Informative Value, Profitability, and Investment Factors*

Manuel Ammann† Tobias Hemauer‡ Simon Straumann§

December 28, 2023

Abstract

Book-to-market, profitability, and investment—the characteristics underlying the Fama-French value, profitability, and investment factors—are imperfect indicators of expected returns. This study narrows down the characteristics' expected return information and uses their informative parts to construct enhanced factors. These informative factors exhibit around 50% higher Sharpe ratios than their standard counterparts. They strongly outperform the standard Fama-French factors regarding the maximum Sharpe ratio criterion and in pricing characteristics-sorted portfolios. Importantly, unlike the standard factors, the informative factors exhibit positive risk prices, making them genuine risk factor candidates. Moreover, our procedure to enhance the factors outperforms other enhancement procedures.

Keywords: Value factor, profitability factor, investment factor, Fama-French five-factor model, cash flow shocks

JEL Classification: G12, G14

*We thank Federico Baldi Lanfranchi (discussant), Lorenzo Bretscher, Jean-Gabriel Cousin (discussant), Sara Easterwood (discussant), Luca Liebi, Martin Nerlinger, João Pinto (discussant), Dennis Umlandt (discussant), and Ajay Venkataraman (discussant) as well as participants of the 2023 Conference of the French Finance Association (AFFI), the 2023 Swiss Finance Institute (SFI) Research Days, the 2023 European Conference of the Financial Management Association (FMAE), the 2023 Annual Meeting of the European Financial Management Association (EFMA), the 2023 World Finance Conference (WFC), the 2023 Annual Meeting of the German Finance Association (DGF), and the 2023 Australasian Finance and Banking Conference (AFBC) for helpful comments and suggestions. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We declare that we have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Send correspondence to Tobias Hemauer, University of St.Gallen, School of Finance, Unterer Graben 21, CH-9000 St. Gallen, Switzerland. Email: tobias.hemauer@unisg.ch.

†University of St.Gallen, School of Finance. Address: Unterer Graben 21, CH-9000 St. Gallen, Switzerland. Email: manuel.ammann@unisg.ch

‡University of St.Gallen, School of Finance. Address: Unterer Graben 21, CH-9000 St. Gallen, Switzerland. Email: tobias.hemauer@unisg.ch

§WHU - Otto Beisheim School of Management, Finance and Accounting Group. Address: Burgplatz 2, D-56179 Vallendar, Germany. Email: simon.straumann@whu.edu

1 Introduction

Characteristics-based factors are ubiquitous in the empirical asset pricing literature, and the value, profitability, and investment factors are among the most prominent factors. Fama and French (2015) motivate the value, profitability, and investment factors by showing that the dividend discount model implies that firms' valuation, profitability, and investment characteristics are related to their expected returns.

Fama and French (2015) combine their value, profitability, and investment factors with market and size factors to obtain a five-factor model.[1] This five-factor model is currently arguably the most established factor model in academia and practice, and Fama and French (2015, 2016) argue that it performs well in explaining the cross-section of stock returns. By contrast, other studies conclude that the model's pricing performance is unsatisfactory (see, e.g., Cooper et al., 2021; Hollstein and Prokopczuk, 2022). Moreover, the factors fail to satisfy the theoretical prediction of the arbitrage pricing theory (APT) that the relation between exposures to the factors and returns should be positive (see, e.g., Daniel et al., 2020; Jegadeesh et al., 2019). This failure indicates that the factors are not mean-variance efficient and questions their interpretation as *risk* factors. Given the factors' unsatisfactory pricing performance and failure to produce an upward-sloping multivariate security market line, they are ill-suited for their typical applications, such as risk-adjusting returns, estimating capital costs, and evaluating investment performance.

The contrast between the Fama-French (2015) factors' theoretical motivation, their shortcomings, and their widespread use spurred researchers to suggest procedures to enhance them (e.g., Daniel et al., 2020; Fama and French, 2020; Ehsani and Linnainmaa, 2022c). We propose a new procedure to enhance the value, profitability, and investment factors. Our enhancement procedure alleviates the factors' deficiencies and also outperforms the existing procedures in this regard. Our approach to enhancing the factors relies on recognizing that their construction methodology neglects a subtle but important aspect: the variation in any of the factors' underlying characteristics—book-to-market, profitability, and investment—reflects not only differences in firms' expected returns but also in other dimensions. We conjecture that factors built from book-to-market, profitability, and investment that are adjusted to be more informative about expected returns generate a better pricing performance than the standard factors.

