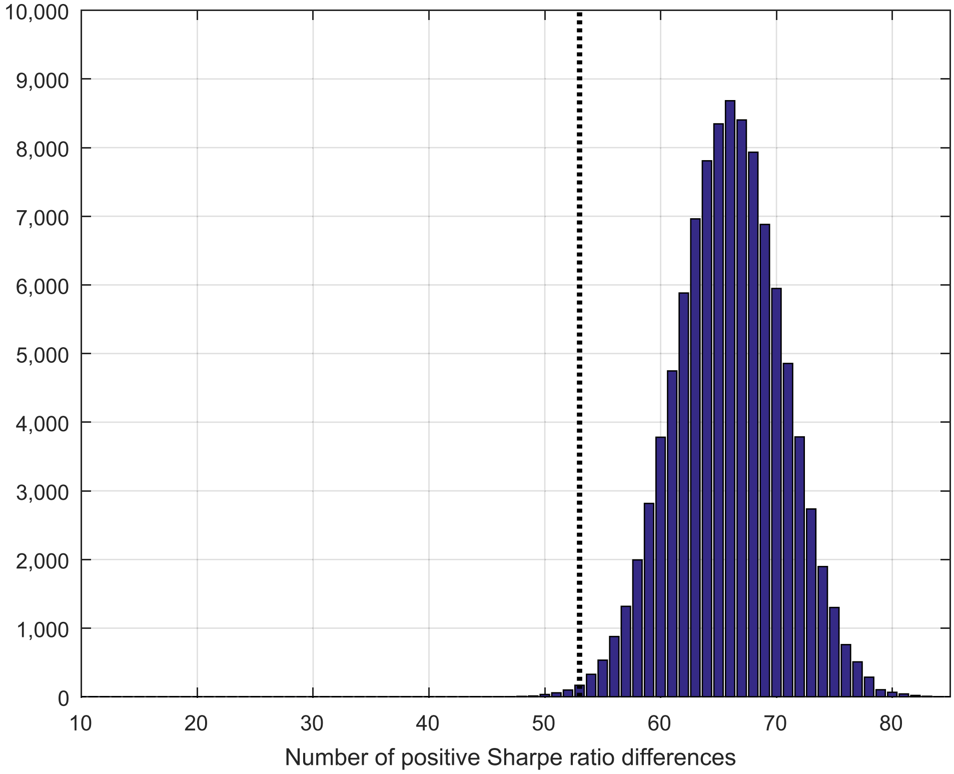

Fig. 1. Bootstrap distribution. The figure shows the bootstrap distribution of the number of positive Sharpe ratio differences across the 103 strategies under the null hypothesis of persistence in volatility with no risk-return relation. The bootstrap procedure is described in the Internet Appendix, and the histogram shows the distribution for 100,000 bootstrap samples. The vertical dotted line indicates that 53 of 103 comparisons produce positive performance differences in the data.

full sample volatility by construction, the scaled portfolio achieves a higher Sharpe ratio as long as $\bar{f}_{\sigma,t} > \bar{f}_t$. The requirement for a positive spanning test alpha, however, is $\bar{f}_{\sigma,t} > \hat{\rho}\bar{f}_t$, where $\hat{\rho}$ is the unconditional correlation between the scaled and unscaled factors.[13] As shown in Table 1, these correlations range from 0.48 to 0.70 for the equity strategies examined by Moreira and Muir (2017).[14] Volatility scaling could, therefore, lead to a 30% or larger drop in Sharpe ratio and at the same time produce a positive spanning regression intercept.

A positive alpha in Eq. (5) does indicate that the optimal ex post combination of scaled and unscaled factors (with positive weight on the scaled factor) expands the mean-variance frontier relative to the original factor (e.g., Gibbons et al., 1989). This point follows from the well-known link between spanning tests and portfolio optimization under mean-variance utility. Consider a mean-variance investor who allocates between excess returns $f_{\sigma,t}$ and f_t. Given the sample moments for $f_{\sigma,t}$ and f_t, the investor's ex post optimal vector of fixed allocations to the volatility-managed and original factors, $a = [x_\sigma^* \quad x^*]^\top$, is the solution to the following problem:

$$\max_a U(a) = a^\top \hat{\mu} - \frac{\gamma}{2} a^\top \hat{\Sigma} a, \tag{6}$$

where $\hat{\mu} = [\bar{f}_{\sigma,t} \quad \bar{f}_t]^\top$ is the 2×1 vector of mean excess returns, $\hat{\Sigma}$ is the 2×2 variance–covariance matrix, and γ is the investor's risk aversion parameter. In this setup, the investor implicitly has access to a risk-free security. The vector of optimal portfolio weights is given by

$$a = \begin{bmatrix} x_\sigma^* \\ x^* \end{bmatrix} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}, \tag{7}$$

and the vector of optimal relative weights in the two risky assets is

$$\begin{bmatrix} w_\sigma^* \\ w^* \end{bmatrix} = \frac{\hat{\Sigma}^{-1}\hat{\mu}}{|1_2^\top \hat{\Sigma}^{-1}\hat{\mu}|}, \tag{8}$$

where 1_2 is a 2×1 vector of ones. The weights in the two risky assets depend on the investor's risk aversion but the relative allocations across the two assets do not. Given that the scaled and unscaled factors have the same full sample standard deviation, the covariance matrix is

$$\hat{\Sigma} = \hat{\sigma}_f^2 \begin{bmatrix} 1 & \hat{\rho} \\ \hat{\rho} & 1 \end{bmatrix}, \tag{9}$$

where $\hat{\sigma}_f^2$ is the unconditional variance of factor returns. The ex post optimal allocation to the volatility-managed

[13] We also show in the Internet Appendix that $\hat{\rho} = \hat{\beta}$, where $\hat{\beta}$ is the estimate of the slope coefficient from the univariate spanning regression in Eq. (5).

[14] The correlations for the broader sample of 103 strategies are between 0.48 and 0.80 (see Internet Appendix).

portfolio is then proportional to the spanning regression alpha,

$$x_\sigma^* = \frac{\hat{\alpha}}{\gamma \hat{\sigma}_f^2 (1 - \hat{\rho}^2)}. \tag{10}$$

The optimal strategy assigns positive weight to the volatility-managed portfolio if and only if this portfolio earns a positive spanning regression alpha.

We combine the optimal investment policy in Eq. (7) with the definition of the volatility-managed portfolio in Eq. (3) to generate the dynamic investment rule

$$y_t^* = x_\sigma^* \left(\frac{c^*}{\hat{\sigma}_{t-1}^2} \right) + x^*. \tag{11}$$

The investor's ex post optimal policy allocates a static weight x_σ^* to the volatility-managed factor and a static weight x^* to the original factor. This policy is equivalent to dynamically adjusting the position (i.e., y_t^*) in the original factor according to Eq. (11). We denote the Sharpe ratio earned by this combination strategy as $SR(y_t^*)$.

Moreira and Muir (2017) link their spanning test results to appraisal ratios and utility gains for investors. These types of metrics can be interpreted in the context of mean-variance portfolio choice. The appraisal ratio for a given scaled strategy is

$$AR = \frac{\hat{\alpha}}{\hat{\sigma}_\varepsilon}, \tag{12}$$

where $\hat{\alpha}$ is the intercept and $\hat{\sigma}_\varepsilon$ is the standard error of the regression in Eq. (5). The squared appraisal ratio reflects the extent to which volatility management can be used to increase the slope of the mean-variance frontier (e.g., Gibbons et al., 1989):

$$AR^2 = SR(y_t^*)^2 - SR(z^*)^2, \tag{13}$$

where $SR(z^*)$ is the Sharpe ratio earned by a mean-variance investor who does not have access to the volatility-managed portfolio.[15] Similarly, we quantify the in-sample utility gains from volatility management by comparing the CER for the investor who optimizes according to Eq. (6) with that of the investor who is constrained to invest in the original factor. We solve for the in-sample CER difference analytically as

$$\Delta CER = \frac{SR(y_t^*)^2 - SR(z^*)^2}{2\gamma}. \tag{14}$$

Based on Eq. (13), Moreira and Muir (2017) note that positive alphas in Eq. (5) indicate that volatility management increases Sharpe ratios relative to the original factors. They also use a relation similar to Eq. (14) as the basis for their conclusion that volatility timing leads to large utility gains for mean-variance investors.[16] These improvements in portfolio performance are based on ex post results, however, which potentially overstate the value of volatility management in practice. An investor could only achieve the utility gains by combining the scaled and unscaled versions of a particular factor using weights that are unknown prior to observing the full sample of factor returns. As such, these types of strategies are not implementable in real time. We empirically demonstrate the link between Moreira and Muir's (2017) in-sample regression results and optimal portfolio choice in the following section.

4.2. In-sample tests

Table 3 reproduces Moreira and Muir's (2017) spanning tests for the nine volatility-managed equity factors. Panel A reports estimates from unconditional regressions of monthly volatility-managed portfolio returns on original portfolio returns following Eq. (5). The volatility-managed alpha is annualized by multiplying the monthly estimate by 12, and the appraisal ratio for the regression is annualized by multiplying the monthly figure by $\sqrt{12}$.

The results in Panel A.1 of Table 3 provide strong empirical support for the in-sample benefits of volatility-managed portfolios. We confirm Moreira and Muir's (2017) finding that the volatility-managed factors often generate positive alphas relative to the original factors. In particular, the volatility-managed *MKT, MOM, RMW, ROE*, and *BAB* portfolios have positive and statistically significant alphas at the 1% level, and the managed *IA* and *HML* alphas are significant at the 10% level. Several of the performance estimates are large in economic magnitude, with the volatility-managed momentum alpha of 12.39% per year standing out as particularly striking.[17]

The large appraisal ratios in Table 3 indicate that volatility management expands the ex post mean-variance frontier relative to the original factors, a conclusion that follows from Eqs. (12) and (13). Building on this point, Moreira and Muir (2017) accentuate the large associated lifetime utility gains for investors timing volatility. The portfolio strategies required to achieve these benefits, however, are not implementable for real-time investors. Taking the market factor as an example, the volatility-managed alpha of 4.63% per year and associated appraisal ratio of 0.32 suggest large potential gains for investors. Panel A.2 of Table 3 highlights that, to attain these portfolio results, an investor must know the scaling parameter for the volatility-managed market factor and, more important, the optimal risky portfolio mixes a 72% weight in the volatility-scaled market factor with a 28% weight in the original market factor. As shown in Panel A.3, this strategy generates a Sharpe ratio of 0.53 compared with 0.42 for the market portfolio. An investor with $\gamma = 5$ has a CER of 2.79% for the ex post optimal combination strategy versus only 1.76% for the original factor.

[15] The investor without access to the volatility-managed factor optimally invests $z^* = \frac{1}{\gamma}\frac{\bar{f}_t}{\hat{\sigma}_f^2}$ in the unscaled factor and earns a squared Sharpe ratio equal to that of the unscaled factor.

[16] To quantify the economic impact of incorporating a given volatility-managed portfolio into the investment opportunity set, Moreira and Muir (2017) focus on the "percentage utility gain" rather than on the CER difference. Their measure is $\Delta U(\%) = \frac{SR(y_t^*)^2 - SR(z^*)^2}{SR(z^*)^2}$.

