IPR JOURNALS

The Journal of Portfolio Management

JPM.iprjournals.com

QUANTITATIVE STRATEGIES: *FACTOR INVESTING*

FOURTH EDITION 2019

The Characteristics of Factor Investing

David Blitz and Milan Vidojevic

Robeco is an International asset manager offering an extensive range of active Investments, from equities to bonds. Research lies at the heart of everything we do, with a 'pioneering but cautions' approach that has been in our DNA since our foundation in Rotterdam in 1929. We believe strongly in sustainability investing, quantitative techniques and constant innovation.

ROBECO
The Investment Engineers

The Characteristics of Factor Investing

David Blitz and Milan Vidojevic

David Blitz
is the head of quantitative equity research at Robeco Asset Management in Rotterdam, the Netherlands.
d.c.blitz@robeco.nl

Milan Vidojevic
is a quantitative researcher at Robeco Asset Management in New York, NY, and holds a position at Vrije Universiteit Amsterdam in Amsterdam, the Netherlands.
m.vidojevic@robeco.com

Traditional asset allocation models are centered around capturing risk premiums offered by distinct asset classes such as public and private equities, corporate and government bonds, and commodities. Decades of empirical research on asset prices have uncovered a set of factors that drive expected returns within and across asset classes. Explicit, top-down allocation to factor premiums (e.g., size, value, momentum, quality, and low volatility) is now becoming mainstream. Yet many questions remain about how to efficiently gain exposures to these premiums.

Factors are systematic drivers of stock returns—they explain why certain stocks go up or down at the same time (co-move) and why certain stocks command higher expected returns than others. In practice, investment portfolios that provide exposures to these factors are constructed by sorting stocks on certain characteristics and buying those that score favorably and, if implemented in a long–short setting, short selling those that score unfavorably. Over the years, researchers have identified a number of factor characteristics that are key determinants of expected stock returns. Some examples include market capitalization (size), book-to-market (value), gross or operating profitability, asset growth (investment), and past return (momentum). In particular, holding all other characteristics constant, stocks with lower market capitalizations (small-caps), higher book-to-market (value), higher profitability and lower asset growth (quality), and higher recent returns (winners) have delivered high returns relative to their peers with lower scores on these characteristics.

Widely used asset pricing models take the capital asset pricing model (CAPM) as a starting point and then augment it with other factors (e.g., the Fama and French 1993, 2015 three- and five-factor models). These models inspire a top-down approach to factor investing, wherein investors allocate a certain amount to each of these factors. For instance, an investor who wishes to capture the value premium could allocate a part of the portfolio to the theoretical Fama–French high minus low (HML) value factor or, more likely, a long-only value index in practice. A pitfall of the top-down approach, however, is that each factor portfolio is constructed in isolation, with a complete disregard for all other factors. To address this problem, we propose a bottom-up approach instead, wherein the expected return of each stock is a function of all its factor characteristics.

To illustrate, take two hypothetical value stocks. These stocks, by definition, have a high book-to-market characteristic, which contributes positively to their long-term expected returns. However, these stocks can be completely different in terms of other characteristics that also influence their expected returns. If one stock scores favorably on other factors and the other has very poor scores on these same factors, these two stocks will have vastly different expected returns, even though they are very similar in the value (book-to-market) dimension. In fact, some value stocks may even have a lower expected return than the market portfolio because their favorable book-to-market score is entirely offset by poor scores on the other factors.

Simple, cookie-cutter approaches to factor investing that construct factor portfolios to deliver exposure to one select factor often ignore the impact of other factor characteristics on expected stock returns and consequently on expected portfolio returns. Many generic factor products, often labeled smart beta, completely disregard the impact of other factors when constructing portfolios with high exposures to any single factor. A value strategy in its simplest form would invest in all stocks with the ratio of book-to-market equity in the top quintile, quartile, or tercile of the investment universe. However, this clearly ignores the fact that many value stocks could have, for instance, poor momentum or profitability characteristics (which is often the case empirically) that completely dominate the positive expected return contribution of the high value characteristic.

In this article, we show that generic factor strategies experience a significant return drag because of their disregard for other factor exposures. To this end, we use a characteristics-based multi-factor model that consists of the aforementioned six well-established stock-level factor characteristics to estimate the historical premiums that these factor characteristics have delivered. Furthermore, at each point in time, for each stock in the universe, we calculate the model-implied return by multiplying the characteristics of the stock with the estimated factor premiums. Once we have the implied returns on each stock in the universe, we analyze the model-implied and realized returns of five prominent factor investing styles: size, value, profitability, investment, and momentum.

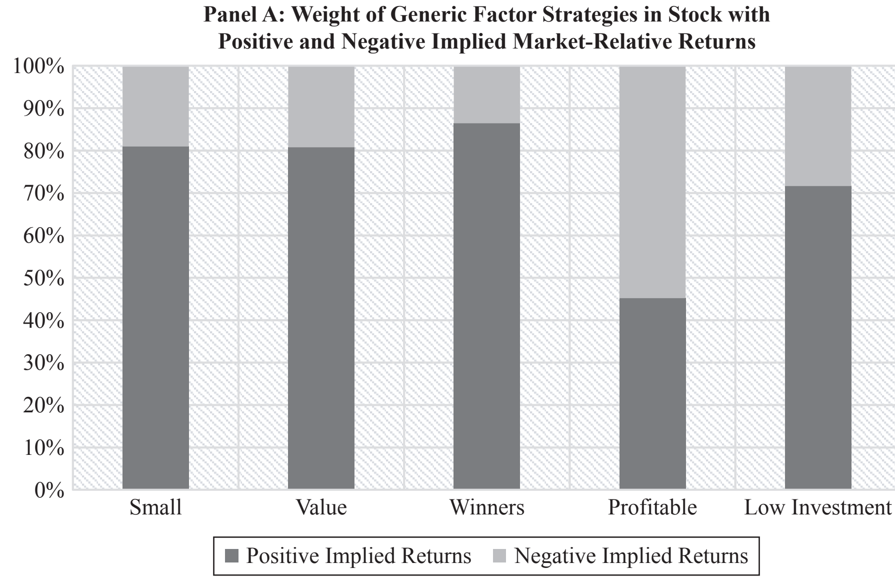
Panel A of Exhibit 1 shows the average weight of each of the five generic single-factor portfolios invested in the stocks with model-implied returns higher (positive) and lower (negative) than the return on the market portfolio. The effect is astonishing: Around 20% of the generic single-factor portfolio weight is invested in stocks that have negative implied market-relative returns. In the case of the profitability style, this number is closer to 50%. This has big implications for the performance of these investment portfolios.

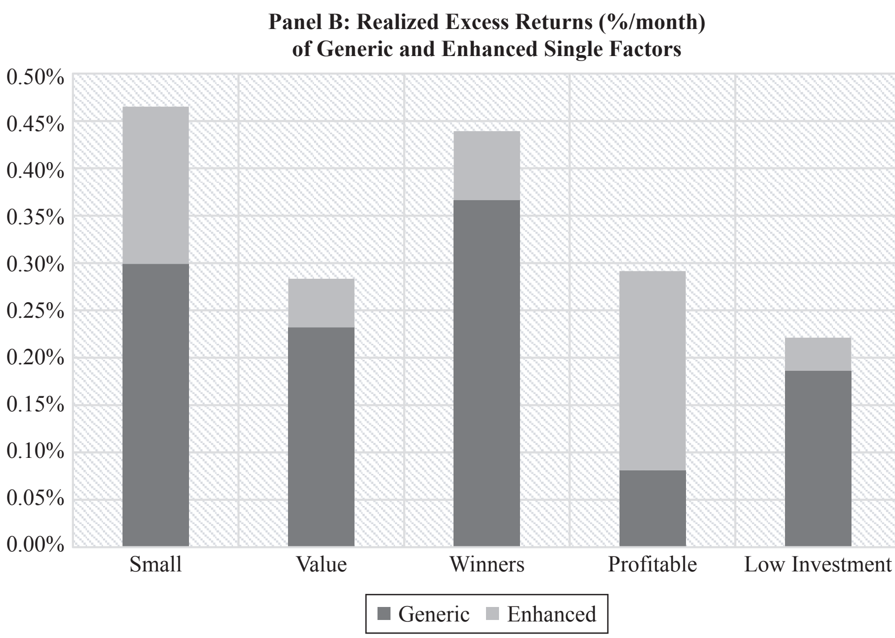
Panel B shows the average realized returns on these portfolios. It also shows the returns on the single-factor portfolios that, at each portfolio rebalancing moment, simply exclude the stocks with such poor characteristics on other factors that their implied returns end up being below those of the market benchmark. The effect on the portfolio performance is substantial, with the largest impact visible for the small-cap and profitability styles.