To narrow down the variation in book-to-market, profitability, and investment that is informative about expected returns, we cancel their variation that is uninformative about expected returns. Besides information about expected returns, all of these characteristics also capture information about expected cash flows. While expected cash flows are relevant for the valuation of stocks based on discounted cash flow models, they are uninformative about expected returns. Therefore, we aim to cancel the information about expected cash flows from book-to-market, profitability, and investment. For this purpose, we obtain a cash flow shock proxy following Hou

[1] The market factor is motivated by the CAPM of Sharpe (1964) and Lintner (1965). The size factor is motivated by extensive empirical evidence that small stocks, on average, outperform big stocks (see, e.g., Banz, 1981; Fama and French, 1992).

and van Dijk (2019) and orthogonalize the characteristics to this cash flow shock proxy. Based on cross-sectional Fama-MacBeth (1973) regressions, we show that our adjusted characteristics in fact capture the original characteristics' information about future returns more precisely. In contrast, the canceled parts exhibit no predictive power for future returns. Hence, our adjusted characteristics are less noisy indicators of expected returns.

We construct new versions of the Fama-French (2015) size, value, profitability, and investment factors based on the adjusted characteristics.[2] We refer to these new versions as informative factors. They substantially improve upon the standard factors, exhibiting higher mean returns, lower volatilities, and higher Sharpe ratios. Thereby, the Sharpe ratios increase by around 50%. Moreover, our informative factors subsume the standard factors, whereas the standard factors cannot price our informative factors. The four-factor model of Hou et al. (2015) with its alternative profitability and investment factors also fails to price our factors. Additionally, our factors are also largely robust to the refinements of Eisfeldt et al. (2022), who adjust book-to-market for intangibles in the value factor's construction, and Ball et al. (2016), who use cash rather than operating profitability in the profitability factor's construction.

Importantly, our informative factors give rise to a significantly higher maximum Sharpe ratio than the standard Fama-French (2015) factors (1.36 vs. 1.04, in annual terms) and thus make the five-factor model more mean-variance efficient. This result also holds in the out-of-sample test proposed by Fama and French (2018). Based on the arguments of Barillas and Shanken (2017), the higher maximum Sharpe ratio implies that our informative factors have higher pricing power for the cross-section of stock returns than the standard factors.

In line with the evidence in the literature, we document that the standard Fama-French (2015) value, profitability, and investment factors exhibit negative risk prices. Thus, they violate the APT's prediction that returns should be positively related to factor exposures. Moreover, the Fama-French (2015) factors generate a significant zero-beta rate, implying that they produce substantial common mispricing in the cross-section of stock returns. By contrast, our informative value, profitability, and investment factors achieve positive risk prices and an insignificant zero-beta rate. Hence, using better indicators of expected returns in the factors' construction leads to negligible common mispricing and an upward-sloping multivariate security market line. Additionally, we show that our informative factors are materially related to the covariance matrix of returns and exhibit reward-to-risk ratios that are consistent with risk-based pricing. Contrary to the Fama-French (2015) factors, our informative factors therefore pass all of the criteria of the factor protocol proposed by Pukthuanthong et al. (2019) and can be considered genuine risk factor candidates.

The recent literature proposes several other procedures to enhance the Fama-French (2015) factors. Most prominent are the hedging procedure of Daniel et al. (2020), the cross-section procedure of Fama and French (2020), and the time-series efficiency procedure of Ehsani and Linnainmaa (2022c). Our procedure differs from these other procedures as we address the

[2] We also construct a new version of the size factor since the construction of the size factor in the Fama-French (2015) model relies on the construction of the value, profitability, and investment factors.

factors' underlying characteristics. We show that our enhancement procedure overall outperforms these other procedures. First, our informative factors achieve higher individual Sharpe ratios than the factors from these procedures. Moreover, our informative factors produce a higher maximum Sharpe ratio than the cross-section and time-series efficient factors. Unlike our informative factors, the value, profitability, and investment factors obtained from the other procedures predominately exhibit negative risk prices and thus fail to satisfy the requirements for valid risk factors. Furthermore, combining our procedure with the other procedures does generally not lead to further improvements—if anything, applying these procedures to our informative factors harms them by leading to negative risk prices and thus turning them into invalid risk factor candidates. This result also implies that our factors hardly suffer from the deficiencies addressed by the other procedures.