[17] The positive unconditional alphas in Table 3 can also be understood in the context of the volatility-timing effects discussed in Lewellen and Nagel (2006) and Boguth et al. (2011). In particular, the conditional alpha for a given volatility-managed portfolio is, by construction, equal to zero. The corresponding unconditional alpha, however, is likely to be positive if the managed portfolio's conditional factor exposure, $c^*/\hat{\sigma}_{t-1}^2$, covaries negatively with the conditional volatility of the unscaled factor.

Table 3
Spanning regressions.
Panel A.1 reports results from univariate spanning regressions of volatility-managed factor returns on the corresponding original factor returns. The spanning regressions are given by $f_{\sigma,t} = \alpha + \beta f_t + \varepsilon_t$, where $f_{\sigma,t}$ (f_t) is the monthly return for the volatility-managed (original) factor. The estimates of α are reported in percentage per year, and the numbers in parentheses are t-statistics based on White (1980) standard errors. For each regression, R^2 is the adjusted R^2 value, and the appraisal ratio is computed as the ratio of alpha to root mean square error. Panel A.2 presents the scaling parameter (c^*) for the volatility-managed factor, the ex post optimal total weight in risky assets ($x_\sigma^* + x^*$), and the ex post optimal relative weights in the volatility-managed (w_σ^*) and original factors (w^*). The vector of portfolio weights is $[x_\sigma^* \quad x^*]^\top = (1/\gamma)\hat{\Sigma}^{-1}\hat{\mu}$, where γ is the risk aversion parameter, $\hat{\Sigma}$ is the 2×2 variance–covariance matrix of $f_{\sigma,t}$ and f_t, and $\hat{\mu}$ is the 2×1 vector of mean excess returns for $f_{\sigma,t}$ and f_t. The vector of relative weights is computed as $[w_\sigma^* \quad w^*]^\top = [x_\sigma^* \quad x^*]^\top / |x_\sigma^* + x^*|$. For each factor, Panel A.3 shows annualized Sharpe ratios and certainty equivalent returns (CERs). The "original factor" results correspond to the ex post optimal combination of original factor and risk-free asset, and the "combination strategy" results correspond to the ex post optimal combination of original factor, volatility-managed factor, and risk-free asset. The results in Panels A.2 and A.3 are for $\gamma = 5$. Panel B adds the Fama and French (1993) three factors as controls in the spanning regressions.

	Factor								
	MKT (1)	*SMB* (2)	*HML* (3)	*MOM* (4)	*RMW* (5)	*CMA* (6)	*ROE* (7)	*IA* (8)	*BAB* (9)
Panel A: Univariate regressions									
Panel A.1: Regression results									
Alpha, α (%)	4.63	−0.76	1.87	12.39	2.23	0.26	4.97	1.18	5.74
	(3.08)	(−0.87)	(1.88)	(7.31)	(2.57)	(0.39)	(5.10)	(1.83)	(5.97)
Beta, β	0.63	0.63	0.57	0.48	0.59	0.68	0.68	0.70	0.62
	(11.32)	(7.75)	(7.65)	(7.13)	(7.10)	(13.82)	(11.12)	(13.59)	(12.97)
R^2	0.40	0.40	0.33	0.23	0.34	0.46	0.46	0.50	0.38
Appraisal ratio, *AR*	0.32	−0.09	0.19	0.86	0.36	0.05	0.77	0.26	0.68
Panel A.2: Ex post optimization parameters									
Scaling parameter, c^*	10.33	2.63	2.95	4.60	1.48	1.53	2.06	1.64	3.20
Risky allocation, $x_\sigma^* + x^*$	0.61	0.34	0.82	1.22	1.45	1.60	2.44	0.70	2.05
Relative factor weights									
Vol-managed factor, w_σ^*	0.72	−0.60	0.46	0.98	0.79	0.12	0.97	0.41	0.78
Original factor, w^*	0.28	1.60	0.54	0.02	0.21	0.88	0.03	0.59	0.22
Panel A.3: Portfolio performance measures									
Sharpe ratio									
Original factor	0.42	0.23	0.40	0.48	0.38	0.53	0.74	0.77	0.77
Combination strategy	0.53	0.25	0.44	0.99	0.52	0.54	1.06	0.81	1.03
Difference	0.11	0.02	0.04	0.50	0.14	0.00	0.32	0.04	0.26
CER (%)									
Original factor	1.76	0.53	1.59	2.35	1.44	2.85	5.46	5.92	5.90
Combination strategy	2.79	0.61	1.94	9.74	2.71	2.88	11.32	6.57	10.52
Difference	1.03	0.08	0.35	7.39	1.27	0.03	5.86	0.65	4.63
Panel B: Additional controls for Fama and French (1993) three factors									
Alpha, α (%)	5.24	−0.56	2.52	10.28	3.02	−0.19	5.51	0.66	5.45
	(3.49)	(−0.65)	(2.52)	(6.56)	(3.49)	(−0.28)	(5.52)	(1.01)	(5.72)
R^2	0.41	0.40	0.35	0.26	0.43	0.47	0.49	0.51	0.39
Appraisal ratio, *AR*	0.37	−0.07	0.26	0.73	0.72	−0.04	0.88	0.15	0.65

Panel A.2 of Table 3 indicates that, with the exception of *SMB*, the ex post optimal combinations feature a positive allocation to the volatility-managed factor. The positive weights follow from the positive spanning test alphas for these factors in Panel A.1 in accord with Eq. (10).[18] Across the nine combination strategies, there is substantial discrepancy in the optimal weight assigned to the volatility-scaled factor. The relative weight for the volatility-managed factor ranges from −60% for *SMB* to 98% for *MOM*, suggesting that there is no obvious fixed trading rule across factors. Finally, the total ex post allocation to the risky asset combination portfolio also differs markedly across the factors. This allocation ranges from 34% for *SMB* to 244% for *ROE*.

Panel A.3 of Table 3 confirms that almost all combination strategies exhibit strong in-sample performance gains relative to the original factors. The most impressive Sharpe ratio improvements correspond to the *MOM* (0.99 for ex post optimal combination portfolio versus 0.48 for original), *ROE* (1.06 versus 0.74), and *BAB* (1.03 versus 0.77) factors. The CER results also indicate large utility gains for several of the combination portfolios. The ex post *MOM* combination strategy, in particular, boasts a CER of 9.74% per year compared with just 2.35% for the original *MOM* factor. Overall, the results are consistent with Moreira and Muir's (2017) finding that incorporating volatility-managed

[18] Panel A.2 of Table 3 also shows that the original version of each factor receives a positive allocation in the ex post optimal combination strategy. These results suggest that "reverse" spanning tests that regress original factors on their volatility-managed versions should generate a similar level of support for the unscaled factors as we observe for the scaled factors in Table 3. We confirm this intuition in the Internet Appendix. Reverse regressions produce positive intercepts for the original factors in all nine cases, and four of the estimates are significant at the 5% level. The results highlight that spanning tests are unsuitable for identifying whether the original or volatility-managed version of a given strategy is superior.

Table 4
Summary of spanning regressions: broad sample.
The table summarizes results from spanning regressions for 103 trading strategies. The spanning regressions are given by $f_{\sigma,t} = \alpha + \beta f_t + \varepsilon_t$, where $f_{\sigma,t}$ (f_t) is the monthly return for the volatility-managed (original) anomaly portfolio. The results in columns (3) and (4) correspond to univariate spanning regressions, and those in columns (5) and (6) are for regressions that add the Fama and French (1993) three factors as controls. Panel A reports results for the full set of 103 trading strategies. Panel B presents separate results for the 9 factors and the 94 anomaly portfolios. Panel C breaks the results down by trading strategy type. For each set of regressions, the table reports the number of alphas that are positive, positive and significant at the 5% level, negative, and negative and significant at the 5% level. We assess statistical significance of the alpha estimates using White (1980) standard errors.

		Univariate regressions		Additional controls for Fama and French (1993) factors	
Sample (1)	Total (2)	$\alpha > 0$ [Signif.] (3)	$\alpha < 0$ [Signif.] (4)	$\alpha > 0$ [Signif.] (5)	$\alpha < 0$ [Signif.] (6)
Panel A: Combined sample					
All trading strategies	103	77 [23]	26 [3]	70 [21]	33 [3]
Panel B: By category					
Factors	9	8 [5]	1 [0]	7 [6]	2 [0]
Anomaly portfolios	94	69 [18]	25 [3]	63 [15]	31 [3]
Panel C: By trading strategy type					
Accruals	10	8 [3]	2 [0]	6 [0]	4 [0]
Intangibles	10	6 [1]	4 [0]	5 [0]	5 [0]
Investment	11	7 [1]	4 [1]	5 [1]	6 [1]
Market	1	1 [1]	0 [0]	1 [1]	0 [0]
Momentum	9	9 [9]	0 [0]	9 [9]	0 [0]
Profitability	22	19 [2]	3 [0]	19 [4]	3 [0]
Trading	21	14 [4]	7 [1]	14 [4]	7 [2]
Value	19	13 [2]	6 [1]	11 [2]	8 [0]

factors into the in-sample portfolio choice problem leads to substantial gains for investors.

Panel B of Table 3 shows that the general conclusions are robust to including the three Fama and French (1993) factors in the spanning tests as additional controls. The spanning regression alphas are positive and statistically significant at the 1% level for the *MKT, HML, MOM, RMW, ROE*, and *BAB* factors. The tests indicate that volatility-managed factors remain important for investors who form ex post optimal portfolios with an investment opportunity set augmented to include the *MKT, SMB*, and *HML* factors.

We also present new evidence on the in-sample benefits of volatility management by applying the tests in Table 3 to the combined sample of 103 trading strategies. The results are summarized in Table 4, with detailed results available in the Internet Appendix. We find that 77 of the 103 scaled portfolios earn positive alphas in univariate spanning tests and, accordingly, are assigned positive weights in the ex post optimal combination portfolios. Twenty-three of the positive estimates are statistically significant at the 5% level. The regression specification that adds the Fama and French (1993) controls produces 70 positive alphas, 21 of which are statistically significant. The broad-based results from Table 4 provide additional support for Moreira and Muir's (2017) general conclusions on the in-sample value of volatility management.

4.3. Out-of-sample tests

The trading strategies suggested by the in-sample spanning tests are not implementable in real time. We therefore examine their out-of-sample counterparts to assess the value of volatility management for real-time, mean-variance investors. When investors are required to form trading strategies based on information available at the time, observed performance may differ from the ex post result for at least two reasons. First, if the conditional risk-return trade-off for a given factor is unstable over time, past data are less likely to be informative about the future potential for volatility management. Along these lines, Whitelaw (1994), Harvey (2001), Brandt and Kang (2004), Ludvigson and Ng (2007), and Lettau and Ludvigson (2010) provide evidence of instability in the risk-return relation for the market factor. Second, estimation risk is a key concern in the real-time portfolio choice problem implied by Eq. (5), as weights are often unstable and the corresponding optimal portfolios tend to perform poorly out of sample (e.g., Black and Litterman, 1992; Green and Hollifield, 1992; Jagannathan and Ma, 2003; DeMiguel et al., 2009a; 2009b).

Given these concerns, we examine whether investors would have benefited from volatility management based on information available in real time. Section 4.3.1 details the design of our out-of-sample tests, and Section 4.3.2 presents the corresponding results. Section 4.3.3 discusses the underlying economic drivers of the differences between our in-sample and out-of-sample evidence.

