

The Journal of Portfolio Management

Fact, Fiction, and the Size Effect

Ron Alquist, Ronen Israel and Tobias Moskowitz

JPM published online 5 October 2018

<http://jpm.ijournals.com/content/early/2018/10/05/jpm.2018.1.082>

This information is current as of October 10, 2018.

Email Alerts Receive free email-alerts when new articles cite this article. Sign up at:
<http://jpm.ijournals.com/alerts>

Institutional Investor Journals

1120 Avenue of the Americas, 6th floor,
New York, NY 10036, Phone: +1 212-224-3589

© 2017 Institutional Investor LLC. All Rights Reserved

IIJ Institutional
Investor
Journals

Downloaded from <http://jpm.ijournals.com/> by guest on October 10, 2018

Fact, Fiction, and the Size Effect

RON ALQUIST, RONEN ISRAEL, AND TOBIAS MOSKOWITZ

RON ALQUIST

is a vice president at AQR Capital Management in Greenwich, CT.
ron.alquist@aqr.com

RONEN ISRAEL

is a principal at AQR Capital Management in Greenwich, CT.
ronen.israel@aqr.com

TOBIAS MOSKOWITZ

is the Dean Takahashi professor of finance and economics at Yale University, a research associate at the NBER, and a principal at AQR Capital Management in Greenwich, CT.
tobias.moskowitz@aqr.com

After confronting the myths surrounding momentum investing¹ and value investing,² we realized two things: 1) We had passed over the first anomaly discovered in academic finance and the one that had been around the longest—size, and 2) despite its longevity and the attention it has received, there is still much confusion and debate surrounding the size anomaly.

The size effect is the phenomenon in which small stocks (i.e., those with lower market capitalizations), on average, outperform large stocks (i.e., those with higher market caps) over time. The size premium is the return achieved by buying (being long in an absolute sense or overweight relative to a benchmark) small stocks and selling (shorting or underweighting) large ones. The size effect was first documented by several academic papers in the early 1980s,³ and it quickly became the first real challenge to the field's preeminent asset pricing framework, the capital asset pricing model (CAPM).

Broadly speaking, researchers responded to its discovery in two ways. On one hand, proponents of market efficiency argued that this evidence simply indicated that the CAPM was misspecified and that size was related to

a second source of priced risk beyond the market. According to this view, both a stock's market beta and its size were now required to understand its expected returns. As long as size was correlated with a fundamental source of risk, rational investors needed to be compensated for holding assets more exposed to that risk. Other scholars interpreted the evidence of a size premium as a more fundamental conceptual challenge to market efficiency, in which small stocks relative to large stocks were simply mispriced, having nothing to do with compensation for risk. The size premium, therefore, represented the first true market anomaly.

Yet, despite size's legacy and its subsequent prominence in the field, there remains much debate about the size effect, including its reliability. The very existence of a size premium, for example, turns out to be a less well-established empirical fact than its younger cousins, value and momentum (and defensive and quality premiums as well)—something we will investigate in depth in this article.

This article is organized around a number of facts and fictions about the size effect that warrant clarification. The facts we present include the following: that the size effect diminished shortly after its discovery and publication; that it is dominated by a January seasonal effect; that it is not applicable or does not work for other asset

¹ Asness et al. 2014.

² Asness et al. 2015a.

³ See Banz 1981, Keim 1983, and Roll 1983.

EXHIBIT 1

Original Size Anomaly Sample Period

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| SMB | 1.9% | 9.8% | 1.21 | 0.19 | −0.3% | −0.22 | −0.03 | 1/31/1936 | 12/31/1975 |
| Decile 1–10 | 7.1% | 25.3% | 1.78 | 0.28 | 2.5% | 0.66 | 0.11 | 1/31/1936 | 12/31/1975 |

classes outside of individual equities; that it can be made much stronger when looked at in conjunction with other factors (namely, defensive/quality factors); that the size premium mostly comes from microcap stocks and is difficult to implement in practice; and, finally, that the size effect continues to receive a disproportionate amount of attention relative to other factors with similar or stronger evidence behind them. The fictions we attempt to clarify include that the size effect is a strong anomaly; that other factors performing better among small stocks is evidence of a size effect; that the size effect is robust to the chosen method of measurement; that it works in other markets and settings; and that it seems to be more than just an illiquidity premium.

Finally, we will address fictional theories that propose an economic story, unrelated to liquidity, in which small stocks should deserve a marginal premium over their other risk characteristics and in which a size premium is consistent with a risk-based efficient markets view of the world. Although a size premium can certainly occur in a world of efficient or inefficient markets, we find economic stories, other than as a proxy for illiquidity, regarding why the size of a firm should matter for pricing to be puzzling.

As done in our prior papers, we address the facts and fictions of the size effect using published and peer-reviewed academic papers and conduct tests using the most well-known and straightforward publicly available data.⁴

⁴ Kenneth French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) provides returns for market (RMRF), small (SMB), value (HML), momentum (UMD), and profitability (RMW) factors, including returns for the long and short sides separately, and for both large- and small-capitalization securities separately, all of which we use in this article. AQR's data library (<https://www.aqr.com/library/data-sets>) provides returns for a betting against beta (BAB) factor from Frazzini and Pedersen (2014), which we use in this article.

Based on the facts we uncover, size does not appear to be on equal footing with other prominent factors, such as value, momentum, and defensive/quality investing. The returns to size are far less stable, less persistent, and less robust than these other factors. Although we do not completely deny the existence of a size effect, and we certainly do not advocate actively betting against or shorting it, we also do not believe that size on its own is a key factor for constructing portfolios. We believe the size effect captures part of a broader effect—an illiquidity premium—that can add value at the margin in conjunction with other factors, but in which it is also (by definition) more difficult and expensive to trade. On its own, a size factor is not a particularly strong source of expected returns in practice, despite its prominence in the literature and the attention it has received from the investment world.

FICTION: THE SIZE EFFECT IS ONE OF THE STRONGEST DOCUMENTED ANOMALIES/FACTORS

Given its pedigree, you would be forgiven for thinking that the size factor is one of the strongest and most robust anomalies in finance. In fact, it is one of the weakest. It is significantly weaker than other well-known anomalies such as value, momentum, profitability, and defensive/quality or low volatility.

Size has never been a very strong effect. Let's start by examining the original study on the size effect. Banz (1981) documented that small stocks outperformed large stocks over his sample period, which spanned January 1936 to December 1975. Exhibit 1 attempts to replicate his results using the same sample period. It reports the annualized mean, volatility, *t*-statistic of the mean, Sharpe ratio, annual alpha, *t*-statistic of the alpha, and information ratio (alpha divided by residual standard deviation) from a regression of the size factor's returns on

the market portfolio (CAPM). All statistics are computed using monthly returns, but the numbers reported are annualized. We use two specifications of long–short portfolios that seek to capture the size premium: The first is Fama and French’s small minus big (SMB) factor, which is constructed from the six size/book-to-market benchmark portfolios and is long the smallest half of stocks (based on NYSE breakpoints) and short the largest half. By construction, then, it is composed of half growth and half value stocks. The second portfolio is long the smallest decile of stocks (based on NYSE breakpoints) and short the largest. Banz’s (1981) original study used Fama and MacBeth (1973) regressions to show a size premium, which is probably closer to the decile portfolio returns approach.⁵

As Exhibit 1 shows, over the 1936 to 1975 period, the evidence in favor of a strong size premium is weak. The first four columns of the table report the annual return, volatility (standard deviation), *t*-statistic of the mean, and Sharpe ratio of the two size strategies. SMB has a 1.9% annualized mean return with almost 10% annual volatility, translating into a 0.19 Sharpe ratio. The mean return of SMB over the 1936 to 1975 period, however, is not statistically significant, with a *t*-statistic of only 1.21. The 1–10 decile portfolio has a much higher mean return of 7.1%, but with more than twice the volatility at 25.3% per year for a Sharpe ratio of 0.28. Here, the *t*-statistic of 1.78 barely meets the 10% significance threshold but does not meet the commonly used 5% threshold. In fact, if Banz’s paper had been written today, and using Harvey, Liu, and Zhu’s (2016) and Harvey’s (2017) suggested threshold value of 3.0 for the data-mining robust *t*-statistic, the statistical evidence for a size effect would be even weaker. These results indicate that the size effect is not particularly strong, even over the original sample period in which it was discovered.

The next three columns of Exhibit 1 report the alphas of the size strategies versus the market portfolio (CAPM alpha). Academic work generally evaluates strategies after controlling for factor exposures, with market beta being the most obvious one, but all too often we still see practitioners presenting results based on raw returns even when the factor exposure is large and intuitive.

If a strategy works only because it has a bigger market beta, then a more efficient and reliable way to capture those returns is simply to allocate more to the market factor itself. This point applies with particular force to the size premium. Because small stocks typically have larger market betas than large stocks, part, or even all, of the size premium may simply be the equity market risk premium in disguise. The CAPM alphas account for the beta difference between small and large stocks. As Exhibit 1 shows, SMB has zero (in fact, slightly negative) alpha with respect to the market when controlling for the betas, and the decile spread portfolio has a positive alpha (2.5%) that is statistically indistinguishable from zero (*t*-statistic of 0.66). These results suggest that the size premium in its original sample is not only weak but seems to be captured by general market exposure.

The poor showing of the size effect in its original sample raises the question of how it received so much initial attention and was considered a challenge to the CAPM when it appears that the CAPM captures it well. The reason is clear if we compare our findings with those in the original studies (e.g., Banz (1981)). In those studies, the evidence of a size effect was much stronger than what we report here. What is causing the difference?

One issue that may have weakened the size effect since the original studies were conducted is that errors in our historical databases of stock prices have been discovered and fixed. The most commonly used database for stock returns is that of the Center for Research in Security Prices (CRSP) at the University of Chicago, which continually fixes data errors it encounters going back in time. One such data error that plagued early studies was a delisting bias. Stocks delisted from the exchanges simply had no return information available for them and were therefore dropped from the analysis. Shumway (1997) painstakingly backfilled these delisting returns by hand collecting delisting events and recording the delisted prices, which, on average, were for negative events.⁶ Because these negative delisting returns were omitted from the original data sources of the original studies on size, and because delisting events are more likely to occur for smaller firms, this

⁵ As Fama (1976) showed, the Fama and MacBeth (1973) regressions tend to place more weight on the smallest, most volatile stocks. Hence, a decile sort, which emphasizes the most extreme stocks, will match these results better.

⁶ Shumway (1997) and Shumway and Warther (1999) found that the delisting return is –55%, on average, for NASDAQ firms and –30%, on average, for NYSE/Amex firms when the delisting is for performance-related reasons.

EXHIBIT 2

Size Anomaly: Full Sample Period

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| SMB | 2.5% | 11.1% | 2.13 | 0.22 | 0.9% | 0.86 | 0.09 | 7/31/1926 | 12/31/2017 |
| Decile 1–10 | 6.1% | 25.6% | 2.29 | 0.24 | 2.2% | 0.89 | 0.09 | 7/31/1926 | 12/31/2017 |

EXHIBIT 3

Major Anomalies: Full Sample Period

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-----|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| BAB | 8.1% | 11.0% | 6.85 | 0.73 | 8.7% | 7.44 | 0.80 | 12/31/1930 | 12/31/2017 |
| HML | 4.6% | 12.1% | 3.65 | 0.38 | 3.4% | 2.73 | 0.29 | 7/31/1926 | 12/31/2017 |
| MOM | 7.9% | 16.3% | 4.62 | 0.48 | 10.2% | 6.32 | 0.67 | 1/31/1927 | 12/31/2017 |
| RMW | 3.0% | 7.7% | 2.89 | 0.39 | 3.7% | 3.68 | 0.50 | 7/31/1963 | 12/31/2017 |
| SMB | 2.5% | 11.1% | 2.13 | 0.22 | 0.9% | 0.86 | 0.09 | 7/31/1926 | 12/31/2017 |

bias made the performance of small stocks look better than it actually was relative to large stocks. Hence, part of the size premium originally discovered by researchers in the late 1970s and early 1980s may have been driven by these data errors that have since been fixed. Thus, a researcher in 1980 might find no return information for a delisted stock in, say, January 1965, but a researcher in 2018 looking at that same stock in January 1965 would find (on average) a –30% return. Therefore, even if one uses the original sample periods of the early studies, the return series during those periods contain fewer errors today than they did at the time researchers were initially investigating them. Consequently, replication of the size anomaly appears weaker than in the original studies, even when the exact same sample period is being examined (see also Asness et al. (2018)).

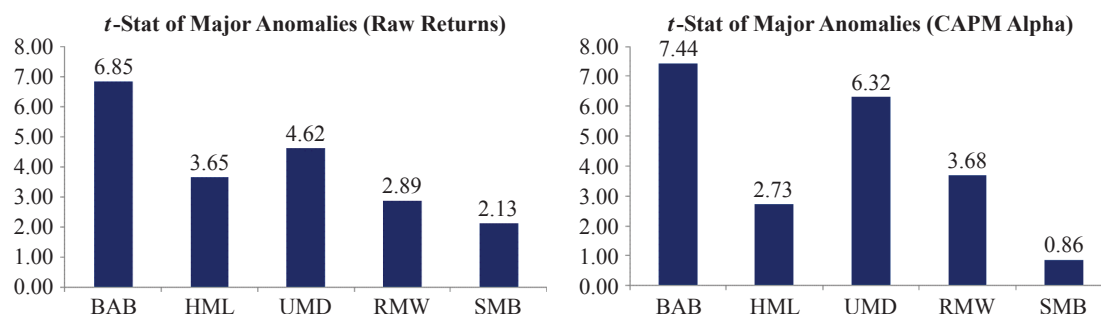
We can also look at the size effect over the much longer period for which we have data, including going back to 1926 and, of course, going forward until 2017. For the full sample period over the last 91 years, the size premium looks to be about the same in magnitude but is statistically a bit stronger because of the larger sample size—as Exhibit 2 shows, SMB has 2.5% annual return with a *t*-statistic of 2.13, and the decile spread portfolio has a 6.1% return with a *t*-statistic of 2.29. However, as the next three columns show, the CAPM still prices

these portfolios nicely: Alphas of both SMB and the Decile 1–10 portfolios are statistically insignificant at conventional levels.