We further demonstrate the usefulness of our informative factors by applying them to the pricing of a large set of characteristics-sorted portfolios. Our informative factors improve upon the Fama-French (2015) factors in explaining the individual portfolios' mean returns and their cross-sectional dispersion. Moreover, they also outperform the factors obtained from the hedging, cross-section, and time-series efficiency procedures in this regard. The factors from these alternative procedures even underperform the standard Fama-French (2015) factors, questioning whether they actually enhance the Fama-French (2015) factors' pricing power.

Enhancing the value, profitability, and investment factors is important for several reasons. First, factor models are the workhorse approach in empirical asset pricing, and many of them include value, profitability, and investment factors (e.g., Fama and French, 2015; Hou et al., 2015; Barillas and Shanken, 2018; Barillas et al., 2020). They experience wide acceptance due to their intuitive appeal and empirical robustness. However, inferences drawn from applying factor models—for example, for risk-adjusting returns or evaluating investment performance—whose factors are not mean-variance efficient and do not satisfy theoretical requirements are potentially misleading. On the one hand, the results, such as estimated alphas, are contaminated by the factors' inefficiencies. On the other hand, as pointed out by Chen et al. (2023), accounting for exposures to factors that earn positive mean returns does not make sense if higher factor exposures are not associated with higher returns. Second, the Fama-French (2015) five-factor model is arguably the leading factor model for determining risk-adjusted returns and evaluating investment performance. Given its widespread use, mitigating the deficiencies of the model's factors is critical for academia and practice. Third, given that they are more mean-variance efficient than the Fama-French (2015) factors, our informative factors represent a tighter benchmark for new anomalies or factors to be detected in the cross-section of stock returns. This is particularly relevant amid the issue of an ever-increasing factor zoo as outlined, for example, by Cochrane (2011) and Harvey et al. (2016). Fourth, and more generally, our results suggest new guidelines on how to construct factors. Specifically, factors based on characteristics that are related to expected returns may be improved by narrowing down the characteristics' variation that is actually informative about expected returns. This logic is also applicable to factors other than the Fama-French (2015) factors.

Beyond the aforementioned studies suggesting enhancement procedures for the Fama-French (2015) factors, our study contributes to further streams of literature. In particular, it also adds to the broad literature on the value, profitability, and investment effects in the cross-section of stock returns. Rosenberg et al. (1985) and Fama and French (1992) are among the first to show that book-to-market is positively related to future returns. Novy-Marx (2013) documents that profitability positively predicts returns. Titman et al. (2004) and Cooper et al. (2008) find that investment is negatively related to future returns. Building on these findings, Fama and French (1993), Novy-Marx (2013), and Xing (2008) introduce value, profitability, and investment factors, respectively, to explain the cross-section of stock returns. Since then, a highly active literature proposes refinements to the factors' standard versions used by Fama and French (2015) (e.g., Asness and Frazzini, 2013; Hou et al., 2015; Ball et al., 2016; Eisfeldt et al., 2022; Gonçalves and Leonard, 2023; Jagannathan et al., 2023) and aims to facilitate the understanding of the factors' underlying drivers and subtleties (e.g., Golubov and Konstantinidi, 2019; Ball et al., 2020; Goyal and Wahal, 2023; Cooper et al., 2024).[3] Particularly related to us, various studies attempt to isolate the variation in the characteristics, especially in book-to-market, that is informative about future returns. Fama and French (2006) attempt to narrow down book-to-market's predictive power for future returns by canceling its information about expected profitability. Daniel and Titman (2006) split the change in book-to-market into a tangible return and an intangible return, finding only the latter to be informative about future returns. Gerakos and Linnainmaa (2018) decompose the change in book-to-market into book equity changes and market equity changes, showing that book-to-market's predictive power emanates only from market equity changes. We expand on these studies by proposing a new approach to narrow down the characteristics' information about future returns and by employing their informative parts to construct factors.