4.3.1. Out-of-sample strategy design

We adopt a standard real-time portfolio choice design. We start with a sample of T monthly excess return observations for a given factor or anomaly portfolio.

Table 5
Real-time combination strategies.
The table reports results for portfolio strategies that combine original factors and volatility-managed factors. We specify an initial training period length of $K = 120$ months and use an expanding-window design for the out-of-sample tests. The out-of-sample period runs from month $K + 1$ to month T, where T is the total number of sample months for a given factor. In Panel A, the "combination strategy (real time)" results correspond to the real-time combination of original factor, volatility-managed factor, and risk-free asset, and the "combination strategy (ex post optimal)" results correspond to the ex post optimal combination of these assets over the out-of-sample period. The "original factor (real time)" results correspond to the real-time combination of original factor and risk-free asset. The strategies in Panel B include the Fama and French (1993) factors in the investment opportunity set. For each strategy, the table shows the annualized Sharpe ratio and certainty equivalent return (CER) in percentage per year over the out-of-sample period. The figures in brackets are p-values for the Sharpe ratio and CER differences. The p-values are computed following the approaches in Jobson and Korkie (1981) and DeMiguel et al. (2009b), respectively. We use a risk aversion parameter of $\gamma = 5$ and impose a leverage constraint that the sum of absolute weights on the risky factors is less than or equal to five.

	Factor								
	MKT (1)	*SMB* (2)	*HML* (3)	*MOM* (4)	*RMW* (5)	*CMA* (6)	*ROE* (7)	*IA* (8)	*BAB* (9)
Panel A: Real-time combination strategies									
Sharpe ratio									
[S1] Combination strategy (real time)	0.42	0.14	0.38	0.92	0.44	0.52	1.13	0.70	1.09
[S2] Original factor (real time)	0.46	0.19	0.43	0.49	0.31	0.56	0.78	0.68	0.79
Difference, [S1]–[S2]	−0.04	−0.06	−0.06	0.44	0.13	−0.03	0.36	0.02	0.30
	[0.64]	[0.37]	[0.41]	[0.00]	[0.53]	[0.20]	[0.00]	[0.74]	[0.00]
[S3] Combination strategy (ex post optimal)	0.53	0.26	0.50	0.99	0.58	0.64	1.21	0.73	1.11
Difference, [S1]–[S3]	−0.11	−0.12	−0.12	−0.07	−0.14	−0.11	−0.07	−0.03	−0.02
	[0.01]	[0.14]	[0.08]	[0.07]	[0.37]	[0.00]	[0.20]	[0.41]	[0.78]
CER (%)									
[S1] Combination strategy (real time)	1.56	0.00	1.41	8.47	1.96	2.74	12.25	4.19	10.88
[S2] Original factor (real time)	1.75	0.38	1.61	2.29	0.91	3.09	5.44	3.68	6.23
Difference, [S1]–[S2]	−0.19	−0.37	−0.20	6.18	1.04	−0.35	6.81	0.51	4.65
	[0.83]	[0.27]	[0.73]	[0.00]	[0.57]	[0.21]	[0.00]	[0.60]	[0.00]
[S3] Combination strategy (ex post optimal)	2.79	0.67	2.47	9.87	3.42	4.04	14.55	5.36	12.34
Difference, [S1]–[S3]	−1.23	−0.66	−1.06	−1.40	−1.46	−1.30	−2.30	−1.17	−1.46
	[0.01]	[0.13]	[0.10]	[0.07]	[0.39]	[0.03]	[0.15]	[0.25]	[0.30]
Panel B: Real-time combination strategies including Fama and French (1993) three factors									
Sharpe ratio									
[S1] Combination strategy (real time)	0.51	0.50	0.53	1.14	0.83	0.77	1.30	0.94	1.19
[S2] Original factor + FF3 (real time)	0.61	0.61	0.61	0.94	0.85	0.80	1.23	0.97	0.98
Difference, [S1] and [S2]	−0.11	−0.11	−0.08	0.20	−0.02	−0.03	0.07	−0.03	0.20
	[0.22]	[0.03]	[0.31]	[0.00]	[0.85]	[0.12]	[0.23]	[0.10]	[0.00]
[S3] Combination strategy (ex post optimal)	0.72	0.71	0.71	1.28	1.11	0.98	1.63	1.09	1.38
Difference, [S1]–[S3]	−0.22	−0.21	−0.18	−0.14	−0.28	−0.21	−0.33	−0.15	−0.20
	[0.00]	[0.01]	[0.03]	[0.01]	[0.00]	[0.01]	[0.00]	[0.01]	[0.00]
CER (%)									
[S1] Combination strategy (real time)	2.51	2.13	2.72	12.88	6.43	5.54	16.25	8.73	13.70
[S2] Original factor + FF3 (real time)	2.52	2.52	2.52	8.75	6.63	6.07	14.88	9.33	9.66
Difference, [S1] and [S2]	−0.02	−0.40	0.20	4.13	−0.19	−0.53	1.38	−0.60	4.04
	[0.99]	[0.21]	[0.77]	[0.00]	[0.92]	[0.13]	[0.25]	[0.11]	[0.00]
[S3] Combination strategy (ex post optimal)	5.21	5.05	5.09	16.39	12.33	9.60	26.53	11.79	19.11
Difference, [S1]–[S3]	−2.71	−2.93	−2.37	−3.51	−5.89	−4.06	−10.28	−3.06	−5.42
	[0.00]	[0.02]	[0.04]	[0.02]	[0.00]	[0.01]	[0.00]	[0.01]	[0.00]

We use the first K months as the initial training period and evaluate portfolio performance over the subsequent out-of-sample period of $T - K$ months. For our base case results, we specify an initial training sample of $K = 120$ months and employ an expanding-window approach to estimate the relevant portfolio parameters. Our choice of initial training sample length allows for a relatively long out-of-sample evaluation period for each strategy, which alleviates the well-known concern with low power in out-of-sample tests (e.g., Inoue and Kilian, 2004). The expanding-window specification also mitigates estimation risk throughout the evaluation period, as investors are able to make full use of past data in estimating return moments.

We acknowledge that readers can reasonably disagree on the preferred design choices. For example, although a longer initial training period would reduce test power, such a design would result in more precise parameter estimates early in the out-of-sample period. Rolling-window estimation could also be preferable to expanding-window estimation if the likelihood and magnitude of structural breaks in the data generating process are large (e.g., Pesaran and Timmermann, 2002; Rossi, 2013). Ex ante, it is challenging to identify which design choices are most appropriate. As such, we conduct an extensive set of robustness checks to confirm that our conclusions are not sensitive to the specified training sample length or type.

As suggested by Eq. (11), a real-time investor needs estimates of the scaling parameter for the volatility-managed factor and the weights to assign to the volatility-managed and original factors in the optimal combination strategy. At the beginning of each month t in the out-of-sample period, we first compute the real-time scaling parameter, c_t, as the constant that allows the original and volatility-managed factors to have the same variance over the training period preceding month t. We then estimate the vector of mean excess returns ($\hat{\mu}_t$) and the covariance matrix ($\hat{\Sigma}_t$) from the training period and construct portfolio weights according to

$$\begin{bmatrix} x_{\sigma,t} \\ x_t \end{bmatrix} = \frac{1}{\gamma} \hat{\Sigma}_t^{-1} \hat{\mu}_t. \tag{15}$$

The investment position in the original factor is given by the real-time version of Eq. (11):

$$y_t = x_{\sigma,t} \left(\frac{c_t}{\hat{\sigma}_{t-1}^2} \right) + x_t. \tag{16}$$

We construct the portfolio excess return for month t as $y_t f_t$, and the outcome of this approach is a time series of $T - K$ monthly excess returns.

Before proceeding, we note that the magnitude of the investment position in a given real-time strategy (i.e., $|y_t|$) is a measure of leverage. The out-of-sample approach to estimating these positions leads to strategies that require extreme leverage for at least two reasons. First, volatility-managed portfolios, by nature, call for the use of substantial leverage to gain aggressive factor exposures following periods of low volatility. Second, sample-based mean-variance optimization often leads to extreme values for estimates of portfolio weights [i.e., Eq. (15)], particularly when the training period is short (e.g., DeMiguel et al., 2009a; 2009b). For our base case, we impose a leverage constraint of $|y_t| \leq 5$. This design choice reflects our desire to maintain the spirit of volatility-managed strategies while simultaneously guarding against our results being driven by extreme outliers. We also confirm that our conclusions are robust to imposing leverage constraints as low as one, as well as allowing for unconstrained investment positions.

In evaluating the performance of the real-time combination portfolios, we focus on whether or not these strategies lead to improved investment outcomes relative to the original factors. This focus is intentionally practical, as the out-of-sample tests are not directly informative about the underlying risk-return trade-off for the strategies of interest. Let $\hat{\mu}_i$ and $\hat{\sigma}_i$ be the mean and standard deviation of excess returns for the strategy that invests in the volatility-managed portfolio, original portfolio, and risk-free asset (i.e., the real-time combination strategy) over the out-of-sample period of length $T - K$. Similarly, $\hat{\mu}_j$ and $\hat{\sigma}_j$ are the mean and standard deviation of excess returns for the real-time strategy that invests in the original portfolio and risk-free asset, and $\hat{\sigma}_{i,j}$ is the covariance between excess returns for the two strategies. We compute the Sharpe ratio difference as

$$\Delta SR = \frac{\hat{\mu}_i}{\hat{\sigma}_i} - \frac{\hat{\mu}_j}{\hat{\sigma}_j} \tag{17}$$

and assess whether the difference is statistically significant using the test proposed by Jobson and Korkie (1981). We also calculate out-of-sample CER gains from having access to the volatility-scaled factor as

$$\Delta CER = \left(\hat{\mu}_i - \frac{\gamma}{2} \hat{\sigma}_i^2 \right) - \left(\hat{\mu}_j - \frac{\gamma}{2} \hat{\sigma}_j^2 \right). \tag{18}$$

We evaluate statistical significance of the CER difference using the method outlined by DeMiguel et al. (2009b).[19]

4.3.2. Out-of-sample results

Panel A of Table 5 shows results from out-of-sample tests for the nine factors. We compare the performance of the real-time combination strategy and that of the real-time strategy that excludes the volatility-managed portfolio. As a benchmark, we also present the performance of the ex post optimal combination portfolio. The Sharpe ratios and CERs of the strategies in Table 5 are calculated over the out-of-sample evaluation period from month $K + 1$ to month T.

For the base case design in Panel A, the combination portfolios outperform the original factors in five of the nine cases. The differences in Sharpe ratios and CERs are positive for the *MOM, RMW, ROE, IA*, and *BAB* factors and negative for the *MKT, SMB, HML*, and *CMA* factors. The improvements for the *MOM, ROE*, and *BAB* factors are statistically significant at the 1% level. Across the remaining six strategies, the Sharpe ratio and CER differences are insignificant, with the CER differences ranging from −0.37% (*SMB*) to 1.04% (*RMW*) per year.