So, the size premium on its own is significant, but adjusting for market beta renders it insignificant. How do these results compare to other well-known factors from the literature? Exhibit 3 reports the performance of five popular academic factors based on the five most prominent asset pricing anomalies found in the academic literature, over the longest sample of data available. The factors include the *betting against beta* factor (BAB) of Frazzini and Pedersen (2014) that is long low-beta stocks and short high-beta stocks, levered to have the same beta and taken from AQR's data library; the *high minus low* value factor (HML) of Fama and French (1993), which is a portfolio long the top 30% of stocks based on high ratios of book-to-market equity (BE/ME) and short the lowest 30% of BE/ME stocks, taken from Ken French's website; the *up minus down* momentum factor (UMD), which is long high-momentum stocks (the top 30%) and short low ones (the bottom 30%), taken from Ken French's website; the *robust minus weak* profitability factor (RMW), which is long highly profitable firms based on the top 30% in terms of profits-to-assets ratio and short the bottom 30%, following Fama and French (2015) and

EXHIBIT 4

t-Statistic of Major Anomalies (raw return vs. CAPM alpha)



taken from Ken French's website; and, finally, the *small minus big* size factor, (SMB).

As Exhibit 3 shows, no matter what metric of performance is used—mean, Sharpe ratio, *t*-statistic, alpha, or IR—the size factor, SMB, has the *worst* performance among the five factors, often by a wide margin.⁷ For example, SMB's full-sample Sharpe ratio is 0.22, whereas the next lowest Sharpe ratio is that of HML at 0.38. UMD and BAB have Sharpe ratios almost two to three times larger (0.48 and 0.73, respectively). The CAPM alphas are significant for all of the factors except SMB, indicating that factors other than size add a return premium above and beyond traditional equity market risk. This evidence also shows that the size effect is in fact not a market anomaly, unlike the other factors. Furthermore, as we will show later, the other four factors have a wealth of out-of-sample evidence showing their efficacy in other time periods, other equity markets, and even other asset classes. The size factor fails to yield significantly positive effects out of sample in all of these settings.

Exhibit 4 shows clearly that on either a raw or risk-adjusted return perspective, the size effect is the weakest of the anomalies.

Thus, although we can debate whether there is in fact a significant size premium at all (and whether there ever was), there is little debate about whether size is one of the strongest anomalies—it is not. It is one of the weakest.

FACT: THE SIZE EFFECT HAS DISAPPEARED OR WEAKENED SINCE ITS DISCOVERY

As noted earlier, there is some debate as to whether a size effect ever existed at all. Even among those who believe there was a healthy size premium, however, many more believe it has significantly weakened over time since its discovery, to the extent that it no longer exists.⁸

Exhibit 5 plots the Sharpe ratio of the SMB over the original sample when it was discovered (1936–1975), as well as decade by decade over the four decades following the original discovery of size: 1976 to 1986, 1987 to 1996, 1997 to 2006, and 2007 to 2017.

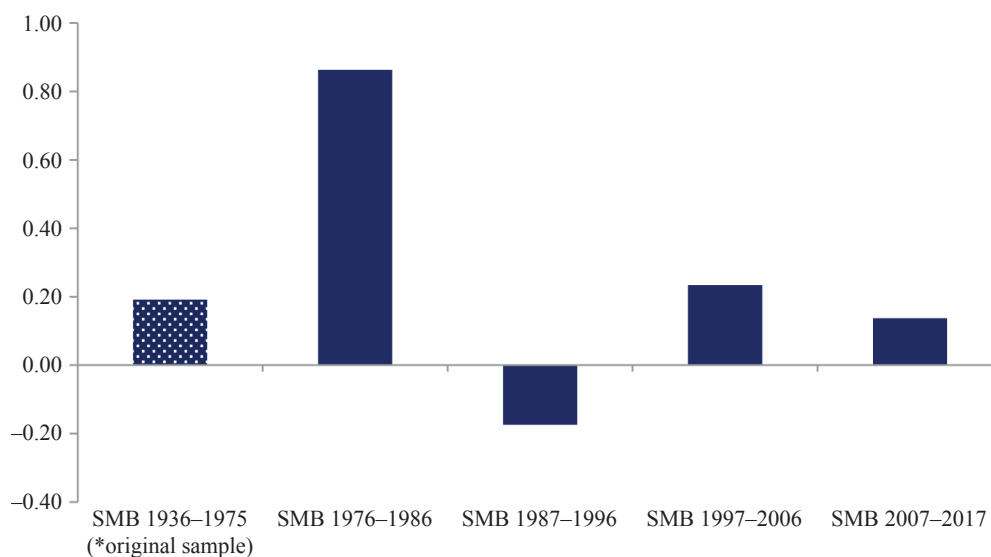
One of the most interesting findings is the very strong performance of SMB in the decade after the end of the original sample, during the first half of which Banz's findings were circulated among academic financial economists and then published in 1981. From 1976 to 1986, the size effect experienced returns almost four times larger than in its original sample. Following Banz's (1981) seminal study, and perhaps spurred on by the immediate increase in the strength of the size effect following that study, a suite of papers by Reinganum (1983a, 1983b), Keim (1983), Schwert (1983), and Roll (1983) all dissected the size anomaly during a decade in which its returns were particularly strong and being noticed by practitioners. However, following the publication of these papers in the 1980s, the returns to size fell precipitously and were actually negative over

⁷ These conclusions also hold if we compare the performance of the factors over the shortest period for which they are available, the common sample period from July 1963 to December 2017.

⁸ See Dichev (1998); Chan, Karceski, and Lakonishok (2000); Horowitz, Loughran, and Savin (2000); Gompers and Metrick (2001); Israel and Moskowitz (2013); Van Dijk (2013); Chordia, Subrahmanyam, and Tong (2014); and McLean and Pontiff (2016).

EXHIBIT 5

Sharpe Ratio over Time



the subsequent decade, turning slightly positive over the next two decades, but essentially remaining flat. Since the slew of publications on the size effect, there has been no significant positive premium associated with small-cap strategies.

Scholars have offered various explanations for the disappearance of the size effect. For instance, Schwert (2003) suggested that the small-firm anomaly disappeared shortly after the initial publication of the papers that discovered it because of an explosion of small-cap funds and indexes that may have priced it away. Gompers and Metrick [2001] argued that institutional investors' continued demand for large stocks in the 1980s and 1990s increased the prices of large companies relative to small companies, which may account for a large part of the size premium's disappearance over this period. Finally, Hou and Van Dijk (2014) argued that small firms experienced a series of negative profitability shocks in the 1980s and 1990s and that these shocks help to explain the disappearance of the size premium during that period.

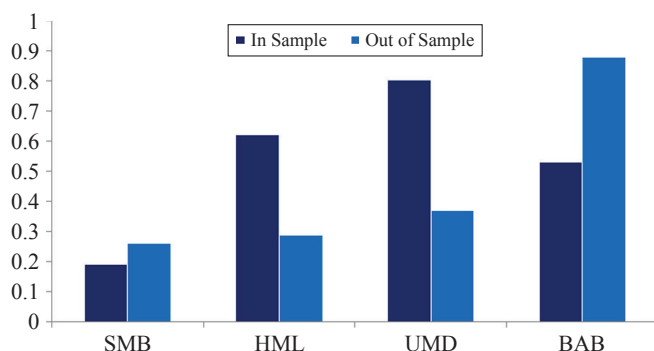
There is also the specter of data mining having exaggerated the original results and thus explaining why the out-of-sample evidence looks poorer. McLean and Pontiff [2016] argued that many anomalies suffer from poorer out-of-sample performance because of both data-mining and arbitrage activity that lowers their

returns.⁹ We compare how other prominent anomalies—value, momentum, and defensive/quality—fare in the out-of-sample periods since their discovery and compare them with the performance of size. We use HML, UMD, and BAB factors to represent each of the other anomalies and define their original sample periods following McLean and Pontiff (2016), who used the seminal papers of Fama and French (1992) for value, Jegadeesh and Titman (1993) for momentum, and Fama and MacBeth (1973) for market beta. The corresponding original sample periods in those studies are 1963 to 1990, 1964 to 1989, and 1926 to 1968, respectively. We therefore report the out-of-sample performance of HML, UMD, and BAB from 1991 to 2017, 1990 to 2017, and 1969 to 2017, respectively.

⁹Interestingly, McLean and Pontiff (2016) showed a bias in published papers, in which the last few years of a paper's data sample tend to exhibit returns that are much stronger than the first few years of out-of-sample data following the original sample. This bias could result from sample-specific data mining or more indirectly from selective updating of data, in which the authors update samples when the added few years improve their results but do not bother if the results are unchanged or weaker. It could also be the case that papers are written *because* the recent sample is so strong, and strong recent performance may make the effect more hotly debated, salient, and interesting. This may also partly explain the proliferation of size-related papers in the years immediately following the original study.

EXHIBIT 6

Sharpe Ratio of Factors: In-Sample versus Out-of-Sample Performance



As Exhibit 6 shows, the out-of-sample performance of the factors is mixed compared with their in-sample performance. SMB's Sharpe ratio increases slightly in the out-of-sample period, but it remains the smallest of any of the other out-of-sample Sharpe ratios. As we saw earlier, there has been significant time variation in the performance of SMB on a decade-by-decade basis, and its strong performance in the 1976–1986 period (Sharpe ratio = 0.86) led to a small increase in the full out-of-sample Sharpe ratio. The Sharpe ratios of the HML and UMD factors decline in the out-of-sample period, whereas the BAB factor's Sharpe ratio improves.¹⁰ More importantly, however, the size premium remains statistically insignificantly different from zero in the out-of-sample period, whereas the other three factors continue to exhibit statistically significant, though in some cases lower, return premiums in their respective out-of-sample periods. Therefore, although the HML and UMD premiums fall out of sample, their returns remain significantly positive.¹¹

¹⁰ The Sharpe ratios for HML and UMD understate the benefits of implementing them in a single portfolio because value and momentum strategies are negatively correlated with each other (Asness, Moskowitz, and Pedersen 2013). A combination of the two factors—something actually done in practice—can lead to better performance than implementing each one individually. Furthermore, using multiple measures of individual factors (e.g., value) can also greatly improve the performance of a portfolio compared with using a single measure (Israel, Jiang, and Ross (2017).

¹¹ In another out-of-sample test, Ilmanen et al. (2018) examined a century of evidence on these factor premiums across many asset classes and tested their out-of-sample performance both before and after the original sample periods in which they were

Regardless of the reason—data mining, arbitrage, and shifting demand for small stocks may all have partly contributed to its demise—the size premium has weakened over time and has become absent since its original discovery in a congregation of papers published in the early 1980s. This evidence stands in stark contrast to the evidence we have for other factors, such as HML, UMD, and BAB, which shows that the returns from these strategies remain economically significant and are robust across sample periods and even asset classes.¹²

FICTION: THE SIZE EFFECT IS ROBUST TO HOW YOU MEASURE SIZE

A single measure of anything seems unlikely to be optimal, given estimation error and data mining concerns and absent any strong theory. Indeed, we showed that for both value and momentum,¹³ multiple measures of each tend to provide better and more stable performance, providing robustness driven by diversification benefits from different measures that diminish data errors, noise, and the influence of missing data that can otherwise limit samples.

For size, we also put this statement to the test. The predominant way to measure size (in academia and practice) is to use the firm's market capitalization, which is the share price of the equity in the firm multiplied by the number of outstanding shares of the stock. However, the size of the firm could be captured in many ways. How robust is the size effect to different measures of size?