Furthermore, our study also relates to the literature examining the pricing of factor exposures. Studies typically find that the theoretically predicted positive relation between factor exposures and returns is empirically very weak or not observable at all for many factor models. This has especially been shown for the CAPM (e.g., Black et al., 1972; Fama and French, 1992; Frazzini and Pedersen, 2014), the Fama-French (1993; 1996) three-factor model (e.g., Daniel and Titman, 1997), and the Fama-French (2015) five-factor model (e.g., Jegadeesh et al., 2019; Daniel et al., 2020; Hollstein and Prokopczuk, 2022). Prominent explanations for the factors' failure to produce a positive relation between factor exposures and expected returns are that they are not true risk factors, that they are imperfect proxies for the mean-variance efficient portfolio, and that measurement errors in the betas lead to biased risk price estimates. Our approach to enhancing factors by narrowing down their underlying characteristics' pricing infor-

[3]We stick to the traditional definitions of book-to-market, profitability, and investment as used by Fama and French (2015) to demonstrate our enhancement procedure relative to the standard factors. Nevertheless, our approach to narrow down the variation in the characteristics that is informative about expected returns is also applicable to alternative definitions of the characteristics. Our procedure can be viewed as complementary to using alternative definitions of the characteristics.

mation addresses the second explanation.[4] In particular, given their higher individual Sharpe ratios as well as the higher maximum Sharpe ratio of their tangency portfolio, our informative factors are better proxies for the mean-variance efficient portfolio than the standard Fama-French (2015) factors. Mitigating the factors' inefficiency while accounting for potential biases in the risk price estimates allows us to evaluate the first explanation, namely whether the factors are true risk factors. We document in fact positive risk prices for our informative factors, suggesting that the Fama-French (2015) factors' mean-variance inefficiency is the reason for their failure to generate an upward-sloping multivariate security market line. The positive risk prices indicate that value, profitability, and investment factors are true risk factors in the sense that they capture risks of hedging concern for investors. Our results also highlight the need to use better indicators of expected returns to obtain efficient factor candidates.

2 Motivation

The most essential prerequisite for a characteristics-based factor to be considered as a potential risk factor candidate is that its underlying characteristic is related to expected returns. However, variation in such characteristics across firms may reflect not only differences in the firms' expected returns but also in other dimensions. Put differently, only part of the characteristics' variation is informative about expected returns, while the remainder is noise. The characteristics are thus imperfect indicators of expected returns. To formalize this idea, consider the following expression for a given characteristic C's cross-sectional correlation with expected returns r:

$$\begin{aligned}\rho_{C,r} &= \frac{cov(C,r)}{\sigma_C \sigma_r} = \frac{cov(C^* + \epsilon_C, r)}{\sigma_C \sigma_r} = \frac{cov(C^*, r) + \overbrace{cov(\epsilon_C, r)}^{=0}}{\sigma_C \sigma_r} \\ &= \frac{cov(C^*, r)\sigma_{C^*}}{\sigma_C \sigma_r \sigma_{C^*}} = \underbrace{\rho_{C^*,r}}_{>\rho_{C,r}} \underbrace{\frac{\sigma_{C^*}}{\sigma_C}}_{<1}\end{aligned} \tag{1}$$

where C^* denotes the characteristic's part that is informative about expected returns, ϵ_C denotes the characteristic's part that is uninformative about expected returns, σ_X denotes the cross-sectional volatility of variable X, $cov(X,Y)$ denotes the cross-sectional covariance between variables X and Y, and $\rho_{X,Y}$ denotes the cross-sectional correlation between variables X and Y. The expression in (1) establishes that the correlation between the characteristic's informative part C^* and expected returns is higher than the correlation between the raw characteristic C and expected returns. Hence, C^* is a better indicator of expected returns than C. Our thesis is that factors based on the characteristics' informative parts achieve a higher pricing power for the cross-section of stock returns than factors based on the raw characteristics.