Our characteristics-based model allows us to address many other important issues. For instance, the market portfolio itself is nothing but a collection of stocks, and at each point in time, each stock has its own model-implied return. This means that we can calculate the weight of the market portfolio that is invested in stocks with such poor factor characteristics that their model-implied returns are even lower than those of government bonds. We find that, on average, around 10% of the market portfolio is invested in stocks with dismal model-implied returns and that removing these stocks from the market each month would result in an increase in the realized return of about 16%, in relative terms.

Recently, there has been a heated debate about the best way to design a strategy that provides exposures to multiple factor premiums simultaneously. Some advocate an integrated approach whereby stocks with the highest combined rank on the selected factor characteristics are chosen; others prefer a mixed-sleeve approach whereby single-factor strategies (sleeves) are used as building blocks for a fund-of-funds–like multi-factor portfolio. Our characteristics-based model allows us to resolve this debate as well. We show that mixing generic single-factor strategies is clearly suboptimal as around 20% or more of each portfolio is invested in stocks that have negative implied market-relative returns. Integrated multi-factor portfolios do not suffer from this issue because they, by construction, select stocks by looking at all of their factor characteristics. However, there is also an in-between approach whereby one first constructs single-factor strategies by selecting stocks that have high characteristics on the targeted factor but do not have extremely poor other factor characteristics; in the second step, these strategies are mixed into one multi-factor portfolio. We show that, so long as

EXHIBIT 1
Generic and Enhanced Single-Factor Strategies





the integrated and enhanced single-factor strategies are designed to provide the same level of factor exposure (i.e., have comparable overall factor characteristics), their model-implied and historically realized returns are in fact not statistically different.

Lastly, through the lens of our model, we also address some longstanding questions about why different weighting schemes have delivered different results or why many factors appear to be stronger in the small-cap segment of the market. We show that the differences in performance among all these portfolios can be fully explained by the differences in their factor characteristics.

Although the characteristics-based model that we apply in this article can be used to forecast out-of-sample stock and portfolio-level expected returns, this is not how we apply it in this article. Here we use the model to ex-post explain why certain factor strategies performed as they did. That is, the fact that we use the full-sample estimates of factor premiums restricts us from making any statements about future expected returns, but because we are applying the model to understand what happened in the past data, using the full-sample estimates of factor premiums is not an unreasonable choice. We further elaborate on this choice in the data and methodology section of the article.

We conclude that to implement factor investing efficiently, investors should always consider all relevant drivers of expected stock returns (i.e., multiple factor characteristics). A characteristics-based multi-factor return model can be used to provide insights into the return drivers of equity portfolios.

DATA AND METHODOLOGY

To estimate the return premiums associated with various stock-level characteristics, we make use of cross-section regressions, building on the Fama and MacBeth (1973) procedure. This means that at the end of each month, we regress stock returns in excess of the risk-free rate on characteristics measured at the end of the previous month. In this way, we obtain a time series of estimated premiums per unit of each characteristic, holding all other characteristics constant. We then average these coefficients over the whole length of the sample to obtain the average premiums per unit of exposure to a particular characteristic (i.e., factor).

The decision to use the full sample estimates of factor premiums can be justified by the fact that the purpose of our model is not to forecast the future; instead, we want to quantitatively describe ex-post the returns that various factor-based investing strategies have delivered. It is well-known that empirical estimates of the magnitude of factor premiums are associated with a wide margin of error—in fact, some claim that one already needs more than 100 years of data to adequately estimate the expected equity risk premium. Because our sample consists of less than 60 independent cross-section years, we use full-sample estimates of the premiums as proxies for their expected values.

One could use our characteristics model to forecast stock returns out of sample, in which case rolling or expanding-window Fama and MacBeth regressions could be used to estimate the factor premiums. However, the issue with the use of rolling regressions is that, following a period during which a certain factor did not do well compared to what is expected over longer periods of time, the estimated premium on this factor could even end up being negative, which would lead us to erroneously conclude that there is a premium of the opposite sign. On the other hand, the expanding-window regressions are very sensitive to the performance of factors during the beginning of the expanding-window period; if certain factors have exhibited returns far different from their long-term expectations, this would significantly affect the estimates for many subsequent periods.

Because our model is not used to forecast factor portfolio returns, but instead to describe the historical performance of the factor portfolios, the fact that we use the ex-post full-sample estimates of the factor premiums is justified. Another way to circumvent the issue of having to estimate the expected factor premiums in the first place is to impose a very conservative assumption that all these factor premiums are equal to one another and set at a certain level (e.g., 15bps a month per unit of standard deviation in the cross-section). In this case, the fit of the model would be somewhat worse than when using the actual premiums estimated from the data; however, it would still lead to qualitatively similar results. In such a model, similar to the model that we use with the full-sample estimates of premiums, all variation in implied returns over time would be driven by the variation in the portfolio characteristics.

To calculate implied returns on each stock in the universe, at each point in time, we multiply a stock's current characteristics by the estimated long-term premiums and add them up. We derive one-month-ahead implied returns for portfolios by taking the weighted average of the implied returns for the individual stocks in the portfolio. Our approach is related to the work of Haugen and Baker (1996) and Lewellen (2015), who showed that expected return forecasts obtained from cross-sectional Fama–MacBeth regressions have strong predictive power for realized stock returns. We apply a similar model, albeit somewhat simpler because we include fewer stock-level characteristics to preserve the parsimony of the model and use static, as opposed to dynamic, estimates of factor premiums. We find that this specification is well suited to the purpose of our study; however, we do acknowledge that one can potentially build a richer model of expected stock returns. The characteristics that we consider are market capitalization (size), book-to-market (value), operating profitability (profitability), investment, momentum, and market beta. Although beta is not a priced characteristic, as we will show later, it is an important control variable in these regressions primarily because of its theoretical underpinnings. The profitability and investment factors are the two quality factors that Fama and French (2015) recently added to their classic three-factor model, turning it into a five-factor model. By also including momentum, for the simple reason that it is too important to ignore, our model uses the characteristics that form the basis of what Fama and French (2018) refer to as the six-factor model.

We obtain security-level data from the Center for Research in Security Prices (CRSP) database. The sample consists of common stocks (share codes 10 and 11) traded on NYSE, AMEX, and NASDAQ exchanges (exchange codes 1, 2, and 3) from June 1963 to December 2017. We exclude stocks with beginning-of-month prices less than \$1 and stocks with market capitalization below the 20th percentile market capitalization of NYSE-listed stocks. These stocks are known as micro-caps; according to Fama and French (2008), they represent around 60% of stocks in the universe but account for less than 3% of the total market capitalization of the market portfolio. Therefore, micro-caps have a big impact on estimation procedures that equally weigh observations, although their importance in the overall market is small. The exclusion of micro-caps from asset pricing tests has been motivated by, among others, Hou, Xue, and Zhang (2017).

The ex-ante estimates of market betas are obtained by running univariate regressions of excess stock returns on the market factor over the 60-month window (minimum 24 months). The market capitalization of a stock is its price times the number of shares outstanding. The balance sheet and income statement information is from Compustat's annual files. Book value is the sum of the book value of stockholders' equity, balance sheet deferred taxes, and investment tax credit (if available), minus the book value of preferred stock. If available, we use the redemption, liquidation, or par value to calculate the book value of preferred stock. Stockholders' equity is obtained either from Moody's industrial manuals or Compustat. If unavailable, we measure stockholders' equity preferably as the sum of the book value of common equity and the par value of preferred stock, or the book value of assets minus total liabilities if the former is not available. For the fiscal years ending in 1993 or later, we do not add deferred taxes to book equity in light of changes in their treatment (Financial Accounting Standards Board 109). The book value of equity is then divided by the market capitalization calculated at the end of the previous calendar year to obtain the book-to-market ratio. Operating profitability is defined as annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity for the last fiscal year end in $t-1$, and investment is the percentage change in firms' total assets from year $t-2$ to $t-1$. Accounting data for a given fiscal year are updated once a year at the end of June of the following calendar year. The 12–2 month total return momentum is the total return from month $t-12$ to $t-2$.