The market factor is an interesting case, as Moreira and Muir (2017, p. 1618) point out that "this strategy would have been easily available to the average investor in real time." Achieving the gains from volatility management in an out-of-sample setting turns out to be difficult, however, as the combination strategy underperforms relative to the original market portfolio based on Sharpe ratio (0.42 versus 0.46) and CER (1.56% versus 1.75% per year).[20] In the Internet Appendix, we demonstrate why the out-of-sample combination strategy for the market portfolio performs poorly. In particular, we show that the strong in-sample performance for volatility-scaled *MKT* is concentrated in the period surrounding the Great Depression, which occurs early in the sample. Out-of-sample investors adopting the combination strategy tend to favor the volatility-managed version of the market factor based on its strong early sample performance. These investors experience unfavorable investment results, however, because the scaled *MKT*

[19] To test the null hypothesis of equal CERs for strategies i and j, we compute the following test statistic, which is asymptotically distributed as a standard normal: $\hat{z}_{DGU} = \frac{(\hat{\mu}_i - \frac{\gamma}{2}\hat{\sigma}_i^2) - (\hat{\mu}_j - \frac{\gamma}{2}\hat{\sigma}_j^2)}{\sqrt{\hat{\theta}}}$, where $\hat{\theta} = \frac{1}{T-K}\left(\begin{bmatrix} 1 & -1 & -\gamma/2 & \gamma/2 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_i^2 & \hat{\sigma}_{i,j} & 0 & 0 \\ \hat{\sigma}_{i,j} & \hat{\sigma}_j^2 & 0 & 0 \\ 0 & 0 & 2\hat{\sigma}_i^4 & 2\hat{\sigma}_{i,j}^2 \\ 0 & 0 & 2\hat{\sigma}_{i,j}^2 & 2\hat{\sigma}_j^4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -\gamma/2 \\ \gamma/2 \end{bmatrix} \right)$.

[20] The real-time analysis for the market factor is similar in spirit to Johannes et al.'s (2014) comparison of the out-of-sample performance of a "stochastic variance-constant mean" model with that of a "constant variance-constant mean" model. They adopt a Bayesian learning approach to optimal allocation that accounts for parameter uncertainty and find that the model specification with stochastic volatility leads to better Sharpe ratios and CERs. Consistent with our evidence on the market factor in Table 5, however, the performance differences tend to be statistically insignificant.

factor underperforms the original *MKT* factor after the first ten years of the sample.

Table 5 also compares Sharpe ratios and CERs for the real-time combination strategies and the ex post optimal combination strategies. These results provide insights into how much the real-time portfolio formation approach leads to a deterioration in performance relative to the in-sample exercise. All nine of the Sharpe ratio and CER differences are negative, and four of the differences for each of the two measures are statistically significant at the 10% level.

Moreira and Muir (2017) provide evidence that volatility-managed factors also tend to produce positive alphas in spanning regressions that include the original factors as well as the *MKT, SMB*, and *HML* factors from the Fama and French (1993) three-factor model. The interpretation of a positive alpha from this test is that an investor who holds the ex post optimal combination of the benchmark factors increases her Sharpe ratio by adding a positive position in the volatility-managed factor. From an economic perspective, including the Fama and French (1993) factors as controls likely provides a better characterization of the investment opportunity set for investors sophisticated enough to consider volatility-managed strategies. As in the univariate spanning tests, however, the implied optimal strategies are not implementable in real time. Optimizing across a larger set of assets may also exacerbate estimation risk.

We examine the real-time performance of strategies that include the Fama and French (1993) three factors in Panel B of Table 5. We report Sharpe ratios and CERs for real-time strategies that either exclude or include the volatility-managed factor. The results indicate that incorporating volatility-managed strategies into the portfolio choice problem is often harmful to investors with the extended investment opportunity set. Based on Sharpe ratio comparisons the strategies that include volatility-managed portfolios underperform for six of the nine factors. The *MOM* and *BAB* factors produce significantly positive Sharpe ratio differences, but the volatility-managed *SMB* and *IA* factors significantly hurt performance. The nine factors produce four positive and five negative CER differences, indicating that volatility timing does not systematically improve performance for investors with access to the Fama and French (1993) three factors. For both the Sharpe ratio and CER comparisons, all nine of the real-time combination strategies exhibit statistically significant underperformance at the 5% level relative to the ex post optimal combination strategies.

In summary, Table 5 shows that volatility management has potential benefits for real-time investors in some factors, but the gains are not systematic and are much less impressive than the corresponding in-sample results. These initial results indicate that real-time implementation issues degrade portfolio performance in the volatility-managed portfolios setting. Panel B of Table 5 also highlights that, relative to the results in Panel A with the investment universe restricted to a single factor, volatility-managed portfolios are even less useful to investors with access to market, size, and value strategies.

To assess whether our conclusions for the nine factors generalize, we turn to the broader sample of 103 trading strategies. Table 6 summarizes the results for out-of-sample tests using our base case design with an expanding, ten-year training sample, a leverage constraint of $|y_t| \leq 5$, and a risk aversion parameter of $\gamma = 5$. We report in Panel A.1 the number of positive and negative differences in Sharpe ratios and CERs. As in Table 5, a positive performance difference indicates that the real-time combination strategy earns a higher Sharpe ratio or CER than the alternative strategy does. We also show the number of significant differences at the 5% level. Panel A.2 presents separate results for the 9 factors and 94 anomaly portfolios, and Panel A.3 breaks the results down by trading strategy type.

The evidence in Panel A of Table 6 suggests that the real-time combination strategies tend to underperform the real-time strategies that exclude the volatility-managed portfolios. Only 45 of the 103 combination portfolios outperform the original portfolios based on Sharpe ratio. Just 31 outperform based on CER. Further, the number of significantly positive CER differences (seven) exactly matches the number of significantly negative ones. All but one of the 103 real-time combination strategies exhibit underperformance relative to the ex post optimal strategies based on Sharpe ratio, and all 103 underperform based on CER.[21] A large proportion of the performance differences relative to the ex post optimal strategies are statistically significant (39 out of 103 based on Sharpe ratio and 41 out of 103 based on CER).

Panel B presents results for real-time trading strategies based on an investment opportunity set that includes the Fama and French (1993) factors. The conclusions are similar to those in Panel A. The real-time strategies that include volatility-managed portfolios outperform those that exclude them in just 32 out of 103 cases based on either Sharpe ratio or CER. The number of significantly positive results is also small in each case (two for Sharpe ratio and three for CER). The majority of the real-time combination strategies also exhibit statistically significant underperformance relative to their ex post optimal versions.

Table 7 summarizes results for robustness tests that modify the base case out-of-sample design. For each design and performance measure, we report the number of positive and negative differences for the real-time strategies and the number of these differences that are statistically significant at the 5% level. We also report the *p*-value from a binomial test of the null hypothesis that positive and negative performance differences are equally likely. The first row in Panel A of Table 7 repeats the results from our base case design in Panel A of Table 6. The remaining rows in this panel display robustness results pertaining to the training sample type, risk aversion parameter, training sample length, and leverage constraint. None of the alternative specifications meaningfully improves performance relative to the base case.

[21] The ex post optimal strategy is based on a fixed scaling parameter for the volatility-managed portfolio and fixed portfolio weights such that it is possible for a real-time strategy with time-varying portfolio weights to outperform the ex post optimal strategy.

Table 6
Summary of real-time combination strategies: broad sample.
The table summarizes results for real-time strategies that combine original and volatility-managed versions of 103 trading strategies. For each factor or anomaly portfolio, we compute the Sharpe ratio and certainty equivalent return (CER) for (i) the real-time strategy that combines the original portfolio, the volatility-managed portfolio, and the risk-free asset ("combination strategy (real time)"), (ii) the real-time strategy that combines the original portfolio and the risk-free asset ("original factor (real time)"), and (iii) the ex post optimal strategy that combines the original portfolio, the volatility-managed portfolio, and the risk-free asset ("combination strategy (ex post optimal)"). Columns (3) and (4) in Panel A summarize the Sharpe ratio differences for these strategies, and columns (5) and (6) summarize the CER differences. Panel B summarizes the Sharpe ratio and CER differences for analogous strategies that include the Fama and French (1993) factors. A positive Sharpe ratio or CER difference indicates outperformance for the real-time combination strategy. The out-of-sample tests are based on a ten-year training period and an expanding-window design. We use a risk aversion parameter of $\gamma = 5$ and impose a leverage constraint that the sum of absolute weights on the risky assets is less than or equal to five. Panels A.1 and B.1 report results for the full set of 103 trading strategies. Panels A.2 and B.2 present separate results for the 9 factors and the 94 anomaly portfolios. Panels A.3 and B.3 break the results down by trading strategy type. For each set of comparisons, the table reports the number of differences that are positive (+), positive and significant at the 5% level (in brackets), negative (−), and negative and significant at the 5% level (in brackets). We assess statistical significance of the Sharpe ratio and CER differences using the approaches in Jobson and Korkie (1981) and DeMiguel et al. (2009b), respectively.

Panel A: Real-time combination strategies

Sample	Total	Sharpe ratio difference: Combination strategy (real time) versus		CER difference: Combination strategy (real time) versus	
		Original factor (real time)	Combination strategy (ex post optimal)	Original factor (real time)	Combination strategy (ex post optimal)
		ΔSR +/−	ΔSR +/−	ΔCER +/−	ΔCER +/−
(1)	(2)	(3)	(4)	(5)	(6)
Panel A.1: Combined sample					
All trading strategies	103	45 [8] / 58 [2]	1 [0] / 102 [39]	31 [7] / 72 [7]	0 [0] / 103 [41]
Panel A.2: By category					
Factors	9	5 [3] / 4 [0]	0 [0] / 9 [2]	5 [3] / 4 [0]	0 [0] / 9 [2]
Anomaly portfolios	94	40 [5] / 54 [2]	1 [0] / 93 [37]	26 [4] / 68 [7]	0 [0] / 94 [39]
Panel A.3: By trading strategy type					
Accruals	10	3 [0] / 7 [1]	0 [0] / 10 [5]	3 [0] / 7 [2]	0 [0] / 10 [4]
Intangibles	10	4 [0] / 6 [0]	0 [0] / 10 [0]	1 [0] / 9 [1]	0 [0] / 10 [4]
Investment	11	5 [0] / 6 [0]	0 [0] / 11 [6]	5 [0] / 6 [0]	0 [0] / 11 [5]
Market	1	0 [0] / 1 [0]	0 [0] / 1 [1]	0 [0] / 1 [0]	0 [0] / 1 [1]
Momentum	9	8 [4] / 1 [0]	0 [0] / 9 [5]	8 [5] / 1 [0]	0 [0] / 9 [5]
Profitability	22	10 [1] / 12 [0]	1 [0] / 21 [7]	6 [1] / 16 [1]	0 [0] / 22 [5]
Trading	21	10 [1] / 11 [1]	0 [0] / 21 [6]	6 [1] / 15 [1]	0 [0] / 21 [8]
Value	19	5 [2] / 14 [0]	0 [0] / 19 [9]	2 [0] / 17 [2]	0 [0] / 19 [9]

Panel B: Real-time combination strategies including Fama and French (1993) three factors