Academia has considered this question. Berk (1995a), for instance, using an argument from Ball (1978), argued that when size is measured by market capitalization, which contains market prices, it can mechanically lead to a negative relationship between size and average returns. The idea is simple: Returns equal today's price

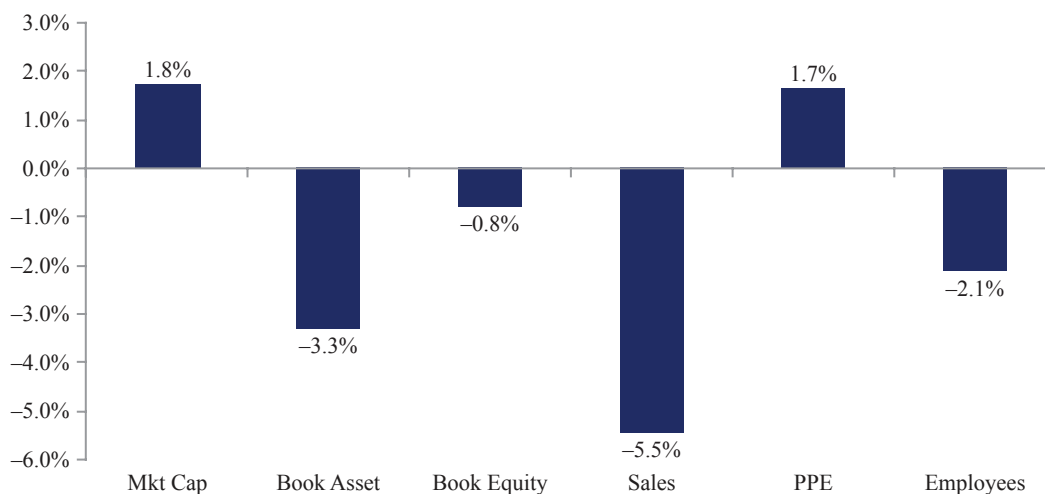
discovered. They found that the out-of-sample evidence for value and momentum strategies in U.S. equities in both the before and after periods is similar, although worse than in the original sample period. The before-discovery sample should be immune from arbitrage trading effects because the anomalies were not yet known, so comparing the performance of the strategies in the pre- and postdiscovery samples provides a test of data mining versus arbitrage-driven return degradation. Thus, the available out-of-sample evidence does not support the claim that the returns to value and momentum strategies are being arbitrated away.

¹² See Ilmanen et al. (2018).

¹³ See Asness et al. (2014, 2015a).

EXHIBIT 7

CAPM Alpha of Various Size Measures (full sample)



plus dividends, divided by yesterday's price, which will have a statistical negative relationship with market cap (which equals shares outstanding times yesterday's price) by construction if prices move. Thus, if running the following regression,

$$\frac{P_t + D_t}{P_{t-1}} = \alpha + \beta(\text{market cap}_{t-1}) + \delta'X_{t-1} + \epsilon_t$$

Return_t market cap_{t-1}

where P_t , D_t , and S_t are price, dividends, and shares outstanding, respectively, at time t , and X_{t-1} is a set of control variables, if the controls do not completely account for all price movements, mechanically there will be a negative relationship between returns and market cap because the price at time $t-1$, highlighted in red, shows up on both sides of the regression.

To address this potential bias, Berk (1995b, 1997) suggested using non-price-based measures of size as a better way to test the true relationship between size and average returns. He found, however, that using non-price-based size measures (e.g., book equity or number of employees) results in no reliable size premium. Hence, the size effect does not appear to be robust to these other measures of size that do not contain market prices.

We examine the robustness of different measures of size for predicting returns by using non-price-based size measures. Specifically, we use the book value of assets; book value of equity; sales; property, plant, and

equipment (PP&E); and the number of employees as alternative non-price-based measures of the size of a firm. For each non-price-based size measure, we form a 1–10 decile portfolio in the same manner as previously and use each non-price size measure to rank and sort stocks.

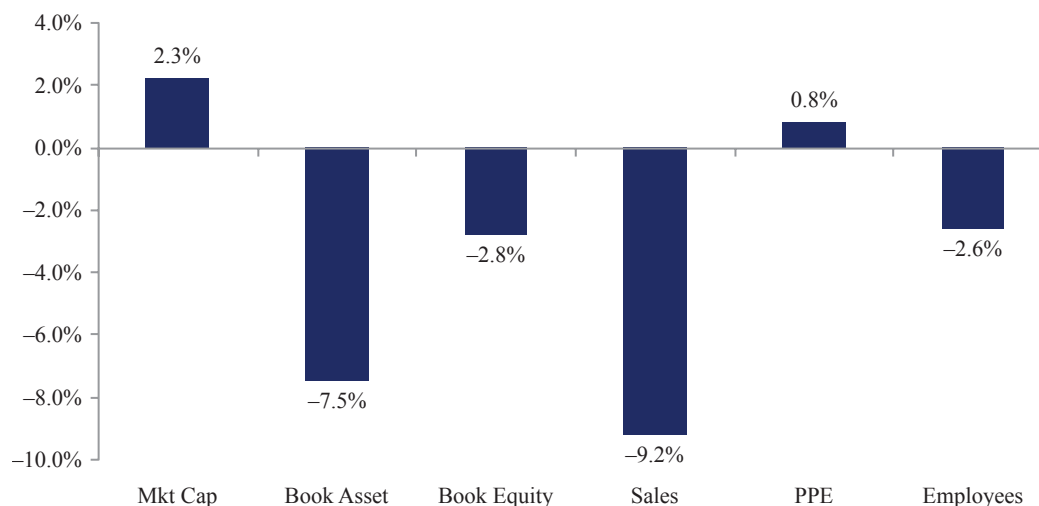
Exhibit 7 shows the alphas with respect to the market (CAPM) of the portfolios based on different measures of size over the full sample or the longest period for which we have data available (January 1951 to December 2017, where accounting numbers are available). The first bar shows the results for market cap as the measure of size, and the remaining bars show the results for the non-price size measures. As the exhibit shows, the market-cap measure of size (which uses prices) delivers the strongest size premium, whereas the non-price-based measures of size are weaker, with four out of the five measures producing a negative result.

Exhibit 8 reproduces Exhibit 7 for the out-of-sample period from 1976 to 2017 after the original study by Banz (1981). Here, the performance of the non-price size measures is even worse, and the only substantial return premium exhibited is that for the market-cap measure of size.

These results are broadly consistent with those of Berk (1995a, 1995b, and 1997) and suggest a much weaker relationship between non-price-based size measures and average returns. What little size premium might be present when using market cap to measure size disappears entirely (and switches sign) when using

EXHIBIT 8

CAPM Alpha of Various Size Measures (1976–2017)



non-price-based measures of size. These results are even starker out of sample following the original research. Hence, the size effect seems to vary considerably with different measures of size and does not appear very robust.

This result is counterintuitive because any individual measure has error (resulting from mismeasurement, missing data for some firms, and random errors), so an average of similar measures should help reduce noise and be more robust. Asness et al. (2015a) and Israel and Moskowitz (2013) showed that multiple measures of value produce more stable value portfolios that deliver higher Sharpe ratios, higher information ratios, and more robust returns. The same is true for momentum (Asness et al. 2014) and for defensive/quality (Asness, Frazzini, and Pedersen 2017). As with any systematic process, unless theory dictates a preference for one metric over all others, an average of sensible measures is generally the best and most robust approach. Although this is true for all of the other commonly used factors, it does not appear true for size.

In addition, using multiple measures to reduce errors generally improves the out-of-sample performance of a strategy. As with any specific sample of data, one will always find some measures that work particularly well in sample and some that do not. However, without theory explaining why one measure should outperform another, this is usually the result of chance. Using multiple measures can therefore guard against the

dangers of data mining—picking one particular measure that happened to work well in one particular sample and that is often overfitted to that sample.

Because only the market-cap measure of size seemed to deliver any sort of premium and all other measures produced a negligible or opposite-sign premium, the robustness of the size effect is questionable. Moreover, the significantly worse performance of the market cap-based measure of size in the out-of-sample period following the original studies is troubling from a data-mining perspective. Combining these results, the size effect does not appear robust. Unlike other factors (e.g., value, momentum, and defensive/quality), the size premium is quite sensitive to changes in how it is measured and over what sample it is examined.

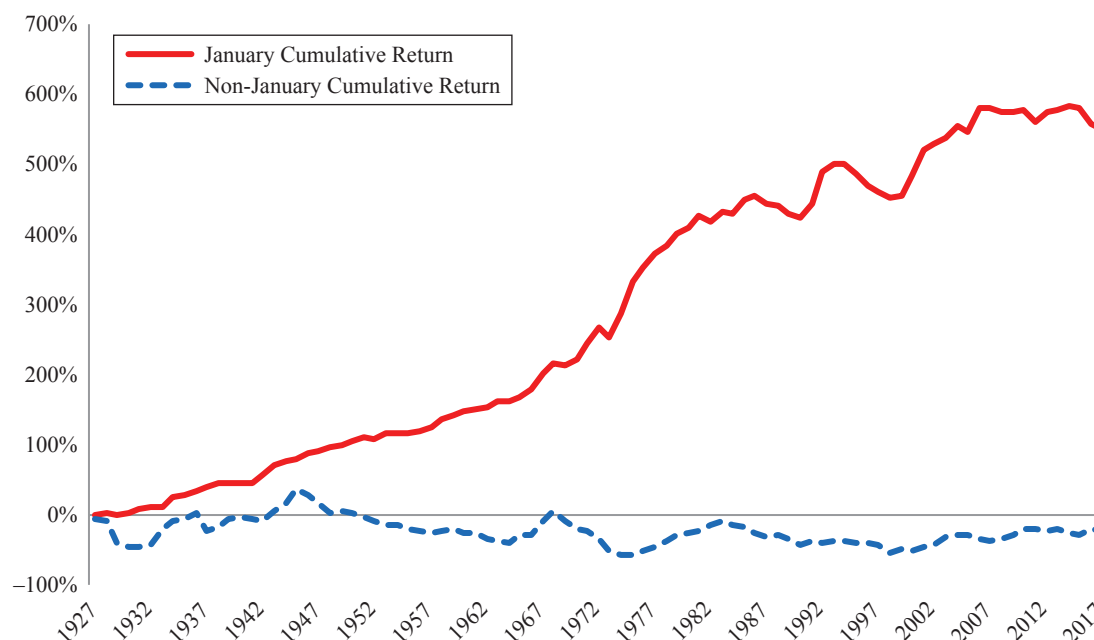
FACT: THE SIZE EFFECT IS DOMINATED BY A JANUARY EFFECT

One of the earliest findings about the size effect was that it mostly occurred in January (see Keim 1983, Roll 1983, and Reinganum 1983a, as well as recent work by Asness et al. 2018). This strong seasonal component to the size effect has long been a focal point—for both advocates and critics—of the size factor.

Again, let's start with the full sample evidence from 1926 to 2017. Exhibit 9 plots the cumulative returns to the size factor, SMB, for the month of January only versus all other months. For the January

EXHIBIT 9

January versus Non-January Cumulative Return



cumulative returns, we invest in SMB in January of each year, then put the returns in cash for the remaining months (February to December). For the non-January cumulative returns, we invest in SMB for all months except January (putting the money in cash for January). A plot of the time series of the two cumulative returns is reported in the exhibit.

As Exhibit 9 clearly shows, there is a substantial return to the size factor in January but absolutely no evidence of any size premium outside of January. The returns to size are completely flat throughout most of the year. Whatever premium the size factor has seems to be generated almost exclusively in January.¹⁴

A more formal test of the size effect in and outside of January is contained in Exhibit 10, in which we report the average monthly return, volatility, t -statistic, Sharpe ratio, and CAPM alphas to SMB in January and non-January months. SMB delivers an impressive 2.1%

¹⁴ Moreover, early researchers (Keim 1983 and Roll 1983) showed that it was in fact the first few trading days of the year that generated nearly all of the January premium and hence all of the size premium as well. This empirical fact has been attributed to year-end tax-loss selling, rebalancing, and cash infusion at the beginning of the year from investors, as well as window dressing by mutual fund managers at the turn of the year.

return in the month of January alone, with a t -statistic of 6.25 that is not captured at all by the CAPM (alpha equals 1.9% with a t -statistic of 5.64). These results are dramatically stronger than what we obtained for SMB over all months over the same sample period. The non-January months exhibit literally zero size premium (average return of 0.0% from 1926 to 2017) and an alpha of -0.1% . All of the returns to size are concentrated in January exclusively, with no evidence of any size effect—economically or statistically—outside of January.

Because of these results, the size effect and the January effect have been inextricably linked. Since its discovery, many researchers have argued that the January effect has weakened over time, driven possibly by increased arbitrage trading that exploited it, less price impact in the market from turn-of-the-year trades as a result of improved market liquidity, more passive index investing, and so on. The weaker January effect may in turn have contributed to the weaker size effect over time.

Exhibit 11 reports the same statistics on SMB in January and non-January months for the more recent sample from 1976 to 2017, following the original size studies. As the exhibit shows, the January effect is indeed much weaker in the more recent sample, but

EXHIBIT 10

SMB Performance in January and Non-January Months (full sample)

| | Monthly Return | St. Dev. | Raw <i>t</i> -Stat | Sharpe Ratio | Monthly Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-------------|-------------------|----------|-----------------------|-----------------|------------------------------|-------------------------------------|----------------|------------|------------|
| January | 2.1% | 3.2% | 6.25 | 0.65 | 1.9% | 5.64 | 0.61 | 1/31/1927 | 1/31/2017 |
| Non-January | 0.0% | 3.1% | 0.32 | 0.04 | -0.1% | -0.88 | -0.10 | 7/31/1926 | 12/31/2017 |

EXHIBIT 11

SMB Performance in January and Non-January Months (1976–2017)

| | Monthly Return | St. Dev. | Raw <i>t</i> -Stat | Sharpe Ratio | Monthly Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-------------|-------------------|----------|-----------------------|-----------------|------------------------------|-------------------------------------|----------------|------------|------------|
| January | 1.0% | 2.7% | 2.39 | 0.37 | 1.0% | 2.24 | 0.35 | 1/31/1976 | 1/31/2017 |
| Non-January | 0.2% | 3.0% | 1.09 | 0.18 | 0.0% | 0.27 | 0.04 | 2/29/1976 | 12/31/2017 |

it still dominates what is left of the size effect in this sample. SMB in January averages only 1.0% per month in this period compared to the 2.1% return it exhibited in January over the longer sample dating back to 1926, with a *t*-statistic of 2.39. The CAPM once again cannot explain these returns. Outside of January, there is no SMB premium in the recent period—the CAPM alpha is 0.0% with a *t*-statistic of 0.27.