As a proxy for the risk-free rate, we use the one-month US Treasury bill rate obtained from the website of Professor Kenneth French, and we obtain returns on a constant maturity portfolio of 10-year government bonds from CRSP. The market portfolio is the value-weighted portfolio of all stocks in our universe.

We assume that the stock-level characteristics are linearly related to average stock returns and estimate the Fama–MacBeth regressions using ordinary least squares. All characteristics are standardized by subtracting their cross-sectional mean and dividing by their standard deviation. Although this does not have a significant impact on the results of the estimation, it does help with the interpretation of the estimated parameters and, more importantly, gives a meaningful interpretation to the constant. In this setting, the constant represents the excess return on a stock with characteristics equal to the universe average (i.e., z-score of zero).

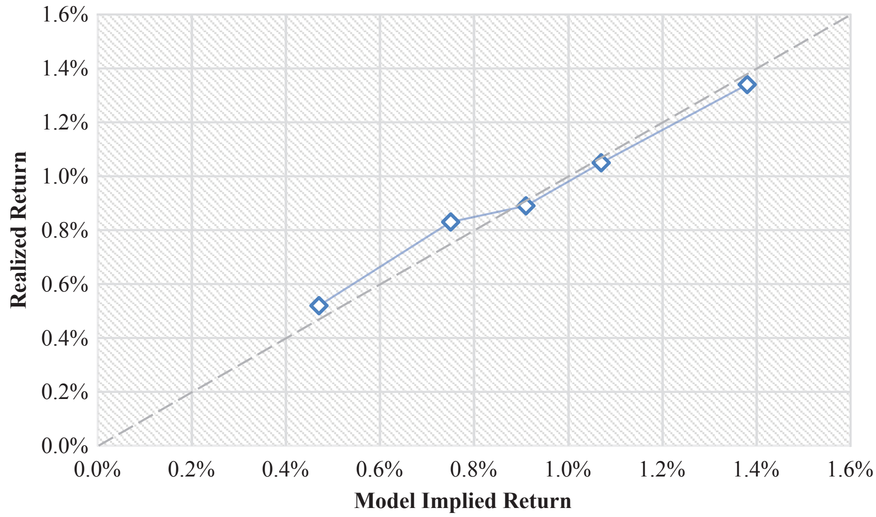
In Exhibit 2, we report the factor premiums estimated over the full sample period. Consistent with the existing literature, we find statistically significant, negative premiums for size and investment (i.e., firms with smaller market capitalization and lower investment, all else constant, have higher average returns) and statistically significant, positive premiums for value, profitability, and momentum. In particular, we find that a stock with a market capitalization that is one standard deviation lower than the capitalization of the average stock in the universe is expected to earn a return that is 10 bps a month higher than that of the average stock, all else constant. Similarly, the reward to a one standard deviation higher book-to-market is 15 bps, the

EXHIBIT 2
Estimated Factor Premiums

	Constant	Beta	ln(Mcap)	ln(BtM)	OP	INV	MOM
Coefficient	0.73**	0.00	–0.10**	0.15**	0.14**	–0.14**	0.26**
t-Stat	3.39	0.04	–2.71	3.01	4.12	–6.39	4.55

*Note: Significant at the ** 1% level and * 5% level.*

EXHIBIT 3
Portfolios Sorted on Implied Return



reward to a one standard deviation higher profitability is 14 bps, the reward to a one standard deviation higher momentum is 26 bps, and the reward to a one standard deviation lower investment is 14 bps, all per month. These results are fully in line with the existing asset pricing literature, which has previously established the existence of these factor premiums.

Contrary to the predictions of the CAPM and the Fama–French asset pricing models, market beta is not priced in the cross-section of stock returns, whereas the constant is positive and significant, instead of zero. This is a manifestation of the well-known low-beta anomaly, which is also fully consistent with the existing literature; see, for instance, Blitz (2014), Clarke et al. (2014), and Blitz and Vidojevic (2017). The positive constant effectively reflects the equity risk premium (i.e., the fact that average stock returns are higher than the risk-free return), which market beta fails to capture. Our estimate of the constant, 0.73% a month, which translates to 8.76% a year, is in line with the realized excess return across all stocks over this sample period.

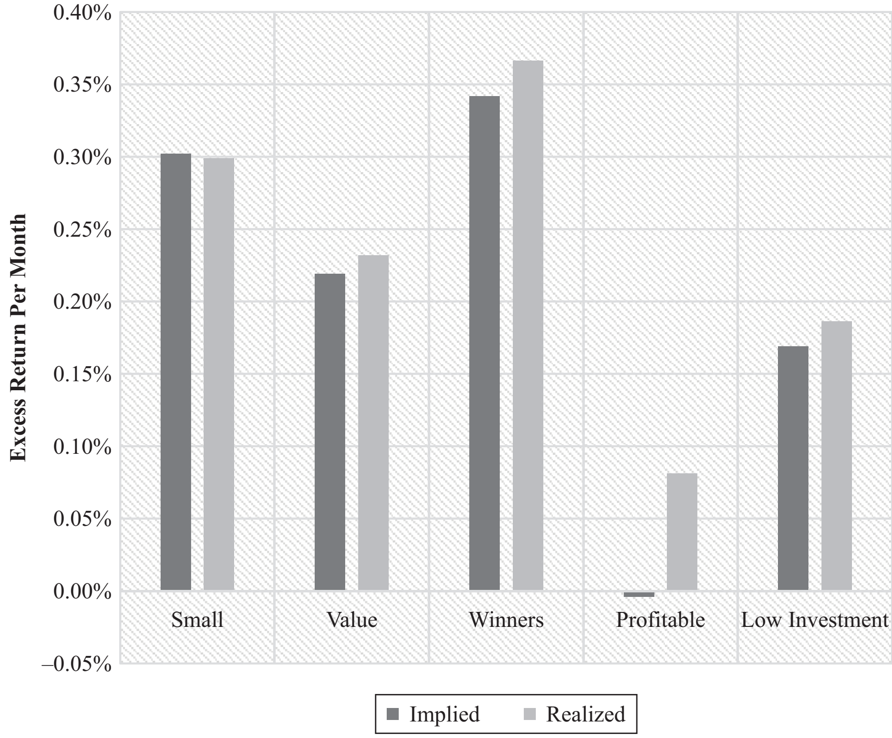
MODEL VALIDATION

Using our multi-factor characteristics-based return model, each month we calculate the implied return of each stock over the following month, based on its current factor characteristics. To validate the predictions of our model, we first sort stocks every month into five portfolios, based on their model-implied returns. We use value weighting (i.e., the weight of a stock in a portfolio is proportional to its market capitalization). We further calculate the implied and realized returns of these five portfolios over the subsequent month. In Exhibit 3, we show that the average implied and realized returns line up closely over the full sample period.

A more formal way to evaluate the fit of our model is to rerun Fama–MacBeth regressions with realized

Exhibit 4

Generic Factor Portfolios: Implied versus Realized Return



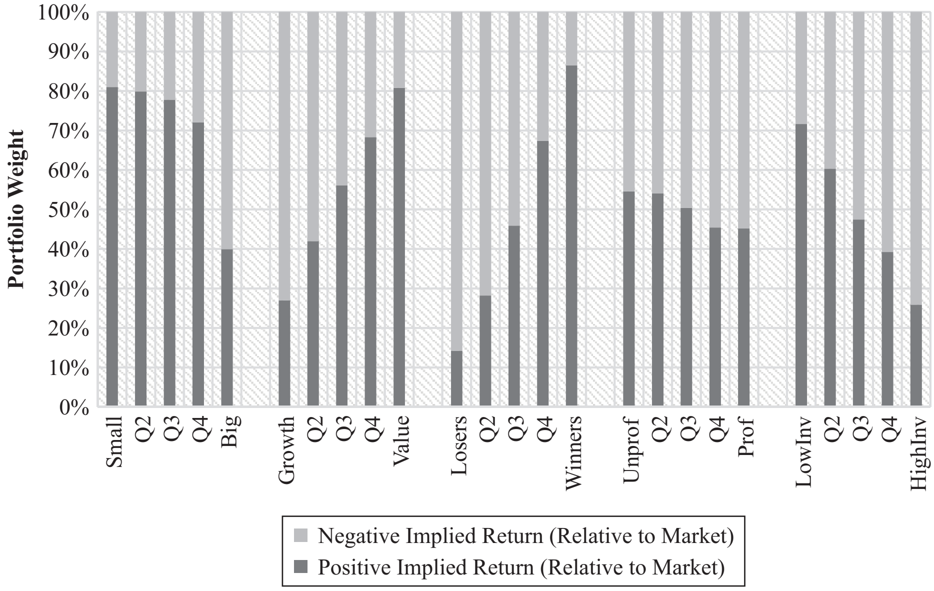
excess stock return on the left-hand side and the implied stock returns on the right-hand side. If the model fit is good, we expect to see a slope coefficient equal to 1. We estimate this model and find a coefficient of 1.05 (t-statistic of 7.11). We further test if this coefficient is statistically different from 1 and find that it is not (t-statistic of 0.34). We conclude that our simple model does a good job matching the first moment of the return distribution.