We apply this idea to the value, profitability, and investment factors of Fama and French (2015). They motivate these factors by establishing that their underlying characteristics—book-

[4]We account for measurement errors in betas when examining the relation between factor exposures and returns using the instrumental variables approach of Jegadeesh et al. (2019).

to-market, profitability, and investment—are related to expected returns. In particular, they derive the following expression by manipulating the dividend discount model:

$$\frac{ME_0}{BE_0} = \sum_{t=1}^{\infty} \frac{\frac{E_0(Y_t)}{BE_0} - \frac{E_0(dBE_t)}{BE_0}}{(1+r)^t} \tag{2}$$

where ME_0 (BE_0) is the current market (book) value, Y_t is total earnings, dBE_t is the change in book equity, and r is the long-term average expected return. All else being equal, a firm's book-to-market ($\frac{BE_0}{ME_0}$) and its expected profitability ($\frac{E_0(Y_t)}{BE_0}$) are positively related to its expected return while the firm's expected investment ($\frac{E_0(dBE_t)}{BE_0}$) is negatively related to its expected return. Thus, book-to-market, expected profitability, and expected investment are indicators of expected returns.[5] Yet, equation (2) also indicates that any of book-to-market, expected profitability, and expected investment is related not only to expected returns but also to the respective other two characteristics. Since the characteristics' variation is not only related to variation in expected returns, they are imperfect indicators of expected returns.

To narrow down the characteristics' variation that is informative about expected returns, we aim to cancel their variation that is uninformative about expected returns. Intuitively, all of these characteristics do not only capture information about expected returns but also about expected cash flows. In the following, we explain separately for each of the three characteristics why they reflect information about expected returns and why they reflect information about expected cash flows. This lays the foundation of our empirical approach to narrowing down the characteristics' expected return information as outlined in Section 3.

First, in the framework of the dividend discount model, firms exhibit low market values, and thus high book-to-market, either because expected dividends are low or because high discount rates are applied to the expected dividends. Since discount rates are in equilibrium equal to expected returns, high book-to-market signals high expected returns only if book-to-market is high because of high discount rates. In contrast, high book-to-market does not signal high expected returns if it is high because of low expected dividends or high book values. Second, in the framework of the net present value rule of investment, firms invest only little either because expected cash flows from their investment projects are low or because the projects' costs of capital are high. Since costs of capital are in equilibrium equal to expected returns, low investment signals high expected returns only if firms invest little because their costs of capital are high.[6] In contrast, low investment does not signal high expected returns if it is low because of low expected cash flows. Finally, and again in the framework of the net present value rule of investment, firms' profitability is high either because they have invested in projects with high expected cash flows and high net present values or because their projects' costs of capital were so high that only projects with high expected cash flows achieved positive net present values.

[5]In the construction of their factors, Fama and French (2015) use current operating profitability and asset growth as proxies for expected profitability and expected investment.

[6]For simplicity, we assume firms' assets to be homogeneous and firms to be all-equity-financed. For a given firm, each investment project thus has the same costs of capital, and the costs of capital equal investors' required returns for holding the firm's stock.

High profitability signals high expected returns only if profitability is high because of high costs of capital. In contrast, high profitability does not signal high expected returns if it is high because of realizing high net present value projects rather than because of high costs of capital.

3 Data and Methodology

3.1 Data Sample

Our sample period runs from July 1963 to December 2019. We obtain stock data from CRSP and firm fundamentals data from Compustat. We supplement the Compustat data with Davis et al.'s (2000) hand-collected book equity data from Kenneth French's website.[7] Our sample includes all stocks that are traded on the NYSE, AMEX, or NASDAQ and have a CRSP share code of 10 or 11. We adjust monthly holding period returns for potential delisting returns. Following Shumway (1997) and Shumway and Warther (1999), we set missing delisting returns for NYSE and AMEX stocks to -30% and for NASDAQ stocks to -55% in case the delisting was performance-related. Finally, we use the one-month T-bill rate retrieved from Kenneth French's website as risk-free rate. The construction of our key variables is described in detail in Appendix A.