Sample	Total	Sharpe ratio difference: Combination strategy (real time) versus		CER difference: Combination strategy (real time) versus	
		Original factor+FF3 (real time)	Combination strategy (ex post optimal)	Original factor+FF3 (real time)	Combination strategy (ex post optimal)
		ΔSR +/−	ΔSR +/−	ΔCER +/−	ΔCER +/−
(1)	(2)	(3)	(4)	(5)	(6)
Panel B.1: Combined sample					
All trading strategies	103	32 [2] / 71 [13]	0 [0] / 103 [77]	32 [3] / 71 [10]	0 [0] / 103 [92]
Panel B.2: By category					
Factors	9	3 [2] / 6 [1]	0 [0] / 9 [9]	4 [2] / 5 [0]	0 [0] / 9 [9]
Anomaly portfolios	94	29 [0] / 65 [12]	0 [0] / 94 [68]	28 [1] / 66 [10]	0 [0] / 94 [83]
Panel B.3: By trading strategy type					
Accruals	10	3 [0] / 7 [1]	0 [0] / 10 [8]	3 [0] / 7 [1]	0 [0] / 10 [8]
Intangibles	10	1 [0] / 9 [1]	0 [0] / 10 [5]	1 [0] / 9 [1]	0 [0] / 10 [9]
Investment	11	3 [0] / 8 [1]	0 [0] / 11 [10]	3 [0] / 8 [1]	0 [0] / 11 [10]
Market	1	0 [0] / 1 [0]	0 [0] / 1 [1]	0 [0] / 1 [0]	0 [0] / 1 [1]
Momentum	9	8 [1] / 1 [0]	0 [0] / 9 [9]	7 [2] / 2 [0]	0 [0] / 9 [9]
Profitability	22	8 [0] / 14 [2]	0 [0] / 22 [14]	9 [0] / 13 [2]	0 [0] / 22 [18]
Trading	21	6 [1] / 15 [5]	0 [0] / 21 [21]	5 [1] / 16 [3]	0 [0] / 21 [21]
Value	19	3 [0] / 16 [3]	0 [0] / 19 [9]	4 [0] / 15 [2]	0 [0] / 19 [16]

Table 7
Summary of real-time combination strategies: robustness tests.
The table summarizes robustness tests for real-time strategies that combine original and volatility-managed versions of 103 trading strategies. For each portfolio, we compute the Sharpe ratio and certainty equivalent return (CER) for (i) the real-time strategy that combines the original portfolio, the volatility-managed portfolio, and the risk-free asset and (ii) the real-time strategy that combines the original portfolio and the risk-free asset. Column (2) in Panel A summarizes the Sharpe ratio differences for these strategies, and column (4) summarizes the CER differences. Panel B summarizes the Sharpe ratio and CER differences for analogous strategies that include the Fama and French (1993) factors. For each set of comparisons, the table reports the number of differences that are positive (+), positive and significant at the 5% level (in brackets), negative (−), and negative and significant at the 5% level (in brackets). A positive Sharpe ratio or CER difference indicates outperformance for the real-time combination strategy that includes the volatility-managed portfolio. We assess statistical significance of the Sharpe ratio and CER differences using the approaches in Jobson and Korkie (1981) and DeMiguel et al. (2009b), respectively. For each set of comparisons, we also present the two-tailed p-value from a binomial distribution test of the null hypothesis that each of the differences is equally likely to be positive or negative. The first row in each panel summarizes results from the base case design in Table 6. This design is characterized by (i) an expanding-window training sample, (ii) an initial training sample length of $K = 120$ months, (iii) a risk aversion parameter of $\gamma = 5$, and (iv) a leverage constraint of $L \leq 5$. The subsequent rows in each panel are based on modified designs as described in the table.

	Sharpe ratio difference: ($N = 103$)		CER difference: ($N = 103$)	
Description (1)	ΔSR +/− (2)	Binomial p-value (3)	ΔCER +/− (4)	Binomial p-value (5)
Panel A: Real-time combination strategies				
Base case design	45 [8] / 58 [2]	0.237	31 [7] / 72 [7]	0.000
Rolling-window training sample	49 [2] / 54 [1]	0.694	17 [1] / 86 [19]	0.000
Risk aversion, $\gamma = 2$	48 [9] / 55 [2]	0.555	35 [8] / 68 [7]	0.001
Risk aversion, $\gamma = 10$	45 [8] / 58 [2]	0.237	31 [7] / 72 [7]	0.000
Initial training sample length, $K = 240$	45 [9] / 58 [10]	0.237	36 [8] / 67 [11]	0.003
Initial training sample length, $K = 360$	40 [9] / 63 [6]	0.030	31 [8] / 72 [8]	0.000
Leverage constraint, $L \leq 1.0$	49 [10] / 54 [2]	0.694	38 [4] / 65 [5]	0.010
Leverage constraint, $L \leq 1.5$	47 [10] / 56 [3]	0.431	38 [7] / 65 [7]	0.010
Leverage constraint, $L \leq \infty$	45 [8] / 58 [2]	0.237	31 [7] / 72 [7]	0.000
Panel B: Real-time combination strategies including Fama and French (1993) three factors				
Base case design	32 [2] / 71 [13]	0.000	32 [3] / 71 [10]	0.000
Rolling-window training sample	32 [0] / 71 [8]	0.000	20 [0] / 83 [16]	0.000
Risk aversion, $\gamma = 2$	22 [3] / 81 [13]	0.000	24 [3] / 79 [10]	0.000
Risk aversion, $\gamma = 10$	31 [3] / 72 [12]	0.000	32 [3] / 71 [10]	0.000
Initial training sample length, $K = 240$	31 [6] / 72 [11]	0.000	35 [9] / 68 [10]	0.001
Initial training sample length, $K = 360$	30 [6] / 73 [9]	0.000	28 [7] / 75 [9]	0.000
Leverage constraint, $L \leq 1.0$	22 [2] / 81 [11]	0.000	27 [3] / 76 [12]	0.000
Leverage constraint, $L \leq 1.5$	21 [3] / 82 [13]	0.000	26 [2] / 77 [9]	0.000
Leverage constraint, $L \leq \infty$	31 [3] / 72 [12]	0.000	32 [3] / 71 [11]	0.000

The robustness design with rolling-window parameter estimation leads to a slightly larger number of positive Sharpe ratio differences but a substantially smaller number of positive CER differences. Using a lower ($\gamma = 2$) or higher ($\gamma = 10$) risk aversion parameter leads to almost identical results to the base case with $\gamma = 5$.

One potential concern with the ten-year training sample in our base case is the relatively small number of observations used to estimate portfolio positions early in the out-of-sample period. We therefore consider specifications with 20-year ($K = 240$) and 30-year ($K = 360$) initial estimation periods. These designs produce roughly the same number of positive Sharpe ratio and CER differences that the base case does. If anything, these robustness results are less favorable for volatility management, as the number of significantly negative performance differences is much higher in the $K = 240$ case.

The final three rows in Panel A of Table 7 detail the impact of alternative leverage constraints. Imposing a leverage constraint could either improve performance if real-time investors avoid taking extreme positions or hurt performance if the constraint prevents investors from capitalizing on the information content in lagged volatility. We consider tighter leverage constraints of $|y_t| \leq 1.0$ and $|y_t| \leq 1.5$ as well as an unconstrained specification, but the performance of the combination strategies does not systematically improve for these cases.

In summary, our main conclusions from Panel A of Table 6 continue to hold across the robustness results in Panel A of Table 7. Incorporating volatility-managed portfolios into the real-time portfolio decision tends to harm performance. More than half of the Sharpe ratio and CER differences are negative under each specification. The CER results, in particular, generate strong statistical evidence against using volatility-managed portfolios in real-time applications. All nine of the binomial p-values for the CER differences produce statistical rejections of the null hypothesis of equal performance at the 5% significance level.

Finally, Panel B of Table 7 summarizes our robustness tests for specifications that include the Fama and French (1993) three factors. Each of the robustness specifications produces similar evidence against volatility management, and all 18 binomial p-values reject the null hypothesis of equal performance differences at the 1% level. Overall, our findings in this section indicate that attempting to use volatility management in real time tends to degrade performance relative to constraining the investment universe to the original factors and anomaly portfolios.

Table 8
Comparison of volatility-managed strategies with traditional anomaly strategies: broad sample.
Panel A summarizes results on the performance of volatility-managed versions of 103 trading strategies. The spanning regression tests are described in Table 4, and the tests for Sharpe ratio differences and certainty equivalent return (CER) differences are described in Table 6. Panel B summarizes results from analogous tests for 102 anomaly-based trading strategies. The in-sample tests are CAPM regressions and Fama and French (1993) three-factor regressions. For each set of regressions, the table reports the number of alphas that are positive (+), positive and significant at the 5% level (in brackets), negative (−), and negative and significant at the 5% level (in brackets). We assess statistical significance of the alpha estimates using White (1980) standard errors. For the out-of-sample tests for a given anomaly portfolio, we compute the Sharpe ratio and CER for (i) the real-time strategy that combines the anomaly portfolio, the market portfolio, and the risk-free asset and (ii) the real-time strategy that combines the market portfolio and the risk-free asset. We also compute Sharpe ratios and CERs for analogous strategies that include the Fama and French (1993) factors. Column (5) in Panel B summarizes the Sharpe ratio differences for these strategies, and column (7) summarizes the CER differences. A positive Sharpe ratio or CER difference indicates outperformance for the real-time strategy that includes the anomaly portfolio. The out-of-sample tests are based on a ten-year training period and an expanding-window design. We use a risk aversion parameter of $\gamma = 5$ and impose a leverage constraint that the sum of absolute weights on the risky assets is less than or equal to five. For each set of comparisons, the table reports the number of differences that are positive (+), positive and significant at the 5% level (in brackets), negative (−), and negative and significant at the 5% level (in brackets). We assess statistical significance of the Sharpe ratio and CER differences using the approaches in Jobson and Korkie (1981) and DeMiguel et al. (2009b), respectively. For each set of comparisons, we also present the two-tailed *p*-value from a binomial distribution test of the null hypothesis that each of the performance measures is equally likely to be positive or negative.

		Alpha:		Sharpe ratio difference:		CER difference:	
Description (1)	Total (2)	α +/− (3)	Binomial *p*-value (4)	ΔSR +/− (5)	Binomial *p*-value (6)	ΔCER +/− (7)	Binomial *p*-value (8)
Panel A: Spanning regressions							
Spanning regressions	103	77 [23] / 26 [3]	0.000	45 [8] / 58 [2]	0.237	31 [7] / 72 [7]	0.000
Spanning regressions with FF3 controls	103	70 [21] / 33 [3]	0.000	32 [2] / 71 [13]	0.000	32 [3] / 71 [10]	0.000
Panel B: Anomaly regressions							
CAPM regressions	102	93 [73] / 9 [3]	0.000	75 [19] / 27 [1]	0.000	68 [18] / 34 [1]	0.001
FF3 regressions	100	81 [60] / 19 [5]	0.000	66 [18] / 34 [1]	0.002	55 [19] / 45 [1]	0.368

4.3.3. Explanations for poor out-of-sample performance

The tests in Section 4.3.2 suggest that, in the volatility-managed portfolios setting, strong in-sample performance metrics often fail to translate into real-time gains for investors. Based on this evidence, it is natural to explore the economic drivers of these results and examine whether or not our findings generalize to more familiar investment settings. As a starting point for this analysis, we compare the performance of the combination strategies based on volatility-managed portfolios with the performance of traditional anomaly strategies.