We further examine the accuracy of our model's predictions by considering various generic factor portfolios. Specifically, we select the most attractive 20% of stocks every month based on their size, value, profitability, investment, or momentum characteristics. Exhibit 4 shows how the average implied returns relate to the average realized returns of these generic factor portfolios. Again we observe that realized and implied returns line up nicely, with the difference between them being statistically indistinguishable from zero in all five cases. The size, value, momentum, and investment portfolios show the best fit. Interestingly, the average implied market-relative return for the profitability portfolio is close to zero; this is because better profitability characteristics tend to go hand in hand with worse size and value characteristics, which, on balance, implies market-relative returns close to zero according to our model. We elaborate on this issue in the next section. Realized returns of the profitability portfolio do show a positive excess return, but much lower than that for the other factors. Although our characteristics-based model seems to have some trouble explaining the performance of the profitability portfolio, the difference between the realized and predicted return is also statistically insignificant for this factor portfolio.

THE CHARACTERISTICS OF GENERIC FACTOR STRATEGIES

Generic factor strategies select stocks with the best scores on one particular factor, regardless of their other factor characteristics. A generic value strategy, for instance, may buy a stock that has strong value characteristics but, at the same time, poor size, momentum, profitability, or investment characteristics. On balance,

EXHIBIT 5
Generic Factor Portfolios: Weight of (Un)Attractive Stocks



taking all this information into account, this may imply a negative, instead of a positive, return for this stock, relative to the market. With our multi-factor characteristics-based return model we are able to assess, at each point in time, how much of a generic factor portfolio is invested in stocks with negative implied returns relative to the market. Panel A of Exhibit 1 shows the average results for the top quintile factor portfolios, and Exhibit 5 presents results for all five quintiles of the five factor strategies. Returning to the example of a generic value portfolio, the graph shows that this strategy invests about 20% in stocks that have a negative predicted market-relative return. This is clearly a non-negligible part of the portfolio. Similar weights are found for the generic small-cap strategy (about 20%), the generic momentum strategy (about 15%), and the generic investment strategy (about 30%). We also see that going from the top to the bottom quintile for each factor, the weight in stocks with implied outperformance declines monotonically.

We next focus on the small, value, winners, profitable, and low investment portfolios (i.e., the top portfolios) and in Exhibit 6 present the result of the factor contribution decomposition implied by our model. For instance, the small-cap portfolio has a full-sample implied outperformance over the market of 30 bps a month (3.6% a year), the vast majority of which is driven by the size (ln(Mcap)) characteristics. However, this portfolio also gets a positive boost from having had a positive exposure to the value factor and a negative contribution from a negative exposure to the profitability factor. The value strategy has also experienced a negative contribution from the profitability factor, on average. The low-investment portfolio has primarily benefited from a high value of the (inverse of the) investment characteristic; however, a substantial part of its outperformance comes from a high value exposure. This is not surprising given that Fama and French (2015) showed that their value (HML) and investment (conservative minus aggressive [CMA]) factors are closely related and that, in fact, the CMA factor subsumes HML in the spanning regressions, making it obsolete.

Exhibit 7 shows the percentage of months in our sample when each of the six characteristics of these factor portfolios were lower (unfavorable) than that of the market (i.e., the targeted characteristic is going against other characteristics). For instance, the small-cap portfolio never has a worse size (ln(Mcap)) characteristic than the (capitalization-weighted) market; in 12% of months it has a lower value (ln(BtM)) characteristic, and in 51% of months it has a worse momentum score. However, the small-cap portfolio systematically goes against the two quality factors—in every single month the profitability

Exhibit 6
Implied Return Decomposition

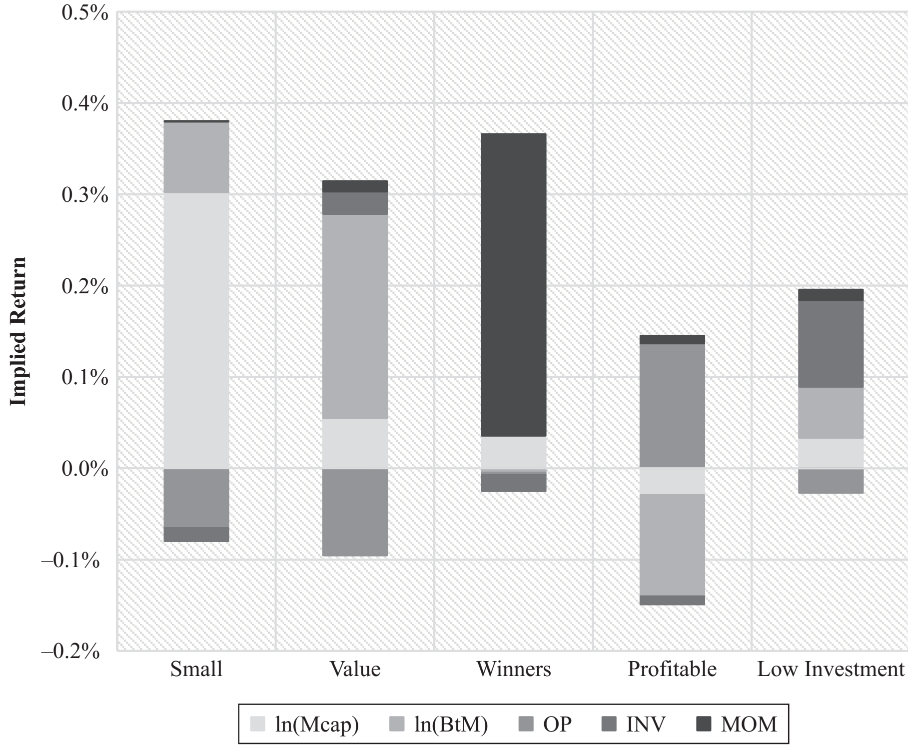


Exhibit 7
Percentage of Months with Unfavorable Characteristics

	ln(Mcap)	ln(BtM)	MOM	OP	INV
Small	0	12	51	100	78
Value	9	0	46	100	14
Winners	29	56	0	52	73
Profitable	99	100	39	0	65
Low Inv.	15	10	45	75	0

characteristic is worse than that of the market, and in the case of investment, this occurs 78% of the time.

Asness et al. (2018) showed that the average small-cap stock is *junky* (i.e., the opposite of high quality), and this detracts from the realized return on the small minus big (SMB) factor. Using ex-post spanning return regressions, they show that if one corrects the SMB factor for its negative exposures to other factors, the small-cap premium emerges as highly economically and statistically significant, although its raw return is quite weak. Our model allows us to look at the performance of the small-cap stocks from a bottom-up characteristics perspective. We confirm that small stocks indeed tend to have poor quality characteristics and conclude that our results are consistent with those of Asness et al. (2018).

Clarke, de Silva, and Thorley (2017) constructed a so-called *pure* size factor that is intended to be orthogonal to other factors and showed that the difference in performance between this pure factor and the generic factor is very small. This finding contradicts that of our article and that of Asness et al. (2018), who also reported substantial improvements in the performance of a size factor that eliminates strong negative exposures to other factors—quality in particular. The reason is that the multivariate regressions used by Clarke, de Silva, and Thorley (2017) ensure that the estimated factor premiums are pure ex-ante but not necessarily ex-post. We conjecture that a pure size premium estimated using their approach will still have substantial exposures to other factors—negative quality exposure in particular.