The bottom line is that other than in January, there is not and never was a size premium. All of the returns to size seem to come from January alone, and the fact that the January effect has diminished over time has contributed to the demise of the size effect.

FICTION: THE SIZE EFFECT WORKS IN OTHER EQUITY MARKETS

Another way to assess the robustness of any factor is to examine its efficacy in other markets. Other equity markets provide a set of out-of-sample tests for any factor and help to guard against data mining. They also can help build a better diversified global factor that offers a more stable return premium because diversification benefits often exist across international equity markets. Much research has shown that factors such as value, momentum, and defensive/quality work extremely well in other markets (Fama and French 1998, 2012, 2017; Rouwenhorst 1998; Liew and Vassalou 2000; Griffin, Ji, and Martin 2003; Chui, Titman, and Wei 2010;

Asness, Moskowitz, and Pedersen 2013; Asness, Frazzini, and Pedersen 2017; and Frazzini and Pedersen 2014). How well does size fare in other equity markets?

We examine 24 international equity markets and compute an SMB portfolio in each market following the same procedure used earlier, which matches that of Fama and French [1993]. The universe of stocks in each country is the MSCI universe. The data are from World Scope and cover the period of January 1984 to December 2017. Exhibit 12 reports the average SMB returns across countries grouped into regions: Europe (Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, the United Kingdom, Greece, Ireland, Israel, Italy, the Netherlands, Norway, Portugal, and Sweden), North America (Canada and the United States), the Pacific (Australia, Hong Kong, Japan, New Zealand, and Singapore), and global. The regional portfolios are equal-weighted averages of the country-specific SMB portfolios in each region. We also consider the global SMB portfolio excluding the United States, available from Ken French's website ("Global SMB ex U.S.").

Exhibit 12 reports the *t*-statistic of the CAPM alphas of these regional SMB portfolios, with the conventional threshold for statistical significance of 2.0 highlighted. As the exhibit shows, none of the *t*-statistics for the regional portfolios is close to being reliably positive, and most are, in fact, negative. Thus, we see no consistent evidence of a positive size premium in these other markets.

EXHIBIT 12

CAPM Alpha t -Statistic of International Size Effect

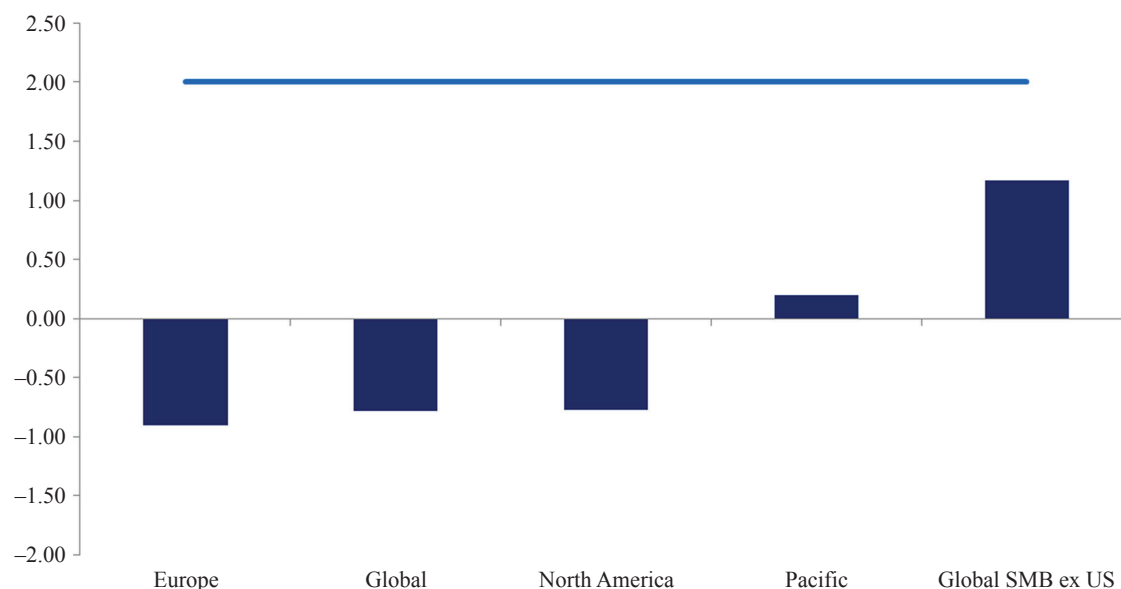
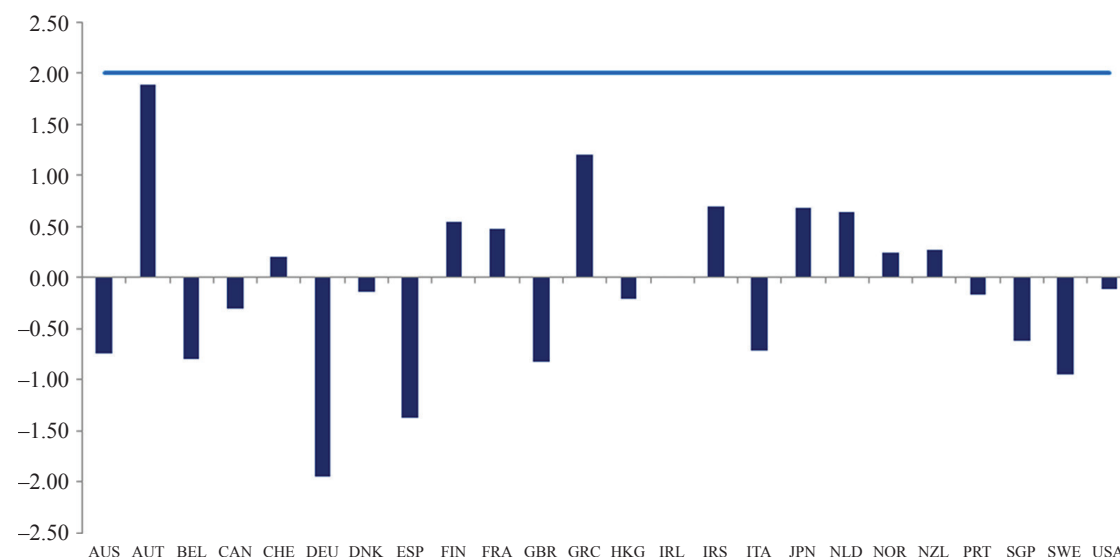


EXHIBIT 13

CAPM Alpha t -Statistic of Country Size Effect



We can also look country by country at how size has fared. Exhibit 13 plots the t -statistics of the CAPM alphas for each of the 24 countries we examine. As the exhibit shows, none of the t -statistics for the CAPM alphas of the country-level SMB portfolios is statistically significant. (The standard 2.0 threshold for significance

for a t -statistic is highlighted). The highest positive t -statistic is for Austria, and it is only 1.89. Moreover, 13 of 24 countries exhibit a *negative* t -statistic, in which the average CAPM alpha for the country-level SMB portfolio is actually negative, not positive. Controlling for market returns, there is more evidence to support a

EXHIBIT 14

Emerging Markets

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|---------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| Emerging Mkts | 3.8% | 14.0% | 1.31 | 0.27 | 3.7% | 1.27 | 0.27 | 7/31/1994 | 12/31/2017 |
| U.S. | 2.6% | 16.6% | 0.75 | 0.16 | 1.6% | 0.47 | 0.10 | 7/31/1994 | 12/31/2017 |

Note: Portfolio construction is the top/bottom 10% of the universe, cap weighted, rebalanced annually.

negative size premium than a positive one, though the evidence is most consistent with there being no size premium at all.

We also look at emerging markets. Here, the historical time series of returns is more limited (beginning in 1994). Exhibit 14 reports results for emerging markets as a region and for the United States over the same time period.

The size premium in emerging markets is positive and larger than it is in the United States, but it still remains insignificant (a *t*-statistic of only 1.31 for raw returns and 1.27 for CAPM alpha). In addition, and per a later discussion on the relationship between size and illiquidity, if we include the lagged return on the market for emerging markets, to account for the nonsynchronous trading effects in less liquid securities such as those that trade in emerging markets, the CAPM alpha declines to 2.3% with a *t*-statistic of 0.80 (not reported).

Finally, the international samples cover a period over which the U.S. size premium is weak (1984–2017). Hence, these are not completely independent tests. Nevertheless, nearly every country fails to deliver a size effect in this sample, so the poor performance of size over this period is robust in every country.

Overall, there is little evidence of a size premium in other equity markets globally. This finding highlights another robustness test the size effect seems to fail.

FACT: THE SIZE EFFECT IS EITHER NOT APPLICABLE OR DOES NOT WORK FOR OTHER ASSET CLASSES

Another virtue of some of the leading asset pricing factors is that they can be applied more broadly to other asset classes. For example, value, momentum, carry, and defensive/quality factors have all been shown to work well in explaining returns in other asset classes, such as

fixed income, credit, currencies, commodities, equity index futures, and options (see Asness, Moskowitz, and Pedersen 2013; Frazzini and Pedersen 2014; Asness et al. 2015b; Israel, Palhares, and Richardson 2018; and Koijen et al. 2018). The application of a factor to other assets is appealing because general theories of asset pricing are not asset specific; they should apply to any financial claim or asset. Furthermore, using the same characteristic to describe returns in many asset classes provides a unifying framework tying those asset classes together. Finally, looking at other asset classes also provides yet another out-of-sample test to guard against data-mining biases.

Does size also help as a unifying concept across asset classes? No. First, the concept of *size* is more difficult to apply outside of equities: What is the size of a currency or government bond or a commodity? Thus, right away the concept of size is ill-suited to describe returns in many asset classes.

Perhaps we can think a bit more creatively about size in some other asset classes to test for a size premium in those asset classes. We can start by looking at country equity indexes, in which size can be defined as the aggregate sum of market caps of all stocks that make up the index in each country. Examining country index portfolios, we can rank countries by their total stock market capitalization and form an SMB portfolio from among the countries. We examine two universes of country index portfolios: (1) developed markets (containing the 24 country index portfolios from January 1975 to December 2017) and (2) emerging markets (containing 25 emerging country index portfolios from January 1988 to December 2017). We go long the smallest half of countries and short the largest, equal weighting the countries in each leg of the strategy. Exhibit 15 reports the performance of these size-based portfolios among country indexes.

EXHIBIT 15

SMB Country Index Portfolios

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|------------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| Developed | 1.3% | 7.8% | 1.06 | 0.16 | 2.0% | 1.68 | 0.26 | 1/31/1975 | 12/31/2017 |
| Emerging | 0.6% | 3.7% | 0.88 | 0.16 | 0.5% | 0.81 | 0.15 | 1/31/1988 | 12/31/2017 |

Notes: Top/bottom 50%, equal weighted, using total market cap as a size measure. The top/bottom 50% is used because of thin breadth; equal weight is used because the cap difference is large.

EXHIBIT 16

SMB Credit Portfolios

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|------------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| Credit HY | -5.2% | 13.7% | -1.73 | -0.38 | -8.0% | -2.91 | -0.64 | 1/31/1997 | 12/31/2017 |
| Credit IG | 0.5% | 3.2% | 0.76 | 0.17 | -0.2% | -0.25 | -0.06 | 1/31/1997 | 12/31/2017 |

Note: Top/bottom 10%, cap weighted, equal weight bonds within an issuer.

There is no evidence of a size premium among developed equity market or emerging market country indexes. The returns to size are positive for both, but statistically insignificant.¹⁵ This contrasts with what researchers have found for value, momentum, carry, and defensive/quality across these same country indexes (see Asness, Moskowitz, and Pedersen 2013; Asness et al. 2015b; and Kojien et al. 2018), in which significantly positive premiums are present.

Because size is a firm attribute that researchers have applied to equities, another natural place to examine size is the other side of a firm's balance sheet—its corporate debt. Using the market capitalization of the firm's equity, we sort firms into deciles based on size, but instead of going long the equities of the smallest 10% of firms and short the stocks of the largest 10% of firms, we go long and short their respective corporate bonds. Exhibit 16 details the results, separating the

universe of U.S. corporate bonds into high yield and investment grade (with the investment-grade universe containing about 500 bonds, on average, and the high-yield universe about 450 bonds, on average). The sample period is January 1997 to December 2017.

As the exhibit shows, there is no size premium at all among corporate bonds in the United States. A portfolio long small-firm credit and short large-firm credit produces a negative average return among high-yield bonds of -5.2%, with a CAPM alpha of -8.0% (*t*-statistic = -2.91). This sign is the opposite of that for the size premium claimed in equities: It is a size discount. Among investment-grade bonds, we find nothing—an insignificant size premium of 0.5% in raw returns, with a -0.2% CAPM alpha. The evidence for a positive size premium in credit is simply not there, which is consistent with results reported by Palhares and Richardson (2018), who found no evidence that the size of a bond issue is correlated with future bond excess returns.

For other asset classes, the notion of size is less direct. For example, we can examine currencies by looking at the size of various countries, using their gross domestic product (GDP) as a measure of economic size (see Hassan 2013). We could do the same for government bonds. However, as Levine et al. (2017) argued,

¹⁵For the developed countries, the CAPM beta-adjusted return is higher than the raw annual return because the loading on the market return is negative and small. Including the lagged market return (not reported) in the regression does not change the estimated CAPM beta very much, suggesting that this effect is unrelated to liquidity effects. In either case, the CAPM alpha is not statistically significant at conventional levels.