Panel A of Table 8 reproduces the main results from Tables 4 and 6 on the in-sample and out-of-sample performance of combination strategies. As described above, these tests focus on the value of including a volatility-managed version of a given portfolio in the investment opportunity set. In the broad sample of 103 trading strategies, positive spanning regression alphas are common, but real-time investors typically earn lower Sharpe ratios and CERs under the expanded investment opportunity set.

Panel B of Table 8 shows results from analogous tests for traditional anomaly strategies. Studies showing cross-sectional anomalies routinely emphasize the alphas earned by these strategies relative to popular asset pricing models such as the Capital Asset Pricing Model (CAPM) or Fama and French (1993) three-factor model. We replicate these types of tests in our broad sample and find that 93 out of 102 portfolios earn positive CAPM alphas, and 81 out of 100 earn positive three-factor alphas.[22] A large proportion of these in-sample alpha estimates are statistically significant. A positive alpha relative to a given factor model implies that the ex post optimal combination of anomaly portfolio and benchmark factors expands the mean-variance frontier relative to the ex post optimal combination of the benchmark factors. As with the volatility-managed portfolios setting, however, real-time investors must construct their portfolios using prior data. For each anomaly portfolio and benchmark model, we conduct an out-of-sample exercise that compares the performance of two strategies: (i) the real-time strategy that combines the anomaly portfolio, the benchmark factors, and the risk-free asset and (ii) the real-time strategy that combines the benchmark factors and the risk-free asset. The out-of-sample design parameters are identical to those introduced in Section 4.3.1.

Although we do see out-of-sample performance degradation in Panel B of Table 8, the effects are much less acute relative to those seen for the volatility-managed portfolios in Panel A. Adding an anomaly portfolio to the CAPM market factor in real time, for example, leads to a Sharpe ratio improvement in 75 out of 102 cases and a CER improvement in 68 out of 102 cases. Real-time performance relative to the Fama and French (1993) three-factor benchmark is less impressive, with positive Sharpe ratio differences for 66 out of 100 strategies and positive CER differences for 55 out of 100 strategies. Nonetheless, these results indicate that real-time anomaly strategies fare substantially better compared with real-time volatility-managed strategies.

The results in Table 8 provide a useful backdrop to examine why the statistical support for out-of-sample combination strategies is particularly weak in the volatility-managed portfolios setting. We consider

[22] The total number of strategies considered in Panel B is less than 103 because we exclude the benchmark factors from the analysis for each model.

Table 9

Summary of multiple structural break tests for spanning regressions and anomaly regressions: broad sample.

Panel A summarizes results from structural break tests in spanning regressions of volatility-managed anomaly portfolio returns on original anomaly portfolio returns. We consider univariate spanning regressions and spanning regressions with additional controls for the Fama and French (1993) three factors. The tests follow Bai and Perron (1998, 2003) and allow for an unknown number of structural breaks. Column (2) reports the total number of volatility-managed strategies considered. For each set of regressions, columns (3)–(7) present the frequency distribution for the number of identified breaks (N_b), and column (8) reports the mean number of breaks ($\bar{N}_b$). Panel B summarizes results from structural break tests in regressions of traditional anomaly portfolio returns on factor returns. We consider CAPM and Fama and French (1993) three-factor regressions.

		Frequency distribution for breaks					
Description (1)	Total (2)	$N_b = 0$ (3)	$N_b = 1$ (4)	$N_b = 2$ (5)	$N_b = 3$ (6)	$N_b \geq 4$ (7)	$\bar{N}_b$ (8)
Panel A: Spanning regressions							
Spanning regressions	103	0	10	52	34	7	2.37
Spanning regressions with FF3 controls	103	1	8	53	35	6	2.37
Panel B: Anomaly regressions							
CAPM regressions	102	15	38	39	9	1	1.44
FF3 regressions	100	10	25	36	21	8	1.92

three potential explanations: (i) estimation risk in the out-of-sample portfolio choice exercise, (ii) low power in the out-of-sample tests, and (iii) structural instability in the conditional risk-return trade-off for the various factors and anomaly portfolios.

A known concern with out-of-sample portfolio optimization is estimation risk. DeMiguel et al. (2009b), for example, note that optimal portfolios constructed from sample moments often exhibit extreme weights that fluctuate dramatically over time. Intuitively, it can be difficult to reliably estimate asset return moments with short training periods, and these moments are the key determinants of portfolio weights [e.g., Eq. (15)]. Although estimation error is always a challenge with real-time portfolio choice applications, we are skeptical that it fully accounts for our results for a variety of reasons. First, our empirical design incorporates several features intended to mitigate estimation risk, including a leverage constraint on portfolio positions, a risk-free asset in the investment opportunity set (Kirby and Ostdiek, 2012), and expanding-window parameter estimation. Second, DeMiguel et al. (2009b) emphasize that estimation risk is less problematic in applications, like ours, in which the number of test assets is small. Third, our main results are based on comparisons of real-time strategies that include volatility-managed portfolios in the investment opportunity set with those that exclude volatility-managed portfolios from the investment opportunity set. Thus, both the combination strategy and the benchmark suffer from estimation risk, and it is not obvious why one of the two would be more adversely impacted. Fourth, if estimation risk is the primary explanation of the poor performance of the combination strategies, then we should see more favorable results under specifications with longer training samples. Table 7 reveals, however, that lengthening the initial training sample has little impact on our conclusions. Finally, Panel B of Table 8 provides direct evidence that in-sample alphas do translate into improved real-time performance measures much more frequently outside of the volatility-managed portfolios setting.

Another common concern with out-of-sample tests is that they lack power relative to in-sample tests because the evaluation period is shorter (e.g., Inoue and Kilian, 2004). Our focus on assessing the value of volatility management for real-time investors necessitates the use of out-of-sample tests. Low power also does not seem to be a satisfactory explanation for our results. If volatility management is systematically beneficial to investors, then we should see a majority of performance differences that are positive in Tables 6 and 7. Low power might be an explanation for why an individual result is statistically insignificant, but it does not account for why most of the performance differences have the wrong sign.

A more plausible economic explanation for the poor out-of-sample performance for the combination strategies is structural instability in the spanning regression parameters from Eq. (5) and the implied optimal weights. We investigate this issue formally using Bai and Perron (1998, 2003) structural break tests. Table 9 summarizes results from these tests for the spanning regressions and the traditional anomaly regressions.[23] In Panel A, we find strong statistical evidence of structural breaks in the spanning tests for the 103 volatility-managed portfolios. For the univariate spanning regressions, none of the tests suggests zero breaks, and 41 out of 103 tests identify three or more breaks. The average number of breaks is 2.37 for both the univariate spanning regressions and the spanning regressions that control for the Fama and French (1993) factors. In contrast, structural breaks are less common in the standard time-series anomaly regressions in Panel B. In the CAPM regressions, for example, 53 out of 102 strategies have one break or less, and the average number of breaks is 1.44. From an economic perspective, structural breaks are direct evidence of instability in the underlying regression parameters and the associated optimal portfolio weights [e.g., Eq. (10)]. In the volatility-managed portfolios setting, the prevalence of breaks often works to the detriment of real-time investors who rely on past data in portfolio construction.

[23] The Internet Appendix provides additional information on the design of these tests and presents detailed results from structural break tests for the nine equity factors.

5. Conclusion

Recent literature suggests that investors can enhance Sharpe ratios and lifetime utility by adopting simple trading rules that scale positions in popular equity portfolios by lagged variance. The trading strategies implied by these studies typically take one of two forms: direct investments in volatility-managed portfolios or combination portfolios that invest in both the volatility-managed version and the original version of an underlying strategy. We show that neither of these methods suggests a pervasive link between volatility management and improved performance for real-time investors.

Direct investments in volatility-scaled strategies are straightforward to implement in real time, and studies following this approach offer compelling empirical evidence that these dynamic portfolios are superior to their static versions. The evidence is isolated to a handful of strategies (e.g., the market, momentum, and betting-against-beta factors), however, making it difficult to draw broad conclusions. We fill this gap by conducting a comprehensive empirical investigation of volatility-managed portfolios. Across a broad sample of 103 equity portfolios, volatility management degrades and improves performance at about the same frequency. From a practical perspective, the results suggest that direct investments in volatility-managed portfolios are not a panacea of improved performance. From an economic perspective, the roughly equal split between positive and negative performance differences is suggestive of a generally positive risk-return trade-off for the individual factors and anomaly portfolios.

Combination strategies that incorporate volatility management, in contrast, exhibit systematically strong in-sample performance. On this point, we extend Moreira and Muir's (2017) spanning regression analysis to our broader set of 103 equity strategies and show that these portfolios tend to exhibit positive alphas. We also demonstrate, however, that structural instability in the spanning regression parameters limits the appeal of this approach to investors conditioning their portfolios on real-time information. The Sharpe ratios and CERs for the out-of-sample combination portfolios are dramatically less impressive than those earned by their in-sample versions. Moreover, the real-time combination strategies routinely underperform simpler strategies constrained to invest in the original, unscaled portfolios.

Appendix A

This appendix provides details on the construction of the anomaly portfolios used in our empirical tests.

As described in Section 2.1, we examine 94 anomaly variables from Hou et al. (2015) and McLean and Pontiff (2016). Our list of firm characteristics includes variables from these studies that are continuous and can be constructed from CRSP, Compustat, and IBES data. We exclude predictors that are based on industry-level variables. Table A1 presents a brief description of the firm characteristics and notes the original study documenting each corresponding anomaly. We construct the anomaly variables following the descriptions provided by Hou et al. (2015) and McLean and Pontiff (2016), and column (4) of Table A1 gives the relevant sources (i.e., "HXZ" or "MP").

The sample includes NYSE, Amex, and Nasdaq ordinary common stocks with return data available on the CRSP monthly and daily stock files for the period from July 1926 to December 2016. When a firm is delisted from an exchange during a given month, we replace any missing returns with the delisting returns provided by CRSP. For a given anomaly variable, we sort firms periodically into ten groups and construct value-weighted portfolios. Our tests focus on the corresponding hedge portfolio that takes a long (short) position in the decile of stocks that is expected to outperform (underperform) based on prior literature.

The portfolios exclude financial stocks (SIC codes 6000–6999) and firms with market capitalization below the first NYSE decile or share price less than $5 at the portfolio formation date. To ensure that accounting data are known prior to the returns they are used to forecast,

Table A1

Anomaly variables.

The table summarizes the firm characteristics used to construct the long-short anomaly decile portfolios in the paper. The panels of the table are organized by anomaly type (i.e., accruals, intangibles, investment, momentum, profitability, trading, and value). For each characteristic, we provide a symbol and brief description and note the original study documenting the corresponding anomaly. We construct the anomaly variables following the descriptions provided by Hou et al. (2015) and McLean and Pontiff (2016) and the relevant source (i.e., "HXZ" or "MP") for a given anomaly is listed in the table. For each anomaly variable, the table also reports the start of the sample period for portfolio returns and the number of monthly return observations.