The profitability factor merits a closer look, especially because it is a relatively newly discovered factor that has gained quite some traction over the recent period. Novy-Marx (2013) showed that stocks with a high ratio of gross profits to assets generate abnormally high returns from the perspective of the, at the time, prominent asset pricing models and, conversely, unprofitable companies generate abnormally low returns. Hou, Xue, and Zhang (2015) used another version of the profitability factor (return on equity, or ROE) in their investment-based Q-factor asset pricing model, and Fama and French (2015) augmented their three-factor model (Fama and French 1993) with a profitability and an investment factors, wherein they use a slightly different proxy for the profitability factor based on firm's operating profitability, which we also use in this article. As a motivation for this new factor, Fama and French used a variant of the dividend discount model (DDM) that implies high expected returns on profitable stocks, holding the level of investment (change in assets) and book-to-market constant. This is an important condition that is often overlooked; namely, the DDM gives the Fama–French factors a conditional interpretation. In this model, a stock that has a higher profitability also has higher expected returns than a peer that trades at the same price multiple (book-to-market) and does the same level of investment. The model does not predict that a stock with higher profitability will outperform unconditionally, for instance, if a stock is more profitable than another but also more expensive.

The generic high-profitability portfolio delivers a fairly small excess return of 0.08% a month in our sample, which is three times smaller than the return of the momentum portfolio. Our characteristics-based model can explain this discrepancy: Profitable stocks tend to score poorly on value and size, which contributed negatively to their model-implied returns. Exhibit 7 shows that there is a systematically negative relationship between the operating profitability (OP) and the book-to-market (ln(BtM)) characteristics—that is, highly profitable stocks also tend to be quite expensive. This is also why Novy-Marx (2013) called (gross) profitability *the other side of value*. The same goes for size because profitable stock also tend to be larger in terms of market capitalization.

For the generic profitability strategy, we find that the part of the portfolio invested in stocks with a negative implied market-relative return amounts to 55%, so from this perspective it is not surprising that this factor strategy shows a much weaker raw performance than the other factor strategies. These results have important implications for the design of strategies that harvest the profitability premium. In particular, generic strategies that target solely the profitability premium are expected to have suboptimal performance unless they take other factors and, crucially, the price paid per unit of fundamental (book) value into account.

ENHANCING FACTOR STRATEGIES

We next examine what happens to performance if, each month, we simply remove stocks that have negative implied market-relative returns from generic factor portfolios. The performance of such enhanced factor strategies is shown in Panel B of Exhibit 1. Compared with the generic factor strategies from which they are derived, the performance improvements are about 20% for the value, momentum, and investment strategies and about 50% for the small-cap strategy. For the profitability strategy, performance more than triples, from 0.08% to 0.29% per month. This large improvement is not surprising because a much bigger adjustment is made to this portfolio than to the other factor portfolios. These results imply that generic factor strategies are suboptimal and that, even when targeting one particular factor premium, investors should not ignore other factor premiums.

Excluding stocks with implied underperformance helps to enhance a single-factor strategy, but the resulting portfolios can still have stocks with negative exposures to other factors that detract from their performance. We next examine how performance changes if, in addition to removing stocks with implied underperformance, we also require stocks to have a non-negative exposure (z-score) to at least one, two, three, or four other, non-targeted premiums. Exhibit 8 shows that realized, full-sample returns of each single-factor strategy tend to increase as we require stocks in the portfolios to have non-negative exposures to more factors. For instance, the raw value strategy has a return of 0.23% a month, which increases to 0.28% per month if we ex-ante exclude stocks with negative implied excess returns. If in addition, at the time of portfolio formation, we require that stocks have non-negative exposures to at least two, three, four, and all five factor premiums, the strategy returns increase to 0.30%, 0.37%, 0.54%, and 0.69%, respectively. As we impose more constraints,

EXHIBIT 8
Further Enhanced Factor Strategies

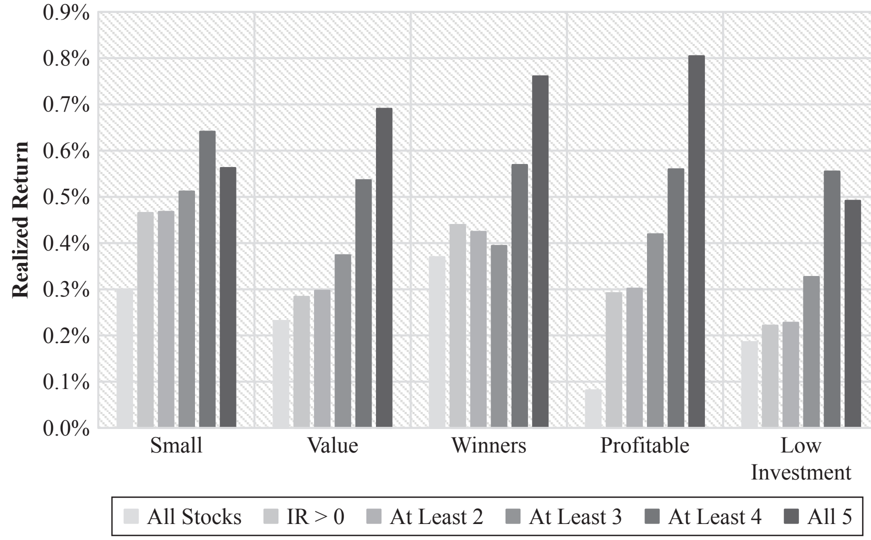
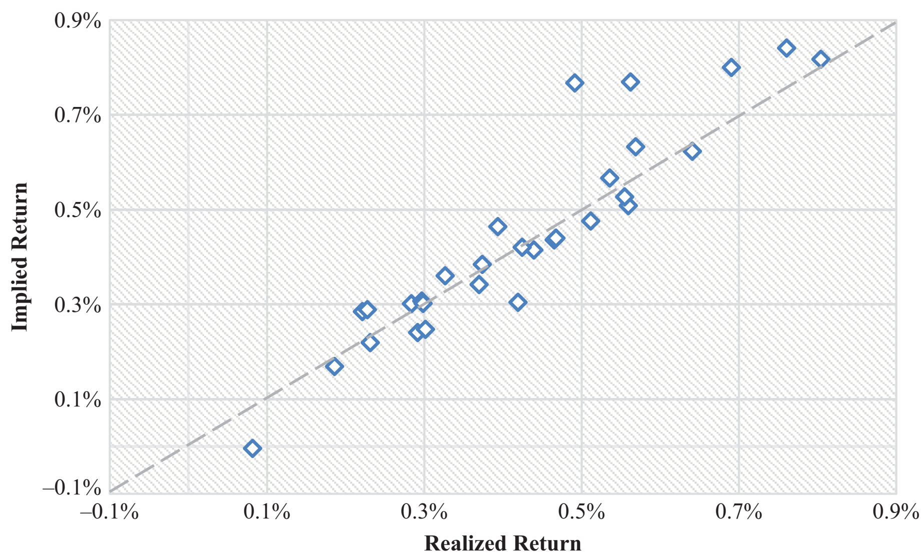


EXHIBIT 9
Realized Versus Implied Returns on Enhanced Single-Factor Strategies

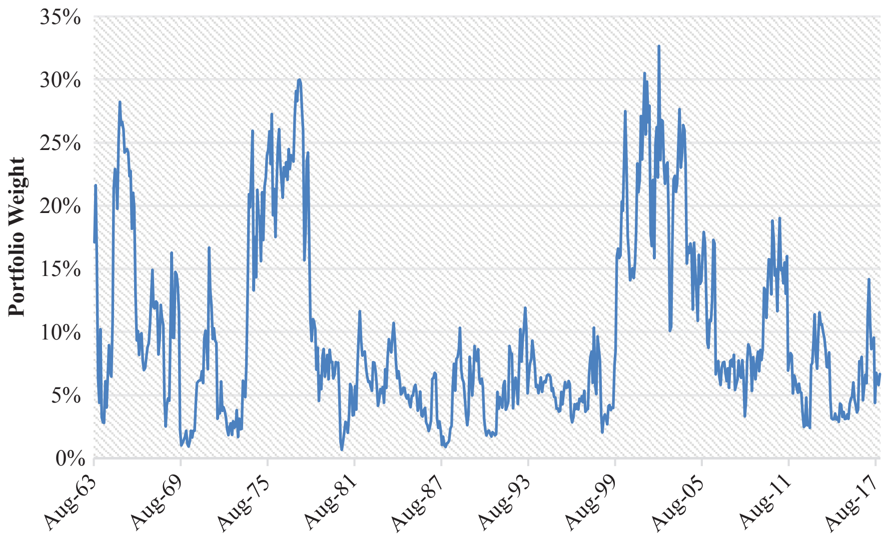


the number of stocks in the portfolio decreases from, on average, 302 with no constraints to 276, 273, 223, 96, and only 13. A similar pattern is observed for other factors, albeit not always monotonic, because very concentrated portfolios can be subject to a fair amount of stock-specific risk that is in general diversified away in broader portfolios.