3.2 A Proxy for Cash Flow Shocks

The discussion in Section 2 outlines that variation in book-to-market, profitability, and investment stemming from variation in dividend and cash flow expectations is uninformative about expected returns.[8] To narrow down the characteristics' variation that is informative about expected returns, we aim to cancel their variation due to cash flow expectations. As we cannot directly observe cash flow expectations, we revert to a proxy for cash flow shocks. We follow Hou and van Dijk (2019) and use firms' estimated profitability shocks as a proxy for their cash flow shocks. In the first step, we implement Hou and van Dijk's (2019) cross-sectional profitability model that yields estimates for firms' expected profitability. Specifically, we run the following cross-sectional regression at the end of each June from 1964 to 2019:[9]

$$\frac{OI_{i,t}}{AT_{i,t-1}} = b_{0,t} + b_{1,t}\frac{FV_{i,t-1}}{AT_{i,t-1}} + b_{2,t}DD_{i,t-1} + b_{3,t}\frac{D_{i,t-1}}{BE_{i,t-1}} + b_{4,t}\frac{OI_{i,t-1}}{AT_{i,t-2}} + \epsilon_{i,t} \tag{3}$$

where $\frac{OI_{i,t}}{AT_{i,t-1}}$ is firm i's operating income after depreciation scaled by lagged assets, $\frac{FV_{i,t-1}}{AT_{i,t-1}}$ is the ratio of market value to book value of assets (market value of assets is calculated as book value of assets plus market equity (from Compustat) minus book equity (calculated as

[7] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[8] In the long run, firms' dividend and cash flow expectations are closely related, given that dividends are cash flows paid out to investors. For simplicity, we therefore refrain from differentiating between the two and refer to them jointly as cash flow expectations in the following.

[9] Following Hou and van Dijk (2019), we exclude firms with total assets of less than \$10 million and book equity of less than \$5 million for the estimation of the model.

described in Appendix A)), $\frac{D_{i,t-1}}{BE_{i,t-1}}$ is the ratio of dividend payments to book equity, and $DD_{i,t}$ is a dummy variable that equals one if the firm does not pay dividends.

Table 1 presents the average coefficients from the annual regressions. Their signs are identical, and their magnitudes are similar to those reported by Hou and van Dijk (2019). In line with intuition, the coefficients indicate that expected profitability is higher for firms with higher valuations, dividend payments, and past profitability.

Like Hou and van Dijk (2019), we use the annual regression coefficients from the profitability model in (3) to calculate firms' profitability shocks. In particular, we forecast firm i's profitability for year t by multiplying the estimated coefficients from the regression in year $t-1$ with the firm's values for the predictor variables in year $t-1$. The firm's profitability shock in year t, $PS_{i,t}$, is then its realized profitability in year t minus its forecasted profitability; that is:

$$PS_{i,t} = \frac{OI_{i,t}}{AT_{i,t-1}} - E_{t-1}\left(\frac{OI_{i,t}}{AT_{i,t-1}}\right) = \frac{OI_{i,t}}{AT_{i,t-1}} - X_{i,t-1}\hat{b}'_{t-1} \tag{4}$$

where $X_{i,t-1}$ is a vector that contains firm i's values for the predictors as of year $t-1$ and $\hat{b}_{t-1}$ is the vector of coefficients estimated from regression (3) in year $t-1$. $PS_{i,t}$ is our proxy for firm i's cash flow shock across the fiscal year that ended in year $t-1$.

In the literature, other approaches, such as vector autoregressions (see, e.g., Vuolteenaho, 2002) or analyst earnings forecasts (see, e.g., Chen et al., 2013), have been used to infer cash flow shocks. These approaches are not suitable for our setting as they are either not implementable in real-time or suffer from selection, survivorship, and behavioral biases. In particular, a vector autoregression approach is subject to a look-ahead bias and has been shown to exhibit low predictive power and to be prone to model misspecification (see, e.g., Chen and Zhao, 2009). Analyst earnings forecasts cover only a small sample of firms and have been shown to be biased (see, e.g., Lin and McNichols, 1998; McNichols and O'Brien, 1997). By contrast, Hou and van Dijk's (2019) profitability shocks can be estimated for a broad sample of stocks in real-time, do not suffer from survivorship, selection, look-ahead, and behavioral biases, and are based on a model with strong predictive power for expected profitability in the cross-section of firms.