Anomaly (1)	Description (2)	Original study (3)	Source (4)	Start (5)	Number (6)
		Panel A: Accruals			
IvC	Inventory changes	Thomas and Zhang (2002)	HXZ	1963:08	641
IvG	Inventory growth	Belo and Lin (2012)	HXZ	1963:08	641
NOA	Net operating assets	Hirshleifer et al. (2004)	HXZ	1963:08	641
OA	Operating accruals	Sloan (1996)	HXZ	1963:08	641
POA	Percent operating accruals	Hafzalla et al. (2011)	HXZ	1963:08	641
PTA	Percent total accruals	Hafzalla et al. (2011)	HXZ	1963:08	641
TA	Total accruals	Richardson et al. (2005)	HXZ	1963:08	641
ΔNCO	Changes in net noncurrent operating assets	Soliman (2008)	MP	1963:08	641
ΔNWC	Changes in net noncash working capital	Soliman (2008)	MP	1963:08	641
NoaG	Growth in net operating assets minus accruals	Fairfield et al. (2003)	MP	1963:08	641

(continued on next page)

Table A1 (*continued*)

Anomaly (1)	Description (2)	Original study (3)	Source (4)	Start (5)	Number (6)
		Panel B: Intangibles			
AccQ	Accrual quality	Francis et al. (2005)	HXZ	1966:08	605
AD/M	Advertisement expense-to-market	Chan et al. (2001)	HXZ	1974:08	509
BC	Brand capital investment rate	Belo et al. (2014b)	HXZ	1980:08	437
H/N	Hiring rate	Belo et al. (2014a)	HXZ	1963:08	641
OC/A	Organizational capital-to-assets	Eisfeldt and Papanikolaou (2013)	HXZ	1963:08	641
OL	Operating leverage	Novy-Marx (2011)	HXZ	1963:08	641
RC/A	R&D capital-to-assets	Li (2011)	HXZ	1980:08	437
RD/M	R&D-to-market	Chan et al. (2001)	HXZ	1976:08	485
RD/S	R&D-to-sales	Chan et al. (2001)	HXZ	1976:08	485
Age	Firm age	Barry and Brown (1984)	MP	1963:08	641
		Panel C: Investment			
ΔPI/A	Changes in PP&E plus changes in inventory	Lyandres et al. (2008)	HXZ	1963:08	641
ACI	Abnormal corporate investment	Titman et al. (2004)	HXZ	1966:08	605
CEI	Composite issuance	Daniel and Titman (2006)	HXZ	1931:08	1025
I/A	Investment-to-assets	Cooper et al. (2008)	HXZ	1963:08	641
IG	Investment growth	Xing (2008)	HXZ	1963:08	641
NSI	Net stock issues	Pontiff and Woodgate (2008)	HXZ	1963:08	641
NXF	Net external financing	Bradshaw et al. (2006)	HXZ	1974:08	509
BeG	Growth in book equity	Lockwood and Prombutr (2010)	MP	1963:08	641
I-ADJ	Industry-adjusted growth in investment	Abarbanell and Bushee (1998)	MP	1965:08	617
		Panel D: Momentum			
Abr-1	Abnormal stock returns around earnings announcements	Chan et al. (1996)	HXZ	1974:08	509
R11-1	Price momentum (11-month prior returns)	Fama and French (1996)	HXZ	1927:08	1073
R6-1	Price momentum (6-month prior returns)	Jegadeesh and Titman (1993)	HXZ	1926:09	1084
RE-1	Revisions in analysts' earnings forecasts	Chan et al. (1996)	HXZ	1976:08	485
SUE-1	Earnings surprise	Foster et al. (1984)	HXZ	1976:08	485
R6-Lag	Lagged momentum	Novy-Marx (2012)	MP	1927:08	1073
Season	Seasonality	Heston and Sadka (2008)	MP	1946:08	845
W52	52-week high	George and Hwang (2004)	MP	1927:08	1073
		Panel E: Profitability			
ATO	Asset turnover	Soliman (2008)	HXZ	1963:08	641
CTO	Capital turnover	Haugen and Baker (1996)	HXZ	1963:08	641
F	*F*-score	Piotroski (2000)	HXZ	1974:08	509
FP	Failure probability	Campbell et al. (2008)	HXZ	1976:08	485
GP/A	Gross profitability-to-assets	Novy-Marx (2013)	HXZ	1963:08	641
O	*O*-score	Dichev (1998)	HXZ	1963:08	641
PM	Profit margin	Soliman (2008)	HXZ	1963:08	641
RNA	Return on net operating assets	Soliman (2008)	HXZ	1963:08	641
ROA	Return on assets	Balakrishnan et al. (2010)	HXZ	1974:08	509
ROE-HB	Return on equity	Haugen and Baker (1996)	HXZ	1974:08	509
RS	Revenue surprise	Jegadeesh and Livnat (2006)	HXZ	1976:08	485
TES	Tax expense surprise	Thomas and Zhang (2011)	HXZ	1976:08	485
TI/BI	Taxable income-to-book income	Green et al. (2017)	HXZ	1963:08	641
ΔATO	Change in asset turnover	Soliman (2008)	MP	1963:08	641
ΔPM	Change in profit margin	Soliman (2008)	MP	1963:08	641
E-con	Earnings consistency	Alwathainani (2009)	MP	1971:08	545
S/IV	Change in sales minus change in inventory	Abarbanell and Bushee (1998)	MP	1963:08	641
S/P	Sales-to-price	Barbee, Jr. et al. (1996)	MP	1963:08	641
S/SG&A	Change in sales minus change in SG&A	Abarbanell and Bushee (1998)	MP	1963:08	641
Z	*Z*-score	Dichev (1998)	MP	1963:08	641
		Panel F: Trading			
β-D	Dimson's beta (daily data)	Dimson (1979)	HXZ	1926:09	1084
β-FP	Frazzini and Pedersen's beta	Frazzini and Pedersen (2014)	HXZ	1931:08	1025
1/P	1/share price	Miller and Scholes (1982)	HXZ	1926:08	1085
Disp	Dispersion of analysts' earnings forecasts	Diether et al. (2002)	HXZ	1976:08	485
Dvol	Dollar trading volume	Brennan et al. (1998)	HXZ	1926:08	1085
Illiq	Illiquidity as absolute return-to-volume	Amihud (2002)	HXZ	1926:08	1085
Ivol	Idiosyncratic volatility	Ang et al. (2006)	HXZ	1926:09	1084
MDR	Maximum daily return	Bali et al. (2011)	HXZ	1926:09	1084

(*continued on next page*)

Table A1 (*continued*)

Anomaly (1)	Description (2)	Original study (3)	Source (4)	Start (5)	Number (6)
ME	Market equity	Banz (1981)	HXZ	1926:08	1085
S-Rev	Short-term reversal	Jegadeesh (1990)	HXZ	1926:08	1085
Svol	Systematic volatility	Ang et al. (2006)	HXZ	1986:08	365
Turn	Share turnover	Datar et al. (1998)	HXZ	1926:08	1085
Tvol	Total volatility	Ang et al. (2006)	HXZ	1926:09	1084
β-*M*	Fama and MacBeth's beta (monthly data)	Fama and MacBeth (1973)	MP	1931:08	1025
σ(*Dvol*)	Dollar volume volatility	Chordia et al. (2001)	MP	1929:08	1049
B-A	Bid-ask spread	Amihud and Mendelson (1986)	MP	1963:08	641
Short	Short interest	Dechow et al. (2001)	MP	1973:08	521
Skew	Coskewness	Harvey and Siddique (2000)	MP	1931:08	1025
Vol-T	Volume trend	Haugen and Baker (1996)	MP	1931:08	1025
		Panel G: Value			
A/ME	Market leverage	Bhandari (1988)	HXZ	1963:08	641
B/M	Book-to-market equity	Rosenberg et al. (1985)	HXZ	1963:08	641
CF/P	Cash flow-to-price	Lakonishok et al. (1994)	HXZ	1963:08	641
D/P	Dividend yield	Litzenberger and Ramaswamy (1979)	HXZ	1927:08	1073
Dur	Equity duration	Dechow et al. (2004)	HXZ	1963:08	641
E/P	Earnings-to-price	Basu (1983)	HXZ	1963:08	641
EF/P	Analysts' earnings forecasts-to-price	Elgers et al. (2001)	HXZ	1976:08	485
LTG	Long-term growth forecasts of analysts	La Porta (1996)	HXZ	1982:08	413
NO/P	Net payout yield	Boudoukh et al. (2007)	HXZ	1974:08	509
O/P	Payout yield	Boudoukh et al. (2007)	HXZ	1974:08	509
Rev	Long-term reversal	De Bondt and Thaler (1985)	HXZ	1931:08	1025
SG	Sales growth	Lakonishok et al. (1994)	HXZ	1967:08	593
An-V	Analyst value	Frankel and Lee (1998)	MP	1976:08	485
σ(*CF*)	Cash flow variance	Haugen and Baker (1996)	MP	1978:08	461
B/P-E	Enterprise component of book-to-price	Penman et al. (2007)	MP	1984:08	389
B/P-Lev	Leverage component of book-to-price	Penman et al. (2007)	MP	1984:08	389
Enter	Enterprise multiple	Loughran and Wellman (2012)	MP	1963:08	641
Pension	Pension funding status	Franzoni and Marin (2006)	MP	1981:08	425

we lag annual Compustat data by six months and assume quarterly Compustat data are known after the report date of quarterly earnings. For the strategies based on annual Compustat data, the decile portfolios are rebalanced annually at the beginning of July. The other trading strategies are rebalanced monthly.

There are some exceptions to the variable selection and portfolio formation rules described above. The composite issuance variable, *CEI*, in Panel C of Table A1 is constructed from CRSP data, but we rebalance the corresponding anomaly portfolios annually. The short interest variable, *Short*, in Panel F is constructed from the Compustat Supplemental Short Interest File. The pension funding status variable, *Pension*, in Panel G is constructed from the Compustat Pension Annual File.

The sample period for a given anomaly portfolio is determined by data availability for the corresponding sorting variable.

References

Abarbanell, J.S., Bushee, B.J., 1998. Abnormal returns to a fundamental analysis strategy. Account. Rev. 73, 19–45.

Alwathainani, A.M., 2009. Consistency of firms' past financial performance measures and future returns. Br. Account. Rev. 41, 184–196.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5, 31–56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. J. Financ. Econ. 17, 223–249.

Ang, A., 2014. Asset Management: A Systematic Approach to Factor Investing. Oxford University Press, Oxford, United Kingdom.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. J. Financ. 61, 259–299.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.

Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. J. Appl. Econom. 18, 1–22.

Balakrishnan, K., Bartov, E., Faurel, L., 2010. Post loss/profit announcement drift. J. Account. Econ. 50, 20–41.

Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. J. Financ. Econ. 99, 427–446.

Baltas, N., Kosowski, R.L., 2017. Demystifying Time-Series Momentum Strategies: Volatility Estimators, Trading Rules, and Pairwise Correlations. Unpublished working paper. Imperial College Business School.

Banz, R.W., 1981. The relationship between return and market value of common stocks. J. Financ. Econ. 9, 3–18.

Barbee Jr., W.C., Mukherji, S., Raines, G.A., 1996. Do sales-price and debt-equity explain stock returns better than book-market and firm size? Financ. Anal. J. 52, 56–60.

Barroso, P., Maio, P., 2018. Managing the Risk of The "Betting-Against-Beta" Anomaly: Does it Pay to Bet Against Beta? Unpublished working paper. University of New South Wales.

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. J. Financ. Econ. 116, 111–120.

Barry, C.B., Brown, S.J., 1984. Differential information and the small firm effect. J. Financ. Econ. 13, 283–294.

Basu, S., 1983. The relationship between earnings' yield, market value and return for NYSE common stocks: further evidence. J. Financ. Econ. 12, 129–156.