EXHIBIT 17

SMB FX Portfolios

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|------------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| Developed | 0.1% | 2.3% | 0.34 | 0.06 | 0.2% | 0.41 | 0.07 | 1/31/1980 | 12/31/2017 |
| Emerging | −1.6% | 4.8% | −1.52 | −0.33 | −1.0% | −0.95 | −0.21 | 1/31/1997 | 12/31/2017 |

Notes: Top/bottom 50%, equal weighted, using total market cap as a size measure. The top/bottom 50% is used because of thin breadth; equal weight is used because the cap difference is large.

the analogy of size in equity markets does not apply easily to commodity markets. For instance, one could use open interest or volume as measures of size, but they seem more related to liquidity. Production weights could also be used as a size measure, which matches the weights in the Goldman Sachs Commodity Index (GSCI). Levine et al. (2017) examined the relationship between various measures of size in commodity futures markets and found no systematic relationship between size and average returns. Hence, size does not appear to be a useful predictor of returns in commodities.

To test one of these markets, we apply a size-based strategy to currencies, in which we use the GDP of each country to rank currencies relative to the U.S. dollar. Specifically, we go long the currencies (relative to the U.S. dollar) of the smallest half of countries, short the currencies of the largest half of countries, and equal weight the countries in the long and short legs. We do this for both developed markets (24 currencies from January 1980 to December 2017) and emerging markets (23 currencies from January 1997 to December 2017). The results in Exhibit 17 show that there is no size premium among currencies either.

Although many studies document significant value, momentum, carry, and defensive/quality return premiums in bonds, country equity index futures, commodities, currencies, and equities globally, we find that size fails to deliver a consistent premium in other asset classes and is less intuitive in other asset classes.

Some might argue that size is really a proxy for liquidity and that if we had looked at liquidity in these other asset classes we might have found a premium. The relationship between size and liquidity is an issue we will discuss later, and the broader concept of liquidity does indeed make more sense and apply more generally to other asset classes. However, the characteristic of size

per se does not seem to capture returns in other asset classes and is often ill-suited as a characteristic for differentiating securities in other asset classes.

FACT: THE SIZE EFFECT MOSTLY COMES FROM MICROCAP STOCKS

One criticism of the size effect is that whatever size premium is present is concentrated in microcap stocks that are extremely small and difficult to trade.¹⁶ We will discuss trading costs and other implementation issues later, but for now we test the conjecture that the returns to size are all concentrated in extremely small stocks. We report size decile 1–10 returns for various subsamples of data, from which we remove the smallest *n*% of firms, and let *n* range from 1% to 30.0%. Exhibit 18 reports the results.

The first row reports the standard result that removes no firms for comparison; we see a 7.2% return (4.3% CAPM alpha) with a *t*-statistic of 2.51 (1.53).¹⁷ As we remove larger fractions of the smallest stocks, the size premium naturally tends to decline. This evidence underscores that the size premium is driven primarily by the smallest stocks.

¹⁶ Horowitz, Loughran, and Savin (2000) found that removing stocks with less than \$5 million in market cap eliminates the small-firm premium. Crain (2011) and Bryan (2014) found that the small-stock effect is concentrated among the smallest 5% of firms.

¹⁷ The entries in the first row differ slightly from those reported in Exhibit 2, with the full-sample statistics for the 1–10 decile portfolio from Ken French's website. For this exhibit, we computed a 1–10 decile portfolio using the sample of U.S. equities from AQR's data library and Fama–French-style construction methods. This series and the one from French's website are 0.95 correlated. We do this step to be internally consistent when we compute the portfolios excluding the smallest *n*% of firms.

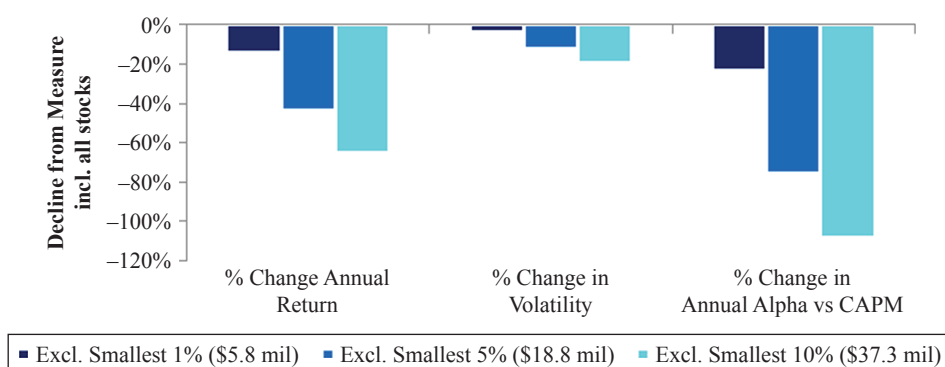
EXHIBIT 18

Size Decile Returns for Data Subsamples

| | Annual Return | Annual Vol | Raw t -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha t -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|-------|---------------|------------|---------------|--------------|-----------------------|--------------------------|-------------|------------|------------|
| 0.0% | 7.2% | 27.4% | 2.51 | 0.26 | 4.3% | 1.53 | 0.16 | 7/31/1926 | 12/31/2017 |
| 1.0% | 6.2% | 26.4% | 2.23 | 0.23 | 3.3% | 1.22 | 0.13 | 7/31/1926 | 12/31/2017 |
| 5.0% | 4.1% | 24.1% | 1.60 | 0.17 | 1.0% | 0.43 | 0.05 | 7/31/1926 | 12/31/2017 |
| 10.0% | 2.5% | 22.0% | 1.09 | 0.11 | -0.3% | -0.16 | -0.02 | 7/31/1926 | 12/31/2017 |
| 20.0% | 0.7% | 18.8% | 0.34 | 0.04 | -1.9% | -1.03 | -0.11 | 7/31/1926 | 12/31/2017 |
| 30.0% | 0.4% | 18.1% | 0.20 | 0.02 | -2.3% | -1.30 | -0.14 | 7/31/1926 | 12/31/2017 |

EXHIBIT 19

Change in Returns, Volatility, and Alpha When Excluding Smallest Firms



Note: June 2016 cap break points are given in parentheses.

On a risk-adjusted basis (relative to the market), there is no size effect because all of the alphas are indistinguishable from zero. However, the point estimate of the alpha declines rapidly when we remove the smallest 5% of firms as well, which corresponds to firms with a market cap of about \$18.8 million—a size that is well below the Russell 3000 minimum, for instance. Hence, to the extent there is a premium for small stocks, the premium does indeed appear to be concentrated among the tiniest 5% of firms.

To see the influence of these tiny firms more clearly, Exhibit 19 plots the percentage change in average returns, volatility, and alpha of the size factor when the smallest 1%, 5%, and 10% of firms are removed from the portfolio. The rapid decrease in returns as the smallest fraction of firms are removed is evident. This fact has been used to argue that the small-firm effect is difficult to trade because the smallest percentage of firms is highly illiquid,

volatile, and expensive to trade. We next take up this issue, which is also related to the extreme price impact experienced by such small stocks in January.

FACT: THE SIZE EFFECT IS DIFFICULT TO IMPLEMENT IN PRACTICE

The fact that the bulk of the size effect seems to come from microcap stocks (as described previously) indicates that implementing a trading strategy to exploit the size effect might be difficult and costly in practice. There are many ways to measure liquidity and no real consensus on which measures are best. To address this issue, we examine a variety of liquidity measures and variables designed to estimate the cost of trading these securities. One very simple and intuitive way to capture the liquidity of a stock is to measure its lagged response to market

EXHIBIT 20

Estimation of Trading Costs for Microcap Stocks (value)

| | Market-Cap Percentile | Lagged Beta | | Frazzini, Israel, and Moskowitz [2018] AQR trading costs (MI in bps) | Amihud [2002] | TAQ Spread | TAQ Price Impact (Kyle [1985] lambda) | Proportion of Zero Return Days | Roll [1984] |
|-------|--------------------------|-------------|--------|--|------------------|---------------|--|--------------------------------------|-------------|
| | | Coeff. | t-Stat | | | | | | |
| Small | Bottom 5% | 0.38 | 8.35 | 63 | 228,042 | 264 | 2.34 | 43% | 159 |
| | Decile 1 | 0.31 | 8.84 | 61 | 132,771 | 225 | 1.82 | 37% | 137 |
| | 2 | 0.19 | 7.52 | 51 | 38,472 | 138 | 1.05 | 27% | 96 |
| | 3 | 0.16 | 8.10 | 38 | 17,851 | 96 | 0.68 | 23% | 74 |
| | 4 | 0.11 | 6.43 | 31 | 8,776 | 72 | 0.46 | 20% | 57 |
| | 5 | 0.08 | 5.57 | 26 | 4,777 | 56 | 0.32 | 18% | 45 |
| | 6 | 0.05 | 4.46 | 22 | 2,544 | 43 | 0.21 | 16% | 31 |
| | 7 | 0.04 | 4.31 | 19 | 1,350 | 33 | 0.14 | 14% | 21 |
| | 8 | 0.01 | 1.29 | 16 | 656 | 24 | 0.08 | 12% | 12 |
| | 9 | 0.00 | -0.69 | 13 | 272 | 16 | 0.04 | 10% | 9 |
| Large | Decile 10 | -0.02 | -4.61 | 10 | 83 | 10 | 0.02 | 7% | 9 |

information (Hou and Moskowitz 2005). An illiquid stock that does not trade continuously or update its price continuously will lag the market. When market news moves stock prices, the most liquid securities will move immediately and contemporaneously because trading takes place nearly continuously in such stocks. Illiquid securities may not trade for some time and prices may not move for some time; hence, they will only reflect the market news with a time lag.

To test this idea, we include lagged market returns in a CAPM regression. The lagged market exposure controls for the possibility of nonsynchronous trading effects in small, illiquid stocks, which do not trade as frequently as larger, more liquid ones, especially early in the sample period. If small stocks slowly lag market movements, their market betas will be biased toward zero. Asness et al. (2018) showed that including the lagged market return increases the market beta exposure of long-short SMB stock portfolios and reduces their alphas with respect to the market. Thus, the CAPM works even better at explaining the size effect after taking lagged market exposure into account.

In addition to reporting lagged market betas as proxies for illiquidity, we also report various other measures of liquidity and trading costs from the academic literature and examine how each varies with firm size. These measures indicate how costly it is to trade various size portfolios in practice and provide a sense of how difficult the strategy is to implement.

Exhibit 20 examines these measures of liquidity and trading costs across size decile portfolios, ranging from the smallest 10% of stocks (decile 1) to the largest 10% of stocks (decile 10). The first two columns report the estimated coefficient on the lagged market return (column 1) and its *t*-statistic (column 2) from a regression of each portfolio's returns on the market portfolio's returns and lagged market returns. The estimated coefficients are strongly decreasing in market cap from decile 1 to decile 10. Portfolios of small stocks have larger lagged betas that are strongly significantly different from zero, whereas portfolios of large stocks exhibit essentially zero lagged beta. The lagged betas of small-stock portfolios are not only statistically significant but economically meaningful. The smallest decile of stocks has a lagged beta of 0.31 (with a *t*-statistic of 8.84), implying that it has a large lagged response to market news. To put this into perspective, if the market risk premium is, on average, 5% per year, then the smallest decile's CAPM alpha will be reduced by an additional $0.31 \times 5\% = 1.55\%$ per year from its lagged response to the market. Put differently, failing to add the lagged return on the market in a CAPM regression would artificially inflate the smallest decile's returns by 1.55% relative to the market. This is a significant difference and suggests a sizable liquidity difference between small and large stocks.

The third column reports the trading cost measure of Frazzini, Israel, and Moskowitz (2018), which

is a measure of market impact from a calibrated model estimated from live executed trades from AQR Capital Management. As the exhibit clearly shows, price impact costs monotonically decrease with firm size: The smallest decile of stocks has an average 61 bps of price impact, whereas the largest decile of stocks experiences only 10 bps of price impact. It is important to recognize that the direction of this effect is tautological. The market value of equity is one of the explanatory variables included in the model. The magnitudes of the effects are nevertheless interesting: The significant differences in cost would severely affect the return differences between small- and large-cap stocks, which were slight anyway.

The remaining columns of the exhibit report results for other cost and liquidity measures, including that of Amihud (2002), who used the daily absolute price change divided by daily share turnover as a measure of illiquidity; the effective bid–ask spread from the Trade and Quote (TAQ) data from the exchanges; a measure of price impact from the TAQ data suggested by the Kyle (1985) model (see Frazzini, Israel, and Moskowitz (2018) and Hasbrouck (2009) for details on how to calculate this measure); the proportion of zero return days (as suggested by Hasbrouck 2009 and Goyenko, Holden, and Trzcinka 2009); and a modified version of the Roll (1984) illiquidity measure, which is the square root of the negative of the autocovariance of daily log prices over the last month and is designed to capture temporary price movements from liquidity trading. For details on the computation of these measures and how they relate to actual transaction costs, see Frazzini, Israel, and Moskowitz [2018].

As can be seen from the exhibit, all of the measures of costs and illiquidity decline steadily when moving from the smallest decile to the largest.¹⁸ The first row depicts the measures for the stocks in the bottom 5% of the size distribution, and the following 10 rows give the measures for each of the size deciles.¹⁹ The illiquidity and cost measures are the highest in the smallest 5%

of stocks and decline monotonically with the size of the stock. This evidence suggests that a portfolio tilted toward smaller stocks and away from larger stocks will suffer from poorer liquidity and larger transaction costs.