Exhibit 9 shows that the realized returns on these portfolios align with their model-implied values; that is, the returns of these portfolios can be attributed to their factor characteristics. None of the differences between the realized and implied returns are statistically significant, although we do observe a couple of more notable deviations from the 45-degree line. The largest

E X H I B I T   10

Market Portfolio Weight in Stocks with Implied Return below Bonds



deviations are for the size and investment portfolios that are constrained to have non-negative exposures to all five factors; these are also the most concentrated portfolios, on average containing just over 10 stocks, so naturally the fit of our return model is somewhat lower. Nevertheless, even for these portfolios, the difference between the average implied and realized returns is statistically indistinguishable from zero.

A similar analysis can be conducted for the equity market portfolio, which many investors track passively to earn the equity risk premium. The existence of factors that are priced in the cross-section implies that the market portfolio is not mean–variance efficient; that is, it is suboptimal to hold the market in the presence of factor portfolios. Using our multi-factor characteristics-based return model, we can calculate at each point in time which stocks have an implied excess return that is not merely below average but is even less than the excess return on 10-year US Treasury bonds. Over our July 1963 to December 2017 sample period, the average excess return on a portfolio of 10-year US bonds amounted to 0.18% per month.

Exhibit 10 plots the weight of the market portfolio that is invested in stocks that have a lower model-implied excess return than that of bonds at each point in time. This weight varies over time between less than 5% and more than 25%, with the peaks occurring in the mid-1960s, late 1970s, and early 2000s. On average, 9.8% of the market portfolio consists of stocks with implied returns lower than bonds. Systematically removing such stocks from the equity market portfolio increases its excess return over 10-year US bonds from 0.36% to 0.42% per month, which amounts to a 16% increase in realized returns for investors, in relative terms. The difference in returns is statistically significant at the most conservative levels (t-statistic of 3.22). Thus, our multi-factor characteristics-based return model can also be used to enhance the capitalization-weighted market portfolio.

MIXED-SLEEVE VERSUS INTEGRATED MULTI-FACTOR PORTFOLIOS

We next compare the performance of the integrated and mixed-sleeve multi-factor portfolios. Clarke, de Silva, and Thorley (2016); Bender and Wang (2016); and Fitzgibbons et al. (2017) all argued that integrated portfolios, which are constructed by investing in stocks with a high combined rank on multiple factors, deliver higher absolute and risk-adjusted returns than multi-factor portfolios constructed by mixing single-factor sleeves. On the other hand, Amenc et al. (2017) emphasized some of the shortcomings and risks of integrated portfolios, such as inefficiency, instability, and

Exhibit 11

Performance of Integrated Multi-Factor Portfolios

	Top 50%	Top 33.3%	Top 20%	Top 10%
Realized	0.36**	0.45**	0.53**	0.64**
Implied	0.37**	0.47**	0.57**	0.70**
Difference	–0.01	–0.01	–0.04	–0.06
t-Stat	–0.10	–0.17	–0.47	–0.53

*Note: Significant at the ** 1% level and * 5% level.*

inability to control factor exposures and non-factor risks. Although both streams of literature focus on keeping a certain risk-related metric, such as tracking error or total volatility, constant across these portfolios, none of these studies explicitly aims to control for the differences in factor exposures between the mixed-sleeve and integrated portfolios. This makes the comparisons inappropriate: An integrated portfolio may simply exhibit superior performance characteristics because it comes with better factor characteristics (exposures) than a mixed-sleeve portfolio.

The main challenge here is to make the comparison fair (i.e., to compare the performance of multi-factor portfolios that offer a similar degree of factor exposure). An integrated portfolio, which buys the top 20% of stocks with the highest combined rank on the five factors in our model, contains around 300 names on average over our sample period. On the other hand, a mix of the five generic single-factor portfolios that each contain 300 names provides much lower factor exposures, implying a lower implied return. We therefore look at integrated portfolios with various degrees of concentration and mixes of not only generic single-factor strategies but also of the enhanced single-factor strategies discussed in the previous section.

The integrated portfolios are based on a joint ranking on the five factors in our model, where each characteristic receives an equal weight. Based on this multi-factor ranking, we create value-weighted portfolios consisting of the top 50%, top 33.3%, top 20%, or top 10% of stocks. The performance of these integrated portfolios is reported in Exhibit 11. The implied and realized returns increase steadily as the integrated portfolio becomes more concentrated, and, once again, realized returns are close to the implied returns, with none of the differences being statistically significant.

We construct mixed single-factor portfolios by calculating an equally-weighted average of the five separate single-factor portfolios. We do this not only for the generic single-factor strategies but also for the enhanced single-factor strategies that exclude stocks with negative implied market-relative returns (IR), and the further enhanced single-factor strategies from which stocks with negative exposures to other factors are removed. The performance of these mixed-sleeve portfolios is reported in Exhibit 12. The implied and realized returns increase steadily as one moves from generic to progressively more enhanced single-factor portfolios. Once more, all realized returns are close to the implied returns, with none of the differences being statistically significant.

Several conclusions can be drawn from these results. First, all realized returns are in line with the implied returns, indicating that the factor characteristics of a portfolio are the key determinant for its return. Second, a top 20% integrated multi-factor portfolio achieves significantly higher returns than a mix of various top 20% generic single-factor portfolios, but that is to be expected because it is much more concentrated and therefore provides more factor exposure. Third, an integrated portfolio diluted to such an extent that its factor exposures are similar to those of a mix of generic single-factor portfolios also ends up with a similar return. Or, the other way around, a mix of enhanced single-factor portfolios that offer more pronounced factor exposures is able to match the performance of an integrated approach. Thus, we conclude that there is no evidence of a factor integration premium; that is, there is no evidence that integrating factors gives more return than what may be expected simply from the resulting exposures to the individual underlying factors.

Ghayur, Heaney, and Platt (2018) also emphasized the importance of making an apples-to-apples comparison between mixed and integrated solutions by matching their factor characteristics. However, their approach is different from ours in a two important ways. First, when constructing multi-factor portfolios using the two approaches, they only matched the characteristics of the targeted factors, ignoring the fact that these portfolios may be different with regard to the non-targeted factors. Granted, the authors were aware of this shortcoming, and they did disclose it in footnote 9 ("The portfolios were exposure-matched for the targeted factors, but they may have had differences in other, non-targeted factor exposures. We did not attempt to

EXHIBIT 12

Performance of Mixed-Sleeve Multi-Factor Portfolios

	All Stocks	IR > 0	IR > 0 and 2	IR > 0 and 3	IR > 0 and 4	IR > 0 and 5
Realized	0.23**	0.34**	0.34**	0.40**	0.57**	0.66**
Implied	0.21**	0.34**	0.34**	0.40**	0.57**	0.80**
Difference	0.03	0.00	0.00	0.01	0.00	–0.14
t-Stat	0.60	0.08	0.05	0.10	0.00	–0.94

*Note: Significant at the ** 1% level and * 5% level.*

control for these ancillary exposures."). In our article, we emphasize the importance of thinking about portfolio characteristics from a holistic standpoint because both the targeted, as well as the non-targeted, factor exposures drive expected returns.