3.3 Identification of Book-to-Market's Pricing Information

In line with the conclusion of Gerakos and Linnainmaa (2018), our discussion in Section 2 suggests that book-to-market's expected return information is embedded in its market equity component rather than book equity component. To identify the variation in book-to-market that is due to variation in market equity, we follow Gerakos and Linnainmaa (2018) and regress book-to-market on lagged market equity changes. Specifically, we run the following cross-sectional regression at the end of each June from 1968 to 2019:

$$BM_{i,t} = b_{0,t} + \underbrace{\sum_{l=1}^{5} b_{l,t} dME_{i,t-l+1}}_{BM^{me}_{i,t}} + \epsilon_{i,t} \tag{5}$$

where $BM_{i,t}$ is log book-to-market and $dME_{i,t}$ is the log change in market equity. The choice of using five lagged market equity changes follows Gerakos and Linnainmaa (2018). It balances two aspects. On the one hand, using more lags identifies the market equity-driven variation in book-to-market more accurately. On the other hand, using more lags decreases the number of stocks for which all of the required data is available.

Panel A of Table 2 shows that the average coefficients on all lagged market equity changes are significantly negative. As indicated by the average adjusted R^2 of 48.8%, the past five years' market equity changes explain half of book-to-market's cross-sectional variation.

In Section 2, we further argue that only the market equity-driven variation in book-to-market that is related to discount rates is informative about expected returns. The variation in book-to-market's market equity-driven part that is related to dividends should not be informative about expected returns. To cancel this variation, we orthogonalize book-to-market's market equity-driven part with respect to our estimated profitability shocks. Specifically, we run the following cross-sectional regression at the end of each June from 1968 to 2019:

$$\widehat{BM}^{me}_{i,t} = b_{0,t} + \sum_{l=1}^{5} b_{l,t} PS_{i,t-l+1} + \underbrace{\epsilon_{i,t}}_{BM^*_{i,t}} \tag{6}$$

where $\widehat{BM}^{me}_{i,t}$ is book-to-market's market-equity driven part, $PS_{i,t}$ is the profitability shock, and $BM^*_{i,t}$ is our adjusted book-to-market measure intended to be a better indicator of expected returns than raw book-to-market. Using five lags is again motivated by balancing data availability with the accuracy of identifying the relevant variation.[10]

Panel B of Table 2 documents that the coefficients on all lagged profitability shocks are significantly negative. In line with intuition, this result suggests that cash flow shocks negatively affect firm valuations. As indicated by the average adjusted R^2 of 38.8%, the profitability shocks across the past five years explain a substantial fraction of the cross-sectional variation of book-to-market's market equity-driven part.

3.4 Identification of Investment's Pricing Information

Section 2 outlines that the variation in investment that is related to cash flows is not informative about expected returns. To cancel this variation, we orthogonalize investment to our estimated profitability shocks by running the following cross-sectional regression at the end of each June from 1968 to 2019:

$$INV_{i,t} = b_{0,t} + \sum_{l=1}^{5} b_{l,t} PS_{i,t-l+1} + \underbrace{\epsilon_{i,t}}_{INV^*_{i,t}} \tag{7}$$

[10] In unreported results, we confirm that our findings are qualitatively robust to using a different number of lags.

where $INV_{i,t}$ is log investment, $PS_{i,t}$ is the profitability shock, and $INV^*_{i,t}$ is our adjusted investment measure intended to be a better indicator of expected returns than raw investment.

The results in Panel C of Table 2 document that the coefficients on all lagged profitability shocks are significantly positive. Consistent with intuition, firms thus increase their investment upon positive cash flow shocks. The profitability shocks across the past five years explain 21.3% of the cross-sectional variation in investment.

3.5 Identification of Profitability's Pricing Information

In Section 2, we argue that the variation in profitability that is related to cash flows is not informative about expected returns. However, given that profitability captures expected return information, our estimated profitability shocks may capture part of the expected return information inherent in profitability. On their own, they are therefore unsuited to cancel the variation in profitability that is related to cash flows.[11] Specifically, the estimated profitability shocks cannot differentiate whether increases in profitability are due to cash flow shocks that make investment projects more profitable or due to shocks to costs of capital that lead firms to realize only highly profitable projects. However, the former case is associated with increasing net present values, whereas the latter case is associated with decreasing net present values. Increasing net present values should also be associated with increases in firms' investment and market values because firms have more positive net present value projects available to invest in and generate more value for investors. Therefore, the part of profitability shocks that genuinely reflects cash flow information rather than expected return information should be associated with increases in investment and market values. Building on this reasoning, we use contemporaneous changes in investment and market values to instrument for the part of profitability shocks that is informative about cash flows. Specifically, we run the following cross-sectional regression at the end of each June from 1964 to 2019:

$$PS_{i,t} = b_{0,t} + \underbrace{b_{1,t} dINV_{i,t} + b_{2,t} dME_{i,t}}_{PS\text{-}Fit_{i,t}} + \epsilon_{i,t} \tag{8}$$

where $PS_{i,t}$ is the profitability shock, $dINV_{i,t}$ is the log change in investment, and $dME_{i,t}$ is the log change in market equity.