Furthermore, the more weight that is given to the smallest stocks at one extreme of the size spectrum, the higher the costs and the worse the liquidity. Because a size-based strategy, like those proposed in the literature, requires investing in microcap stocks in deciles 1 and 2, the returns to these strategies are significantly affected by transaction costs that tend to eliminate what little premium might exist. For example, take the most optimistic of our results on the size premium using a portfolio long decile 1 and short decile 10 that generated a 2.0%–2.5% alpha over the market. The trading costs associated with that long–short portfolio would eliminate most of the return premium. Using the full CRSP universe, trading SMB at \$200 million incurs about 88 bps per year; at \$2 billion, about 152 bps per year; and at \$5 billion, 240 bps per year (Frazzini, Israel, and Moskowitz 2018). Focusing purely on January, in which all of the size returns seem to occur, would similarly be hampered significantly by trading costs (and constrained by liquidity as well, especially at larger portfolio sizes). Hence, a size-based strategy is hindered by liquidity and transaction costs that make it difficult to implement in practice.

That said, there are, in principle, ways to reduce the costs of implementing a size-based strategy. For example, in practice, people do not trade exactly to SMB and allow some tracking error to it. An actual size-based strategy would take into account expected trading costs and would deviate from the SMB portfolio as it traded off tracking error and implementation costs. Given the tight relationship between the attractiveness of size and transaction costs documented earlier, however, the scope for making this trade-off is likely to be more limited than it is for other factors like HML and UMD (see Frazzini, Israel, and Moskowitz 2018). A separate but related point about the costs of implementing trading strategies was made by Frazzini, Israel, and Moskowitz (2018). They showed that trading costs can be reduced by combining multiple factors that are not perfectly correlated to each other. Much like the diversification benefits a portfolio can achieve with less-correlated return sources, a portfolio can also benefit from diversification in trading costs. Combining size with other factors can lower trading costs at the margin that may

¹⁸ As another sign of illiquidity, we find that smaller deciles load more positively on lagged market returns, consistent with non-synchronous trading for small, illiquid stocks.

¹⁹ The results for the lagged beta measure of liquidity for the stocks in the bottom 5% are constructed using AQR's data library and are consistent with standard Fama–French sorts. The portfolio is market cap weighted, rebalanced annually, and spans the sample period of July 1950 to December 2017. The decile portfolios used in the regressions are from Ken French's website.

EXHIBIT 21

Size Portfolio Regressed on Liquidity Factors, Market Return, and Lagged Market Return

| | Alpha vs. Liquidity | Alpha <i>t</i> -Stat vs. Liquidity | Liquidity Beta | Liquidity Beta <i>t</i> -Stat | Start Date | End Date |
|--|------------------------|---------------------------------------|----------------|----------------------------------|------------|------------|
| Panel A: SMB | | | | | | |
| Lagged Beta | 0.1% | 0.09 | 0.13 | 7.45 | 8/31/1926 | 12/31/2017 |
| Frazzini, Israel, and Moskowitz [2018] | −4.3% | −3.05 | 0.51 | 19.25 | 7/31/1988 | 6/30/2017 |
| AQR trading costs (MI in bps) | | | | | | |
| Amihud [2002] | −2.6% | −2.74 | 0.41 | 27.91 | 7/31/1963 | 6/30/2017 |
| TAQ Spread | −0.7% | −0.41 | 0.48 | 15.73 | 7/31/1993 | 6/30/2017 |
| TAQ Price Impact (Kyle [1985] lambda) | 0.7% | 0.34 | 0.38 | 8.68 | 7/31/1993 | 6/30/2017 |
| Proportion of Zero Return Days | −2.1% | −1.74 | 0.38 | 15.00 | 7/31/1963 | 6/30/2017 |
| Modified Roll [1984] | 0.9% | 0.90 | 0.38 | 23.79 | 7/31/1963 | 6/30/2017 |
| Panel B: 1–10 Decile | | | | | | |
| Lagged Beta | −0.1% | −0.03 | 0.33 | 8.77 | 8/31/1926 | 12/31/2017 |
| Frazzini, Israel, and Moskowitz [2018] | −7.1% | −4.05 | 0.84 | 25.79 | 7/31/1988 | 6/30/2017 |
| AQR trading costs (MI in bps) | | | | | | |
| Amihud [2002] | −4.9% | −4.13 | 0.74 | 39.12 | 7/31/1963 | 6/30/2017 |
| TAQ Spread | −1.1% | −0.49 | 0.76 | 19.56 | 7/31/1993 | 6/30/2017 |
| TAQ Price Impact (Kyle [1985] lambda) | 1.2% | 0.41 | 0.62 | 10.11 | 7/31/1993 | 6/30/2017 |
| Proportion of Zero Return Days | −4.5% | −2.61 | 0.73 | 20.28 | 7/31/1963 | 6/30/2017 |
| Modified Roll [1984] | 1.2% | 0.81 | 0.62 | 26.22 | 7/31/1963 | 6/30/2017 |

make a size factor valuable in combination with other factors. Given other research on size's interaction with quality and how the two factors are negatively correlated and can enhance each other's returns, combining size with quality can also mitigate some of the transaction costs.²⁰ Taken together, a portfolio of size and quality has both a higher return premium and lower trading costs than a stand-alone size strategy and hence may be more implementable.

FICTION: THE SIZE EFFECT IS LIKELY MORE THAN JUST A LIQUIDITY EFFECT

As alluded to by the previous facts and fictions, size is closely related to measures of liquidity. In fact, many

scholars have argued that size is really just a proxy for a liquidity effect and that better and more direct measures of liquidity can explain the size effect. In short, these papers argue that there is no size effect, per se, but that it is instead a poor proxy for a stronger liquidity effect. Size may just be a proxy for illiquidity and liquidity risk, and investors generally require compensation for holding such securities.²¹

We can test this idea using factors that attempt to capture liquidity return premiums more directly and see whether the size premium is related to these liquidity factors. For the lagged beta, we regress either the standard SMB (Exhibit 21, Panel A) or the standard 1–10 decile (Exhibit 21, Panel B) portfolio's returns (from Ken French's website) on the market and the lagged market returns. For the other liquidity measures, we form long–short portfolios that invest in the 10% least liquid securities (based in turn on the measures from Frazzini, Israel, and Moskowitz (2018) and Amihud (2002), TAQ

²⁰ Asness et al. (2017) studied the link between liquidity and size when controlling for quality and found that although there is a tight relationship between size and liquidity, there is little relationship between liquidity and quality measures—high-quality small stocks face liquidity similar to that of junky small stocks. They argued that these results are consistent with liquidity-based theories for the size premium, in which size is also correlated with a quality factor that is unrelated to liquidity, and so the size–liquidity relationship may be partly obscured by quality. Hence, size seems to be related to both illiquidity (positively) and quality (negatively), but where liquidity and quality are not strongly related.

²¹ A large amount of literature argues that the returns to size are captured by measures of illiquidity (Brennan and Subrahmanyam 1996; Amihud 2002; Hou and Moskowitz 2005; Sadka 2006; Ibbotson et al. 2013) and measures of liquidity risk, such as those of Pastor and Stambaugh (2003) and Acharya and Pedersen (2005). Crain (2011) summarized this evidence.

effective spread and lambda, proportion of zero return days, and Roll [1984]) and short the 10% most liquid and compute their returns over time. We then regress the returns of the size portfolios on the liquidity factors, the market return, and the lagged market return and report the results in Exhibit 21.

The first two columns report the alphas and the t -statistics of the alphas of the size portfolios on each liquidity factor portfolio. None of the alphas is significantly positive, indicating that the size factor does not deliver excess returns above any of the liquidity factors. The next two columns report the liquidity betas and their t -statistics. The liquidity betas are all economically large and highly statistically significant, which shows that the returns of the size strategy are highly correlated with the returns of the liquidity strategies, consistent with the results of the previous section. What is more, Panel B shows that the liquidity betas of size 1–10 portfolios are consistently larger than those of the SMB portfolio. This finding makes sense because the 1–10 portfolio difference is more extreme than the SMB portfolio, and it underscores the importance of liquidity effects in size portfolios. These results are largely consistent with the literature claiming that size is a proxy for liquidity and that any detectable size premium is well captured by a liquidity premium.

FICTION: THERE IS A STRONG ECONOMIC STORY, EX LIQUIDITY, IN WHICH SMALL STOCKS SHOULD DESERVE A MARGINAL PREMIUM OVER THEIR OTHER RISK CHARACTERISTICS

Another possible explanation for the size premium is that size is distinctly related to expected returns beyond just liquidity. Here, size would play a special role that is linked to a return premium above and beyond the one that compensates investors for liquidity. Although, as discussed earlier, the data seem to confirm that size has trouble predicting returns beyond liquidity measures, we can also evaluate this statement on theoretical grounds.

If size per se carries a return premium, then there must be an economic story for why size, separate from liquidity, should be related to expected returns. Aside from liquidity, why would size be a characteristic that could drive returns? We can appeal to the two leading paradigms for thinking about return predictability: risk-based explanations consistent with efficient markets and

behavioral mispricing explanations consistent with less-than-perfectly efficient markets.

Among risk-based stories, the size of an asset has to be related to the covariance of that asset's return to some underlying economic source of risk for which investors require compensation. Small stocks may simply have higher betas on those sources of risk—such as the market, macro variables, and so on. For example, one explanation for the size effect is that smaller firms have greater exposure to earnings growth, which is a fundamental source of nondiversifiable risk (e.g., Penman et al. 2018). However, this explanation is not about size per se, but rather about size being an indirect measure of the relevant variable—expected earnings growth.²² In fact, we already showed that the CAPM does a good job of explaining size's returns; hence, there is not a size effect per se, and size is just picking up higher beta stocks rather than another risk factor related to size, such as expected earnings growth. This, therefore, is not a story about size carrying a premium.

Similarly, on the behavioral side, scholars have suggested that small stocks are harder to arbitrage and trade (indeed, we found they have much higher trading costs and illiquidity), and hence there will be more mispricing associated with them. For such stocks to carry a return premium, however, they must be systematically underpriced. In principle, mispricing should be equally likely to cause overpricing as underpricing, but if anything, behavioral-finance theory following Miller (1977) predicts that small stocks are more likely to be overpriced than to be underpriced. The idea is that it is harder to short sell small stocks, so their prices primarily reflect the views of optimists and are therefore overvalued. This implies a *negative* size premium. To explain why size itself is compensated, it must be that people demand a larger return (lower price) to trade in small, illiquid, and costly-to-trade (and short) stocks—but this sounds exactly like an illiquidity premium story. The case for size itself to matter seems harder to make.

Finally, size as a characteristic that drives returns is a strange notion compared with other characteristics such as value (book-to-price), momentum (past year

²² Penman et al. (2018) argued that accounting standards encourage firms to report risky activities through deferrals and accruals, which depress current earnings during risky times and cause earnings and book multiples to diverge. Book values become more important when earnings growth is higher and more uncertain, as it is for small stocks.

EXHIBIT 22

Factor Performance among Small-Cap and Large-Cap Stocks

| | Annual Return | Annual Vol | Raw <i>t</i> -Stat | Sharpe Ratio | Annual Alpha vs. CAPM | Alpha <i>t</i> -Stat vs. CAPM | IR vs. CAPM | Start Date | End Date |
|---------------------|------------------|---------------|-----------------------|-----------------|-----------------------------|-------------------------------------|----------------|------------|------------|
| BAB Large | 4.2% | 11.5% | 3.45 | 0.37 | 6.0% | 5.12 | 0.55 | 1/31/1931 | 12/31/2017 |
| BAB Small | 10.5% | 12.3% | 7.96 | 0.85 | 10.0% | 7.54 | 0.82 | 1/31/1931 | 12/31/2017 |
| HML Large | 3.4% | 13.9% | 2.32 | 0.24 | 1.4% | 1.03 | 0.11 | 7/31/1926 | 12/31/2017 |
| HML Small | 5.9% | 12.6% | 4.46 | 0.47 | 5.3% | 4.04 | 0.43 | 7/31/1926 | 12/31/2017 |
| MOM Large | 6.5% | 17.9% | 3.45 | 0.36 | 8.9% | 4.98 | 0.53 | 1/31/1927 | 12/31/2017 |
| MOM Small | 9.3% | 16.3% | 5.46 | 0.57 | 11.6% | 7.12 | 0.75 | 1/31/1927 | 12/31/2017 |
| Profitability Large | 2.2% | 8.7% | 1.86 | 0.25 | 3.1% | 2.73 | 0.37 | 7/31/1963 | 12/31/2017 |
| Profitability Small | 3.8% | 9.1% | 3.12 | 0.42 | 4.4% | 3.58 | 0.49 | 7/31/1963 | 12/31/2017 |

returns), and quality (profits-to-assets). For example, if the cost of capital were a function of size, then this by itself would be a reason to merge; we would observe more mergers, even across very different industries and types of businesses, than we actually do. Thus, size would be an unusual return predictor in an economic model. On the other hand, when two firms merge, their value, momentum, and quality characteristics are averaged because they are ratios. Hence, the cost of capital predicted by these characteristics following a merger would be the average cost of capital of the two firms. This makes intuitive sense.