Another difference in our approaches is in how the exposures are matched: Ghayur, Heaney, and Platt (2018) matched factor exposures over the full sample period and not at each point in time (footnote 8 in their paper). This means that their approach could lead to substantial deviations in any given time period. In our article, we account for both the targeted and the non-targeted factor characteristics at each point in time because the implied stock and portfolio returns are calculated monthly. Our results are consistent with those of Leippold and Rueegg (2017), who concluded that the two approaches yield similar results if the portfolios are constructed in comparable ways.

EXPLAINING OTHER FACTOR PORTFOLIO RETURN DIFFERENCES

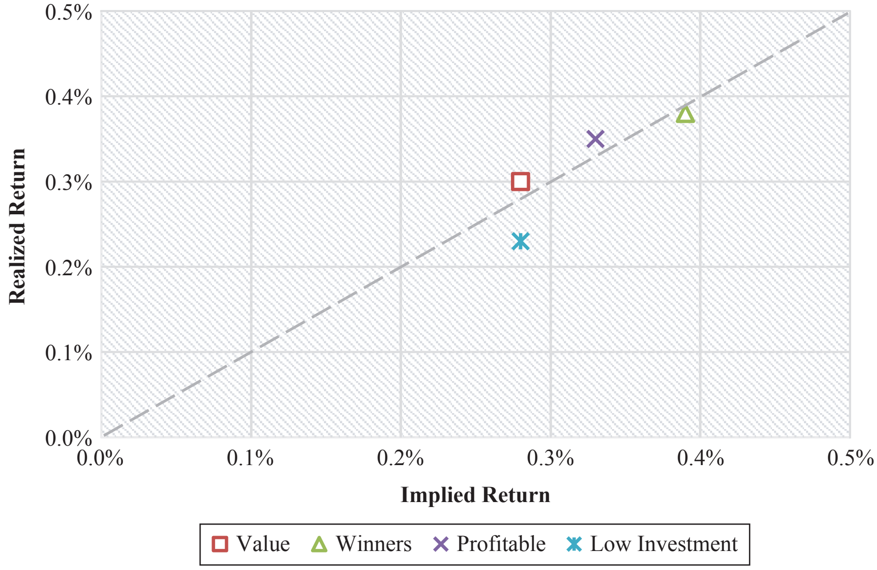
It is well-known that small-cap factor portfolios tend to generate higher returns than large-cap factor portfolios (Fama and French 2008), but it is not entirely clear what drives this spread. One possibility is that the outperformance of the factors in the small-cap universe is driven by a higher exposure of these portfolios to the size premium. Another possibility is that the outperformance comes from a higher exposure to the targeted factor—for instance, if the small-cap value strategy has higher value characteristics than the large-cap value strategy. A third possibility is that the outperformance comes from different exposures to the non-targeted factor premiums (e.g., if the small-cap value strategy has better momentum, investment, or profitability characteristics than the large-cap value strategy). A fourth possibility is that the small-cap and large-cap portfolios have comparable factor characteristics but that the factor premiums are larger in the small-cap segment, in which case a linear expected return model is not appropriate. Lastly, it could be the case that factors that are not accounted for drive these differences.

Another well-known phenomenon is that the equally-weighted factor portfolios tend to generate higher returns than their value-weighted counterparts. Because equally-weighted factor portfolios invest more in smaller stocks, all of the aforementioned explanations may also apply here. In addition, equally-weighted factor portfolios might profit from the periodic rebalancing that is required to maintain equal weights over time (i.e., a potential rebalancing premium).

In this section, we decompose the implied returns of these portfolio pairs using our characteristics-based multi-factor model and pinpoint the reason for the observed return differences between small-cap and large-cap factor portfolios and between equally-weighted and value-weighted factor portfolios.

Exhibit 13 shows the implied and realized performance difference between top quintile value-weighted factor portfolios in the small-cap space and top quintile value-weighted factor portfolios in the large-cap space, where stocks are split into two size groups at the median market capitalization value of our stock universe. Exhibit 14 shows the implied and realized performance difference between equally-weighted and value-weighted factor portfolios. Both graphs show a close match between implied and realized returns, which indicates that our model is able to explain the observed return differences quite well. The tables below the graphs present these numbers in a tabular format and confirm that the difference between realized and implied returns is not statistically significant. In other words, the superior performance of the small-cap factor

Exhibit 13
Small-Cap versus Large-Cap Factor Portfolios



	Value	Winners	Profitable	Low Investment
Realized	0.30**	0.38**	0.35**	0.23
Implied	0.28**	0.39**	0.33**	0.28**
Difference	0.02	0.00	0.02	–0.04
t-Stat	0.15	–0.03	0.13	–0.29

*Note: Significant at the ** 1% level and * 5% level.*

portfolios compared to the large-cap factor portfolios, and the superior performance of equally-weighted factor portfolios compared value-weighted factor portfolios, appears to be fully in line with the implied returns that can be derived from the factor characteristics of these portfolios. This means that there is no need to assume that the same factor exposure is rewarded more in the small-cap space than in the large-cap space, nor is there a need to assume that equally-weighted factor portfolios benefit from some kind of rebalancing premium.

Exhibit 15 shows the breakdown of the implied return contributions across the factor characteristics for the small-big factor portfolios. Clearly, the reason why small-cap factor portfolios outperform their large-cap counterparts can be mostly attributed to a higher exposure to the size factor (i.e., lower market capitalization). All small-cap factor legs tend to have better value characteristics but worse profitability scores. In the case of the past winners, it appears that the small-cap winners portfolio not only benefits from a higher exposure to the small-cap factor but also from a higher exposure to the momentum factor itself; that is, small stocks have higher momentum scores than large stocks, on average.

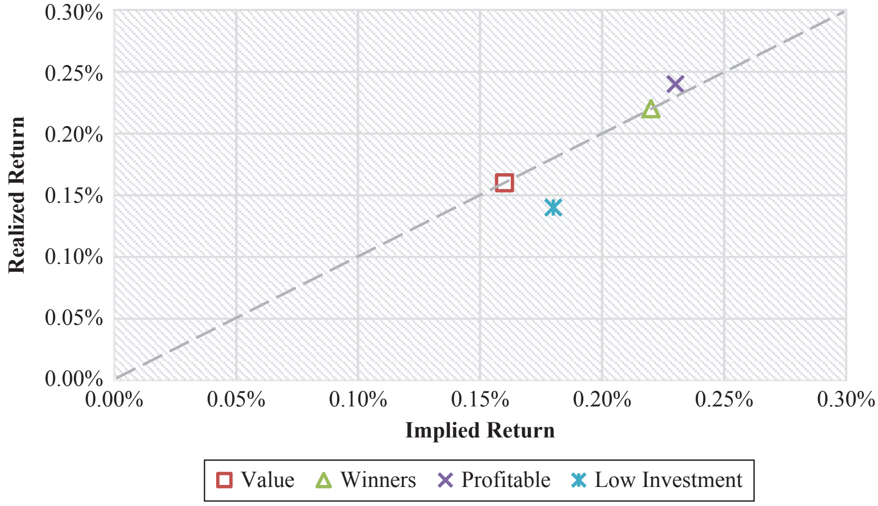
Exhibit 16 shows these breakdowns for the equal versus value-weighted portfolios. We find that the difference in performance across factors constructed using these two weighting schemes can to a large extent be attributed to the small-cap and value tilts that equal-weighting and rebalancing back to the equal weight introduce.