Belo, F., Lin, X., 2012. The inventory growth spread. Rev. Financ. Stud. 25, 278–313.

Belo, F., Lin, X., Bazdresch, S., 2014. Labor hiring, investment, and stock return predictability in the cross section. J. Polit. Econ. 122, 129–177.

Belo, F., Lin, X., Vitorino, M.A., 2014. Brand capital and firm value. Rev. Econ. Dyn. 17, 150–169.

Bhandari, L.C., 1988. Debt/equity ratio and expected common stock returns: empirical evidence. J. Financ. 43, 507–528.

Black, F., Litterman, R., 1992. Global portfolio optimization. Financ. Anal. J. 48, 28–43.
Boguth, O., Carlson, M., Fisher, A., Simutin, M., 2011. Conditional risk and performance evaluation: volatility timing, overconditioning, and new estimates of momentum alphas. J. Financ. Econ. 102, 363–389.
Boudoukh, J., Michaely, R., Richardson, M., Roberts, M.R., 2007. On the importance of measuring payout yield: implications for empirical asset pricing. J. Financ. 62, 877–915.
Bradshaw, M.T., Richardson, S.A., Sloan, R.G., 2006. The relation between corporate financing activities, analysts forecasts and stock returns. J. Account. Econ. 42, 53–85.
Brandt, M.W., Kang, Q., 2004. On the relationship between the conditional mean and volatility of stock returns: a latent VAR approach. J. Financ. Econ. 72, 217–257.
Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. J. Financ. Econ. 49, 345–373.
Campbell, J.Y., 1987. Stock returns and the term structure. J. Financ. Econ. 18, 373–399.
Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. J. Financ. 63, 2899–2939.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Financ. 52, 57–82.
Chan, L.K., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. J. Financ. 51, 1681–1713.
Chan, L.K., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. J. Financ. 56, 2431–2456.
Chordia, T., Subrahmanyam, A., Anshuman, R.V., 2001. Trading activity and expected stock returns. J. Financ. Econ. 59, 3–32.
Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. J. Financ. 63, 1609–1651.
Daniel, K., Hodrick, R.J., Lu, Z., 2017. The carry trade: risks and drawdowns. Crit. Financ. Rev. 6, 211–262.
Daniel, K., Moskowitz, T.J., 2016. Momentum crashes. J. Financ. Econ. 122, 221–247.
Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. J. Financ. 61, 1605–1643.
Datar, V.T., Naik, N.Y., Radcliffe, R., 1998. Liquidity and stock returns: an alternative test. J. Financ. Mark. 1, 203–219.
De Bondt, W.F., Thaler, R., 1985. Does the stock market overreact? J. Financ. 40, 793–805.
Dechow, P.M., Hutton, A.P., Meulbroek, L., Sloan, R.G., 2001. Short-sellers, fundamental analysis, and stock returns. J. Financ. Econ. 61, 77–106.
Dechow, P.M., Sloan, R.G., Soliman, M.T., 2004. Implied equity duration: a new measure of equity risk. Rev. Account. Stud. 9, 197–228.
DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R., 2009a. A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. Manag. Sci. 55, 798–812.
DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: how inefficient is the $1/N$ portfolio strategy? Rev. Financ. Stud. 22, 1915–1953.
Dichev, I.D., 1998. Is the risk of bankruptcy a systematic risk? J. Financ. 53, 1131–1147.
Diether, K.B., Malloy, C.J., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. J. Financ. 57, 2113–2141.
Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. J. Financ. Econ. 7, 197–226.
Eisdorfer, A., Misirli, U., 2020. Distressed stocks in distressed times. Management Science, (forthcoming).
Eisfeldt, A.L., Papanikolaou, D., 2013. Organization capital and the cross–section of expected returns. J. Financ. 68, 1365–1406.
Elgers, P.T., Lo, M.H., Pfeiffer Jr., R.J., 2001. Delayed security price adjustments to financial analysts' forecasts of annual earnings. Account. Rev. 76, 613–632.
Engle, R., 2004. Risk and volatility: econometric models and financial practice. Am. Econ. Rev. 94, 405–420.
Fairfield, P.M., Whisenant, S.J., Yohn, T.L., 2003. Accrued earnings and growth: implications for future profitability and market mispricing. Account. Rev. 78, 353–371.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on bonds and stocks. J. Financ. Econ. 33, 3–53.
Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. J. Financ. 51, 55–84.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116, 1–22.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. J. Polit. Econ. 81, 607–636.
Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. J. Financ. 56, 329–352.
Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using "realized" volatility. J. Financ. Econ. 67, 473–509.
Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. Account. Rev. 59, 574–603.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2005. The market pricing of accruals quality. J. Account. Econ. 39, 295–327.
Frankel, R., Lee, C.M., 1998. Accounting valuation, market expectation, and cross-sectional stock returns. J. Account. Econ. 25, 283–319.
Franzoni, F., Marin, J.M., 2006. Pension plan funding and stock market efficiency. Journal of Finance 61, 921–956.
Frazzini, A., Pedersen, L.H., 2014. Betting against beta. J. Financ. Econ. 111, 1–25.
French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. J. Financ. Econ. 19, 3–29.
George, T.J., Hwang, C.-Y., 2004. The 52-week high and momentum investing. J. Financ. 59, 2145–2176.
Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 87, 1121–1152.
Green, J., Hand, J.R.M., Zhang, X.F., 2017. The characteristics that provide independent information about average US monthly stock returns. Rev. Financ. Stud. 30, 4389–4436.
Green, R.C., Hollifield, B., 1992. When will mean-variance efficient portfolios be well diversified? J. Financ. 47, 1785–1809.
Grobys, K., Ruotsalainen, J., Äijö, J., 2018. Risk-managed industry momentum and momentum crashes. Quant. Financ. 18, 1715–1733.
Hafzalla, N., Lundholm, R., Van Winkle, M.E., 2011. Percent accruals. Account. Rev. 86, 209–236.
Harvey, C.R., 2001. The specification of conditional expectations. J. Empir. Financ. 8, 573–637.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Financ. 55, 1263–1295.
Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. J. Financ. Econ. 41, 401–439.
Heston, S.L., Sadka, R., 2008. Seasonality in the cross-section of stock returns. J. Financ. Econ. 87, 418–445.
Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? J. Account. Econ. 38, 297–331.
Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Rev. Financ. Stud. 28, 650–705.
Inoue, A., Kilian, L., 2004. In-sample or out-of-sample tests of predictability: which one should we use? Econom. Rev. 23, 371–402.
Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. J. Financ. 58, 1651–1683.
Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. J. Financ. 45, 881–898.
Jegadeesh, N., Livnat, J., 2006. Revenue surprises and stock returns. J. Account. Econ. 41, 147–171.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Financ. 48, 65–91.
Jobson, J., Korkie, B.M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. J. Financ. 36, 889–908.
Johannes, M., Korteweg, A., Polson, N., 2014. Sequential learning, predictability, and optimal portfolio returns. J. Financ. 69, 611–644.
Kelly, B., Pruitt, S., Su, Y., 2019. Characteristics are covariances: a unified model of risk and return. J. Financ. Econ. 134, 501–524.
Kirby, C., Ostdiek, B., 2012. It's all in the timing: simple active portfolio strategies that outperform naive diversification. J. Financ. Quant. Anal. 47, 437–467.
Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. J. Financ. Econ. 135, 271–292.
La Porta, R., 1996. Expectations and the cross-section of stock returns. J. Financ. 51, 1715–1742.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. J. Financ. 49, 1541–1578.
Lettau, M., Ludvigson, S.C., 2010. Measuring and modeling variation in the risk-return trade-off. In: Ait-Sahalia, Y., Hansen, L.P. (Eds.), Handbook of Financial Econometrics. Elsevier, Amsterdam, pp. 617–690.
Lewellen, J., Nagel, S., 2006. The conditional CAPM does not explain asset-pricing anomalies. J. Financ. Econ. 82, 289–314.
Li, D., 2011. Financial constraints, R&D investment, and stock returns. Rev. Financ. Stud. 24, 2974–3007.
Litzenberger, R.H., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: theory and empirical evidence. J. Financ. Econ. 7, 163–195.
Lockwood, L., Prombutr, W., 2010. Sustainable growth and stock returns. J. Financ. Res. 33, 519–538.
Loughran, T., Wellman, J.W., 2012. New evidence on the relation between the enterprise multiple and average stock returns. J. Financ. Quant. Anal. 46, 1629–1650.

Ludvigson, S.C., Ng, S., 2007. The empirical risk-return relation: a factor analysis approach. J. Financ. Econ. 83, 171–222.
Lyandres, E., Sun, L., Zhang, L., 2008. The new issues puzzle: testing the investment-based explanation. Rev. Financ. Stud. 21, 2825–2855.
Maurer, T.A., To, T.D., Tran, N., 2018. Optimal Factor Strategy in FX Markets. Unpublished Working Paper. Washington University in Saint Louis.
McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? J. Financ. 71, 5–32.
Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Financ. Lett. 1, 21–23.
Miller, M.H., Scholes, M.S., 1982. Dividends and taxes: some empirical evidence. J. Polit. Econ. 90, 1118–1141.
Moreira, A., Muir, T., 2017. Volatility-managed portfolios. J. Financ. 72, 1611–1644.
Moreira, A., Muir, T., 2019. Should long-term investors time volatility? J. Financ. Econ. 131, 507–527.
Moskowitz, T., Ooi, Y.H., Pedersen, L.H., 2012. Time series momentum. J. Financ. Econ. 104, 228–250.
Novy-Marx, R., 2011. Operating leverage. Rev. Financ. 15, 103–134.
Novy-Marx, R., 2012. Is momentum really momentum? J. Financ. Econ. 103, 429–453.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108, 1–28.
Penman, S.H., Richardson, S.A., Tuna, I., 2007. The book-to-price effect in stock returns: accounting for leverage. J. Account. Res. 45, 427–467.
Pesaran, M.H., Timmermann, A., 2002. Market timing and return prediction under model instability. J. Empir. Financ. 9, 495–510.
Piotroski, J.D., 2000. Value investing: the use of historical financial statement information to separate winners from losers. J. Account. Res. 38, 1–41.
Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. J. Financ. 63, 921–945.
Richardson, S.A., Sloan, R.G., Soliman, M.T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. J. Account. Econ. 39, 437–485.
Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. J. Portf. Manag. 11, 9–16.
Rossi, B., 2013. Advances in forecasting under instability. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam, pp. 1203–1324.
Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Account. Rev. 71, 289–315.
Soliman, M.T., 2008. The use of DuPont analysis by market participants. Account. Rev. 83, 823–853.
Thomas, J., Zhang, F.X., 2011. Tax expense momentum. J. Account. Res. 49, 791–821.
Thomas, J.K., Zhang, H., 2002. Inventory changes and future returns. Rev. Account. Stud. 7, 163–187.
Titman, S., Wei, K.J., Xie, F., 2004. Capital investments and stock returns. J. Financ. Quant. Anal. 39, 677–700.
White, H., 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica 48, 817–838.
Whitelaw, R.F., 1994. Time variations and covariations in the expectation and volatility of stock market returns. J. Financ. 49, 515–541.
Xing, Y., 2008. Interpreting the value effect through the q-theory: an empirical investigation. Rev. Financ. Stud. 21, 1767–1795.