Therefore, although the data do not seem to yield a large size premium above and beyond any illiquidity premium, theory, too, struggles with why size per se would provide a return premium separate from market risks and liquidity.

FICTION: MANY ANOMALIES BEING STRONGER AMONG SMALL STOCKS IS EVIDENCE OF A SIZE EFFECT

The first part of this statement is true, but the latter part is false. The size effect—that small stocks outperform large stocks—is often confused with other factors, such as value, being stronger among small stocks than among large stocks. Many anomalies (though not all) are indeed stronger among small stocks, but this has nothing to do with the “size effect,” or more precisely a return premium for size per se. This statement is about other return premiums being stronger (at least gross of trading costs) among smaller-cap stocks. This could be because

of illiquidity, more limited arbitrage, higher volatility, or more retail investors associated with small stocks, all of which may exacerbate any return premium associated with other factors, but none of which necessarily have anything to do with a premium associated with small firms themselves.

For example, when looking at other factors among small- versus large-cap stocks, the factors are neutral to size. When we look at another factor, such as value, that is long value stocks (high BE/ME) and short growth stocks (low BE/ME) within the small-cap universe, the value stocks are, on average, the same size stocks as the growth stocks. A value factor that is long value and short growth among small-cap stocks would net out any small-cap exposure. The only return premium being picked up here is a value premium *among* small stocks. Any size premium is effectively hedged away.

Exhibit 22 reports the performance of various factors—HML, BAB, UMD, RMW—formed among small-cap and large-cap stocks separately as described earlier, over the longest sample period for which we have data. Each factor is neutral to size because the longs and the shorts of each leg of the factor have equivalent size characteristics.

As the exhibit shows, for every factor, the returns, Sharpe ratio, alpha, and information ratio are a bit higher when formed among small-cap stocks than among large-cap stocks, especially for BAB and HML.²³ Thus, other factor returns do indeed seem to be a bit stronger (gross of transaction costs) among small-cap stocks.

²³ Note the lack of any value premium among large-cap stocks, a point discussed by Asness et al. (2014).

However, none of the additional returns to these factors is driven by a size premium because each factor is neutral to size.

FACT: OTHER FACTORS WORKING BETTER IN SMALL CAPS CAN BE A REASON TO BE OVERWEIGHT SMALL STOCKS EVEN THOUGH THERE IS NO PURE SIZE EFFECT

Although other factors working better among small stocks is not evidence of a size effect per se, as the previous section highlights, it still might warrant overweighting small stocks. Other factors may exhibit slightly higher returns among small stocks simply because those stocks are less liquid, are more difficult to trade, have more idiosyncratic volatility, may have more retail (less sophisticated) investors present, and simply face more limited arbitrage capital, all of which could contribute to enhancing the return premium associated with other factors. However, this does not indicate there is a premium for size per se, only that other premiums are larger when implemented among small-cap stocks.²⁴ Other premiums being stronger among small caps may be a rationale to want to overweight small-cap stocks even if there is no size premium. For example, the value premium in small stocks may be so large that it justifies being overweight small stocks even though there is no stand-alone size effect. Of course, simply being overweight small is not nearly as profitable as being overweight small value. Hence, absent a pure stand-alone size effect, an investor is always better off being overweight certain kinds of small stocks (e.g., those with high value, momentum, and quality) rather than generic small stocks.

Finally, as with most of our analysis here, we are looking at gross of trading cost returns. In practice, returns net of transaction costs often ameliorate any factor performance differences across different size universes because small-cap stocks are more expensive and more difficult to trade. On a net-of-trading-cost basis, the performance of many of these factors is not very different among small-versus large-cap stocks because

small stocks are more difficult and more costly to trade, as our earlier evidence showed. Being overweight small stocks to take advantage of the increase in factor premiums among small caps critically depends on the size of implementation costs and the net-of-cost returns to those factors among small-cap stocks.

FACT AND FICTION: THE SIZE EFFECT IS MUCH STRONGER WHEN CONTROLLING FOR OTHER FACTORS

This one depends critically on what other factors are controlled for. We have already shown that controlling for the market portfolio (CAPM) diminishes the size effect, rendering it insignificant in most cases. As we showed earlier, the CAPM works even better at explaining the size effect if we take account of lagged market exposure to control for the illiquidity of small stocks. But what about other factors? We can also run a regression of SMB's returns on the Fama and French factor HML, as well as the market portfolio. With some abuse of terminology, we refer to this model as FF3 because it contains the relevant factors from the Fama–French three-factor model, excluding SMB of course.²⁵

These regressions are run for SMB formed from market capitalization—the classic measure of size—as well as SMB portfolios formed from the non-price-based measures of size we used earlier (book assets, book equity, employees, PP&E, and sales). We then repeat these regressions by adding the momentum factor—UMD. We refer to this model as FF3 + UMD. The first three bars in each section of Exhibit 23 report the alphas from all of these regressions, along with the CAPM alphas for comparison.

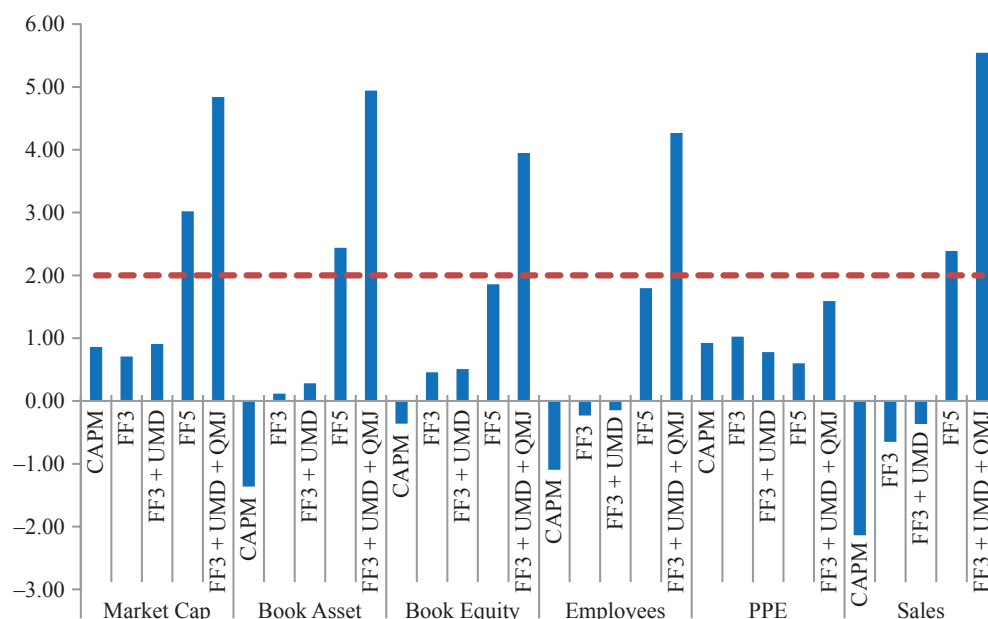
As the exhibit shows, none of the CAPM alphas are significant, with some of the alphas negative and none having a *t*-statistic anywhere close to +2.0. Adding HML does little to change that conclusion, and adding UMD also has a negligible effect on SMB's alpha. This suggests that although controlling for the market makes

²⁴ As another example, researchers often find that factor return premiums are stronger when applied to emerging markets for similar reasons. Again, this indicates factor return premiums are stronger within emerging markets (gross of transaction costs), but that does not mean there is an emerging market premium per se.

²⁵ The real Fama and French (1993) three-factor model contains the market portfolio, SMB, and HML, but because we are using SMB in our analysis as the dependent variable, we obviously cannot include it on the right-hand side of the regression as a control. Thus, we instead refer to this model as FF3.

EXHIBIT 23

t-Statistic of SMB Alpha



the size effect weaker, it is relatively immune to controls for value and momentum.

The exhibit also shows the results of a regression of the same SMB portfolios against factors from Fama and French's new five-factor model (FF5), which contains MKT, HML, and the two new factors from Fama and French (2015): RMW, a profitability factor that is long "robust" or profitable firms (high operating profits-to-assets) minus "weak" unprofitable firms; and CMA, an investment factor that is long "conservative" firms with low investments-to-assets and short "aggressive" firms with high investments-to-assets.²⁶ Here, the story changes considerably. Suddenly, the alpha of SMB is positive and strongly significant with a *t*-statistic of 3.02. Moreover, and even more interestingly, the SMB alphas from the non-price-based measures of size are also now significantly positive, with *t*-statistics ranging from just under 2.0 to 2.4. In other words, the size effect seems to have been made substantially stronger by including the two new Fama and French factors RMW and CMA.

²⁶ Again, the Fama and French (2015) five-factor model contains the market portfolio, SMB, HML, RMW, and CMA, but because we are using SMB in our analysis as the dependent variable, we drop SMB as a control and refer to this model as FF5.

Digging into this result, it is the relationship between SMB and RMW that is driving it.

Why does the size effect become significantly stronger when controlling for the profitability factor? Because, as Asness et al. (2018) showed, the size effect is confounded by a very powerful quality versus junk effect. They investigated the relationship between size and quality and found that controlling for quality not only resurrects the size premium and elevates it significantly but also helps resolve some of the aforementioned patterns associated with size.

The interaction between size and quality is especially interesting for several reasons. First, quality can be defined as a characteristic of an asset that, all else equal, commands a higher price. As such, size, which is based on market values, should have a strong connection to quality, in which size's relationship to average returns is made much clearer once controlling for quality. We can measure firm quality in a variety of ways. Asness, Frazzini, and Pedersen (2017) and Asness et al. (2018) measured it by using profitability, payout, growth, and safety, taking an average of these measures to form a quality factor that is long high-quality stocks and short low-quality/junk stocks; they call this *quality minus junk* (QMJ).

Exhibit 23 adds QMJ and UMD to our FF3 model (FF3 + UMD + QMJ). As the exhibit indicates, the size premium is substantially increased after controlling for QMJ. The t -statistic of the alpha for size jumps from 0.91 when controlling for the FF3 + UMD factors to 4.84 when adding the QMJ factor. The same substantial increase in the size premium also occurs for the non-price-based measures of size because the alpha of each SMB portfolio associated with the non-price size measures jumps from statistically no different from zero (t -statistics all well below 1.0) under the FF3 + UMD factors to highly significantly positive (t -statistics close to and above 5.0).

The interaction between size and quality/junk is far stronger than size's interaction with other factors (beta, value, and momentum), and accounting for it produces a more significant size premium. Regardless of the quality metric used—metrics that vary substantially both qualitatively and in terms of measured correlation—we find a much stronger and more stable size effect when controlling for a firm's quality.

Firm size is highly confounded with firm quality, which distorts the relationship between size and expected returns. Large firms tend to be high-quality firms, whereas small firms tend to be junky. Because high-quality stocks outperform junk stocks, on average, the basic size effect is fighting a strong quality effect. Going long small stocks and short large stocks, a size-based strategy is long a potential size premium but also short a quality premium, which both understates the actual size effect and introduces additional variation from the quality factor.

In addition to resurrecting the size premium, controlling for quality also reconciles many of the empirical irregularities associated with the size effect that we (and the literature) have documented. Controlling for quality resurrects the size effect after the 1980s and explains its time variation, restores a linear relationship between size and average returns among all firms and not just the smallest ones, and revives the returns to size in months other than January while simultaneously diminishing them in January. What is more, conditioning on quality uncovers a larger size effect in almost two dozen international equity markets, where size has been notably weak. These results are robust to using non-market-based size measures, making the size premium a much stronger and more reliable effect after controlling for quality (Asness et al. 2018).

We hasten to add that these findings do not resurrect the pure size effect, the evidence for which is weak. What these findings instead show is that it is possible to uncover a conditional size effect and resolve many of its empirical irregularities when controlling for quality.

FACT: THE SIZE EFFECT RECEIVES DISPROPORTIONATELY MORE ATTENTION THAN OTHER FACTORS WITH SIMILAR OR MUCH STRONGER EVIDENCE

Finally, despite its pretty mediocre evidence and lack of theory, the size effect has received disproportionately more attention than other factors with much stronger evidence and theory behind them. For instance, as we showed earlier, value, momentum, and defensive/quality provide much stronger return evidence than size. Liquidity also seems to provide stronger empirical premium than size. Yet, the size effect has received a lot more attention in the literature than some of these factors.

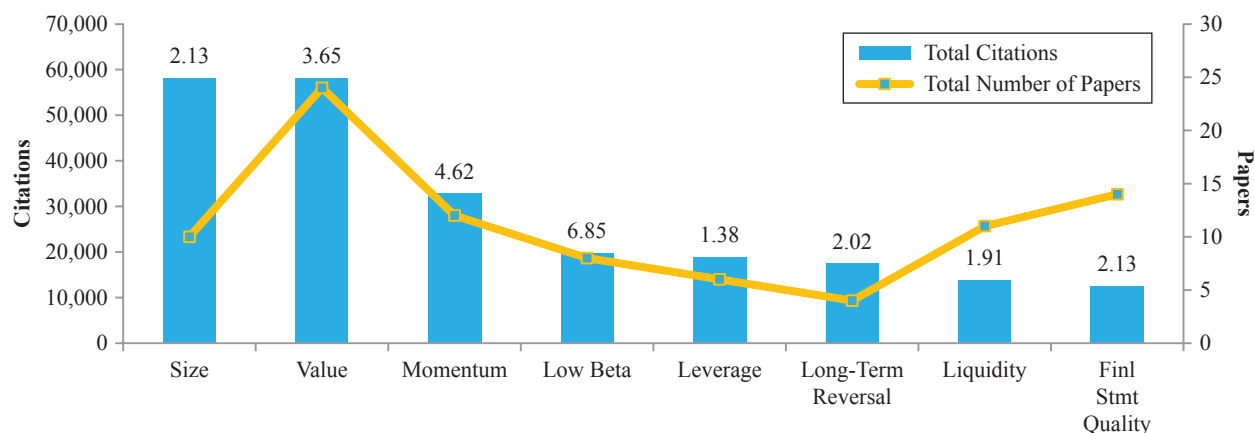
Using Google Scholar, we added up the number of papers that have explicitly focused on the size effect (excluding this one) and added up all of the citations to those articles in the academic finance, accounting, and economic literatures. We then did the same for several other prominent factors in the literature: value, momentum, beta, leverage, reversals, liquidity, and quality (broadly defined as *financial statement quality* [FSQ]). The results are plotted in Exhibit 24 for each factor, along with t -statistics of the raw returns associated with each factor over the longest possible sample period, which begins around 1926–1931 for size, value, beta, reversals, and momentum and around 1951 for liquidity, FSQ, and leverage.