CONCLUSION

Efficient implementation of factor investing requires an understanding of how factors perform independently and how they are related to one another. In this article, we apply a characteristics-based multi-factor

Exhibit 14

Equal versus Value-Weighted Factor Portfolios



	Value	Winners	Profitable	Low Investment
Realized	0.16**	0.22**	0.24**	0.14
Implied	0.16**	0.22**	0.23**	0.18**
Difference	0.00	0.00	0.01	–0.04
t-Stat	–0.03	–0.05	0.14	–0.46

*Note: Significant at the ** 1% level and * 5% level.*

Exhibit 15

Implied Return Decomposition for Factor Portfolios in Small- versus Large-Cap Universes

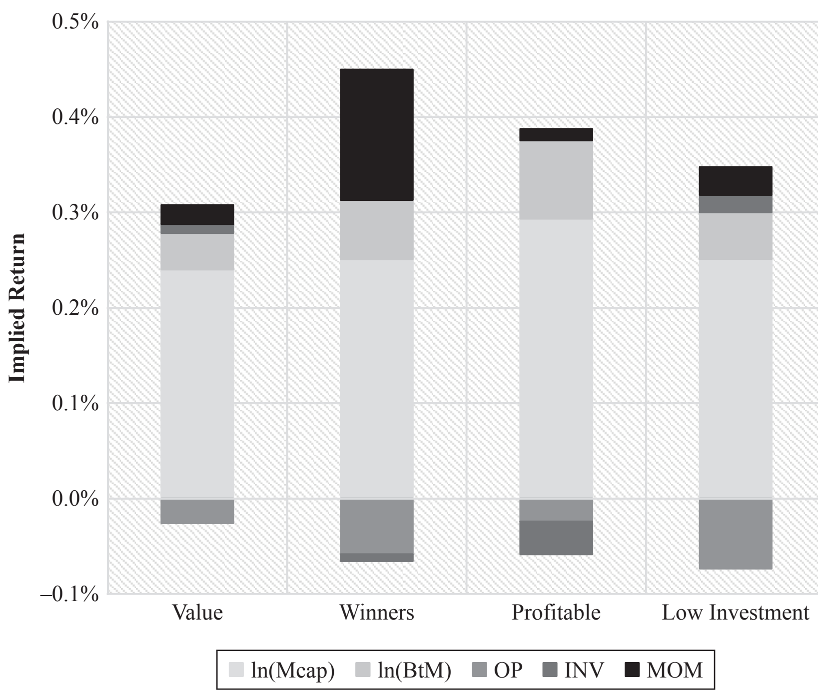
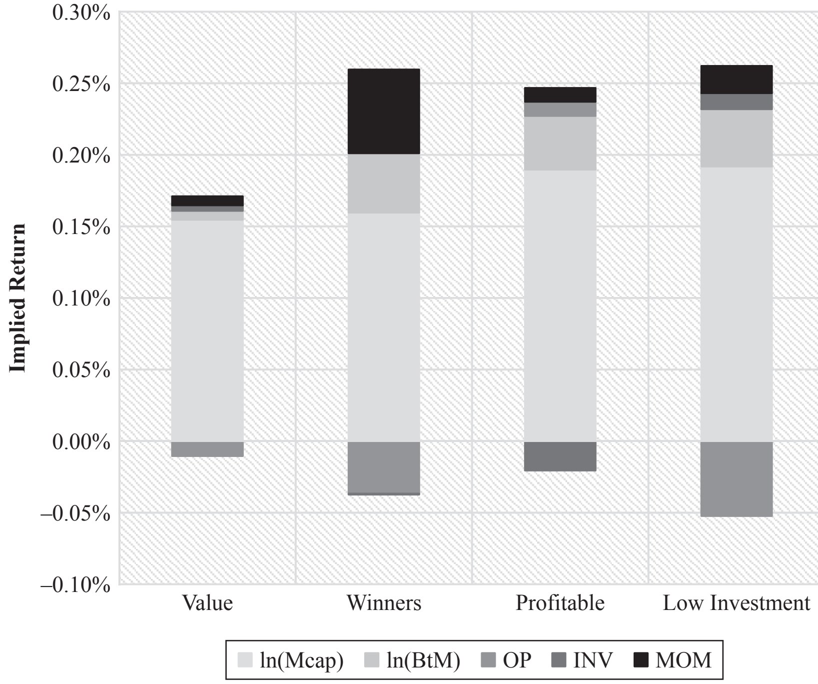


EXHIBIT 16

Implied Return Decomposition for Equal- versus Value-Weighted Factor Portfolios



return model to show that investing in generic single-factor portfolios leads to suboptimal investment results because stocks that score highly on one factor can go strongly against other factors and, on net, have negative market-relative model-implied and realized returns. We show that removing such stocks from the portfolio yields substantially higher returns. We further show that if, in addition to excluding stocks with implied underperformance relative to the market, we also exclude stocks that have negative exposures to one, two, three, or four non-targeted factors, we can enhance these single-factor strategies even more. Our model can also attribute performance differences between mixed-sleeve and integrated multi-factor portfolios to differences in their factor characteristics. As generic single-factor sleeves are an inefficient way to obtain factor exposure, mixing them into a portfolio also results in suboptimal multi-factor portfolios. However, a mixed-sleeve portfolio of enhanced single factors that do not invest in stocks with implied underperformance is able to match the performance of a bottom-up integrated multi-factor portfolio with similar factor exposures. Finally, we show that our model can explain performance differences between factor portfolios in the small-cap and in the large-cap space, as well as equal- and value-weighted portfolios. Therefore, an efficient implementation of factor-based, or any kind of equity strategies requires an understanding of the underlying portfolio factor characteristics.

The discussion in this article is confined to the question of how investors should approach building an expected return model for equity portfolios. In the real world, investors' preferences and constraints cause portfolios to have a significant exposure to factors that are not rewarded with higher returns but certainly do contribute to the risk. An example is industry exposures—an investor should not expect to be compensated for taking industry bets unless they are a result of the rewarded factor characteristics of the underlying industry. Although an effective management of the unrewarded sources of risk is of paramount importance, one has to be cognizant of the fact that, in the long run, expected portfolio returns depend on the priced portfolio characteristics.

ACKNOWLEDGMENTS

The views expressed in this article are solely those of the authors and are not necessarily shared by Robeco or its subsidiaries. We thank Matthias Hanauer for valuable comments.

REFERENCES

Amenc, N., F. Ducoulombier, M. Esakia, F. Goltz, and S. Sivasubramanian. 2017. "Accounting for Cross-Factor Interactions in Multi-Factor Portfolios without Sacrificing Diversification and Risk Control." *The Journal of Portfolio Management* 43 (5): 99–114.

Asness, C., A. Frazzini, R. Israel, T. Moskowitz, and L. Pedersen. 2018. "Size Matters, If You Control Your Junk." *Journal of Financial Economics* 129 (3): 479–509.

Bender, J., and T. Wang. 2016. "Can the Whole Be More Than the Sum of the Parts? Bottom-Up versus Top-Down Multi-Factor Portfolio Construction." *The Journal of Portfolio Management* 42 (5): 39–50.

Blitz, D. 2014. "Agency-Based Asset Pricing and the Beta Anomaly." *European Financial Management* 20: 770–801.

Blitz, D., and M. Vidojevic. 2017. "The Profitability of Low Volatility." *Journal of Empirical Finance* 43: 33–42.

Clarke, R., H. de Silva, and S. Thorley. 2014. "The Not-So-Well-Known Three-And-A-Half-Factor Model." *Financial Analysts Journal* 70: 13–23.

——. 2016. "Fundamentals of Efficient Factor Investing." *Financial Analysts Journal* 72 (6): 9–26.

——. 2017. "Pure Factor Portfolios and Multivariate Regression Analysis." *The Journal of Portfolio Management* 43 (3): 16–31.

Fama, E., and K. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33: 3–56.

——. 2008. "Dissecting Anomalies." *The Journal of Finance* 63 (4): 1653–1678.

——. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22.

——. 2018. "Choosing Factors." *Journal of Financial Economics* 128 (2): 234–252.

Fama, E., and J. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81: 607–636.

Fitzgibbons, S., J. Friedman, L. Pomorski, and L. Serban. 2017. "Long-Only Style Investing: Don't Just Mix, Integrate." *The Journal of Investing* 26 (4): 153–164.

Ghayur, K., R. Heaney, and S. Platt. 2018. "Constructing Long-Only Multi-Factor Strategies: Portfolio Blending vs. Signal Blending." *Financial Analysts Journal* 74 (3): 70–85.

Haugen, R., and N. Baker. 1996. "Commonality in the Determinants of Expected Stock Returns." *Journal of Financial Economics* 41 (3): 401–439.

Hou, K., C. Xue, and L. Zhang. 2015. "Digesting Anomalies: An Investment Approach." *The Review of Financial Studies* 28 (3): 650–705.

——. "Replicating Anomalies." Working paper, 2017.

Leippold, M., and R. Rueegg. 2017. "The Mixed vs the Integrated Approach to Style Investing: Much Ado about Nothing?" *European Financial Management* 24 (5): 1–27.

Lewellen, J. 2015. "The Cross-Section of Expected Stock Returns." *Critical Finance Review* 4: 1–44.

Novy-Marx, R. 2013. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics* 108 (1): 1–28.