As the exhibit shows, size has received much more attention than just about every factor except value. Comparing the citations versus historical performance of each factor, it is arguable that size has received much more attention than it deserves. The evidence behind size is far more meager than that for many other factors that receive much less attention, and other factors that have similar strength of evidence behind them receive a lot less rumination.

The undue prominence of the size effect in the academic literature and practice is likely because it was the first anomaly to challenge standard asset pricing theory (namely, the CAPM), and a focus in science can often

EXHIBIT 24

Citations, Number of Articles, and *t*-Statistics by Factor Category



be path-dependent. This path dependence also affected the investment industry. Based on those early findings, the investment management industry decided to organize product offerings by size. The truth, however, is that the premium associated with size is not very strong, not very persistent, and not very robust, and it may never have existed in the first place (if not for data errors and improper risk measurement).

CONCLUSION

The bottom line is that, after addressing the facts and fictions of the size effect, we find neither strong empirical evidence nor robust theoretical support for a prominent size premium. This raises a question: Should academics and investors be using a size-based factor?

The answer, we think, depends on the application. If one is trying to understand what drives or explains differences in expected returns across stocks or portfolios or is trying to predict returns, then the answer is no. A size factor simply does not add much and is not very useful at capturing or forecasting average returns. If, on the other hand, one wishes to understand what actual portfolio managers do and evaluate their performance, or explain the time-series variation in managers' portfolios, then a size factor can be quite helpful. Many mutual fund and institutional money managers select specific size-based portfolios and benchmark to size-based indexes such as the Russell 2000. Hence, a size factor can be quite useful in decomposing the returns to these managers.

More broadly, should an investor overweight small stocks to enhance returns? Again, the answer depends on how it is done. Simply generically tilting toward small stocks is unlikely to provide much of a premium. However, our evidence suggests that the success of some other factors, such as value, among small-cap equities implies being overweight those firms will enhance returns, assuming the higher transaction costs of such a strategy permit it. Our research also shows the importance of controlling for quality in identifying a conditional size premium, the practical implications of which we and the industry are still exploring. In sum, we endorse a more nuanced view of the size effect informed by the currently available evidence and recommend rethinking how the notion of size is used to answer academic and practical questions.

ACKNOWLEDGMENTS

We thank Cliff Asness, Jeremy Getson, Antti Ilmanen, Sarah Jiang, Jared Kizer, John Liew, Lasse Pedersen, Scott Richardson, Caroline Sasseville, Rodney Sullivan, and Ryan Wei for useful comments and suggestions.

REFERENCES

- Acharya, V. V., and L. Pedersen. 2005. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics* 77 (2): 375–410.
- Amihud, Y. 2002. "Illiquidity and Stock Returns: Cross-section and Time-Series Effects." *Journal of Financial Markets* 5 (1): 31–56.

- Asness, C., A. Frazzini, and L. Pedersen. 2017. "Quality Minus Junk." June 5, 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2312432.
- Asness, C., T. J. Moskowitz, and L. Pedersen. 2013. "Value and Momentum Everywhere." *Journal of Finance* 68 (3): 929–985.
- Asness, C., A. Frazzini, R. Israel, and T. J. Moskowitz. 2014. "Fact, Fiction, and Momentum Investing." *The Journal of Portfolio Management* 40 (5): 75–92.
- . 2015a. "Fact, Fiction, and Value Investing." *The Journal of Portfolio Management* 42 (1): 34–52.
- Asness, C., A. Iltanen, R. Israel, and T. J. Moskowitz. 2015b. "Investing with Style." *Journal of Investment Management* 13 (1): 27–63.
- Asness, C., A. Frazzini, R. Israel, T. J. Moskowitz, and L. Pedersen. 2018. "Size Matters, If You Control Your Junk." *Journal of Financial Economics* 129 (3): 479–509.
- Ball, R. 1978. "Anomalies in Relations Between Securities' Yields and Yield-Surrogates." *Journal of Financial Economics* 6 (2–3): 103–126.
- Banz, R. 1981. "The Relationship between Return and Market Value of Common Stocks." *Journal of Financial Economics* 9 (1): 3–18.
- Berk, J. 1995a. "A Critique of Size-Related Anomalies." *Review of Financial Studies* 8 (1): 275–286.
- . 1995b. "An Empirical Reexamination of the Relation between Firm Size and Return." Working paper, University of Washington.
- . 1997. "Does Size Really Matter?" *Financial Analysts Journal* 53 (5): 12–18.
- Brennan, M., and A. Subrahmanyam. 1996. "Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns." *Journal of Financial Economics* 41 (3): 441–464.
- Bryan, A. 2014. "Does the Small-Cap Premium Exist?" Chicago, IL: Morningstar Inc.
- Chan, L. K. C., J. Karceski, and J. Lakonishok. 2000. "New Paradigm or Same Old Hype in Equity Investing?" *Financial Analysts Journal* 56 (4): 23–36.
- Chordia, T., A. Subrahmanyam, and Q. Tong. 2014. "Have Capital Market Anomalies Attenuated in the Recent Era of High Liquidity and Trading Activity?" *Journal of Accounting and Economics* 58 (1): 41–58.
- Chui, A. C., S. Titman, and K. J. Wei. 2010. "Individualism and Momentum around the World." *Journal of Finance* 65 (1): 361–392.
- Crain, M. A. 2011. "A Literature Review of the Size Effect." Working paper, Financial Valuation Group.
- Dichev, I. D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" *Journal of Finance* 53 (3): 1131–1147.
- Fama, E. F. 1976. *Foundations of Finance*. New York: Basic Books.
- Fama, E. F., and K. R. French. 1992. "The Cross-section of Expected Stock Returns." *Journal of Finance* 47 (2): 427–465.
- . 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.
- . 1998. "Value versus Growth: The International Evidence." *Journal of Finance* 53 (6): 1975–1999.
- . 2012. "Size, Value, and Momentum in International Stock Returns." *Journal of Financial Economics* 105 (3): 457–472.
- . 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22.
- . 2017. "International Tests of a Five-Factor Asset Pricing Model." *Journal of Financial Economics* 123 (3): 441–463.
- Fama, E. F., and J. D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–636.
- Frazzini, A., and L. Pedersen. 2014. "Betting Against Beta." *Journal of Financial Economics* 111 (1): 1–25.
- Frazzini, A., R. Israel, and T. J. Moskowitz. 2018. "Trading Costs." Working paper, Yale University.
- Gompers, P., and A. Metrick. 2001. "Institutional Investors and Equity Prices." *Quarterly Journal of Economics* 116 (1): 229–259.

- Goyenko, R., C. W. Holden, and C. Trzcinka. 2009. "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics* 92 (2): 153–181.
- Griffin, J. M., X. Ji, and J. S. Martin. 2003. "Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole." *Journal of Finance* 58 (6): 2515–2547.
- Harvey, C. 2017. "Presidential Address: The Scientific Outlook in Financial Economics." *Journal of Finance* 72 (4): 1399–1440.
- Harvey, C., Y. Liu, and H. Zhu. 2016. "... and the Cross-Section of Expected Returns." *The Review of Financial Studies* 29 (1): 5–68.
- Hasbrouck, J. 2009. "Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data." *Journal of Finance* 64 (3): 1445–1477.
- Hassan, T. 2013. "Country Size, Currency Unions, and International Asset Returns." *Journal of Finance* 68 (6): 2269–2308.
- Horowitz, J. L., T. Loughran, and N. E. Savin. 2000. "Three Analyses of the Firm Size Premium." *Journal of Empirical Finance* 7 (2): 143–153.
- Hou, K., and T. J. Moskowitz. 2005. "Market Frictions, Price Delay, and the Cross-Section of Expected Returns." *Review of Financial Studies* 18 (3): 981–1020.
- Hou, K., and M. A. Van Dijk. 2014. "Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns." Working paper, Ohio State University.
- Ibbotson, R., Z. Chen, D. Y. J. Kim, and W. Hu. 2013. "Liquidity as an Investment Style." *Financial Analysts Journal* 69 (3): 30–44.
- Ilmanen, A., R. Israel, T. J. Moskowitz, A. Thapar, and F. Wang. 2018. "What Can We Learn from a Century of Style Premia?" Working paper, Yale University.
- Israel, R., and T. J. Moskowitz. 2013. "The Role of Shorting, Firm Size, and Time on Market Anomalies." *Journal of Financial Economics* 108 (2): 275–301.
- Israel, R., S. Jiang, and A. Ross. "Craftsmanship Alpha: An Application to Style Investing." Available at SSRN: <https://ssrn.com/abstract=3034472>, September 8, 2017.
- Israel, R., D. Palhares, and S. A. Richardson. 2018. "Common Factors in Corporate Bond Returns." *The Journal of Investment Management* 16 (2).
- Jegadeesh, N., and S. Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1): 65–91.
- Keim, D. B. 1983. "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics* 12 (1): 13–32.
- Koijen, R., T. J. Moskowitz, L. Pedersen, and E. Vrugt. 2018. "Carry." *Journal of Financial Economics* 127 (2): 197–225.
- Kyle, A. S. 1985. "Continuous Auctions and Insider Trading." *Econometrica* 53 (6): 1315–1335.
- Levine, A., Y. H. Ooi, M. Richardson, and C. Sasseville. 2017. "Commodities for the Long Run." *Financial Analysts Journal* 74 (2): 55–68.
- Liew, J., and M. Vassalou. 2000. "Can Book-to-Market, Size and Momentum Be Risk Factors that Predict Economic Growth?" *Journal of Financial Economics* 57 (2): 221–245.
- McLean, D., and J. Pontiff. 2016. "Does Academic Research Destroy Stock Return Predictability?" *Journal of Finance* 71 (1): 5–32.
- Miller, E. M. 1977. "Risk, Uncertainty, and Divergence of Opinion." *Journal of Finance* 32 (4): 1151–1168.
- Palhares, D., and S. A. Richardson. "(Il)liquidity Premium in Credit Markets: A Myth?" March 29, 2018, <https://ssrn.com/abstract=3130213>.
- Pastor, L., and R. Stambaugh. 2003. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy* 111 (3): 642–685.
- Penman, S. H., F. Reggiani, S. A. Richardson, and I. A. Tuna. 2018. "A Framework for Identifying Accounting Characteristics for Asset-Pricing Models, with an Evaluation of Book-to-Price." *European Financial Management* (forthcoming): 1–33.
- Reinganum, M. 1983a. "Portfolio Strategies Based on Market Capitalization." *The Journal of Portfolio Management* 9 (2): 29–36.

———. 1983b. “The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects.” *Journal of Financial Economics* 12 (1): 89–104.

Roll, R. 1983. “Vas Is Das? The Turn-of-the-Year Effect and the Return Premia of Small Firms.” *The Journal of Portfolio Management* 9 (2): 18–28.

———. 1984. “A Simple Implicit Measure of the Effective Bid–Ask Spread in an Efficient Market.” *Journal of Finance* 39 (4): 1127–1139.

Rouwenhorst, K. G. 1998. “International Momentum Strategies.” *Journal of Finance* 53 (1): 267–284.

Sadka, R. 2006. “Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk.” *Journal of Financial Economics* 80 (2): 309–349.

Schwert, G. 1983. “Size and Stock Returns, and Other Empirical Regularities.” *Journal of Financial Economics* 12 (1): 3–12.

———. 2003. “Anomalies and Market Efficiency.” In: Constantinides, G. M., M. Harris, and R. M. Stulz, eds., pp. 939–974. Elsevier, Amsterdam: *Handbook of the Economics of Finance*.

Shumway, T. 1997. “The Delisting Bias in CRSP Data.” *Journal of Finance* 52 (1): 327–340.

Shumway, T., and V. A. Warther. 1999. “The Delisting Bias in CRSP’s Nasdaq Data and Its Implications for the Size Effect.” *Journal of Finance* 54 (6): 2361–2379.

Van Dijk, M. A. 2013. “Is Size Dead? A Review of the Size Effect in Equity Returns.” *Journal of Banking and Finance* 35 (12): 3263–3274.

Disclaimer

AQR Capital Management is a global investment management firm and may or may not apply investment techniques or methods of analysis similar to those described herein. The views expressed here are those of the authors and not necessarily those of AQR.

To order reprints of this article, please contact David Rowe at d.rowe@pageantmedia.com or 646-891-2157.