The results in Panel D of Table 2 reveal that the coefficients on the investment and market equity changes are positive and highly significant. That is, profitability shocks are, as expected, positively associated with increases in investment and market values.

To eventually cancel its variation that is related to cash flows, we orthogonalize profitability to the fitted values from the regression in (8) by running the following cross-sectional regression

[11] Note that this is not a problem for canceling book-to-market's and investment's cash flow information using the estimated profitability shocks because expected return information and expected cash flow information are reflected with opposite signs in book-to-market and investment—expected returns are positively (negatively) related to book-to-market (investment) while cash flows are negatively (positively) related to book-to-market (investment)—but with the same sign in profitability shocks—expected returns and cash flows are both positively related to profitability and thus profitability shocks.

at the end of each June from 1968 to 2019:

$$OP_{i,t} = b_{0,t} + \sum_{l=1}^{5} b_{l,t} PS\text{-}Fit_{i,t-l+1} + \underbrace{\epsilon_{i,t}}_{OP^*_{i,t}} \tag{9}$$

where $OP_{i,t}$ is operating profitability, $PS\text{-}Fit_{i,t}$ is the fitted profitability shock, and $OP^*_{i,t}$ is our adjusted operating profitability measure intended to be a better indicator of expected returns than raw operating profitability.

Panel E of Table 2 shows that the coefficients on all but the five-year lagged fitted profitability shock are significantly positive. As indicated by the average adjusted R^2 of 40.4%, the fitted profitability shocks across the past five years explain a substantial fraction of the cross-sectional variation in operating profitability.

3.6 Characteristics' Predictive Information

Before constructing and evaluating the factors based on our adjusted characteristics, we verify that they in fact narrow down the raw characteristics' pricing information. For this purpose, we conduct monthly cross-sectional Fama-MacBeth (1973) regressions that regress one-month ahead excess returns on the raw and adjusted characteristics. All characteristics are measured at the end of the most recent June (except size, which is measured at the end of each month). We winsorize all independent variables at the 1% and 99% levels and use weighted least squares with stocks' market capitalizations as weights.

Table 3 presents the average monthly coefficient estimates across the period from June 1968 to November 2019. t-statistics are based on Newey-West (1987) standard errors with one lag. In line with the size and investment effects, columns (1) and (4) reveal that raw size and raw investment significantly negatively predict returns on an individual basis. Moreover, in line with the value and profitability effects, columns (2) and (3) indicate that raw book-to-market and raw operating profitability are positively, albeit insignificantly, associated with future returns. Column (5) shows that size and operating profitability significantly predict returns when all four characteristics are simultaneously included. Book-to-market and investment are insignificantly but with the correct signs (i.e., positively, respectively, negatively) associated with future returns.

To examine whether our adjusted characteristics capture the raw characteristics' informative variation, we split the raw characteristics into the adjusted characteristics and the canceled parts. The canceled parts are the differences between the raw and adjusted characteristics. Columns (6) to (8) show that the canceled parts individually have no significant predictive power and that most of their coefficients even have the wrong sign. When included jointly, the canceled parts of operating profitability and investment exhibit significant predictive power, but the coefficients again have the wrong signs. Thus, the canceled parts do not contain any of the raw characteristics predictive information for future returns.

By contrast, columns (10) to (12) reveal that all of the adjusted characteristics individually exhibit significant or only marginally insignificant predictive power for future returns and that the coefficients have the correct signs. When included jointly (column (13)), only the adjusted book-to-market exhibits significant predictive power, but the coefficients of the adjusted operating profitability and investment have the correct signs and are just marginally insignificant.

Columns (14) to (16) further show that the adjusted characteristics' predictive power remains significant when the respective characteristics' canceled parts are added. In contrast, the canceled parts still exhibit small and insignificant coefficients that usually have the wrong sign. When all adjusted characteristics and canceled parts are jointly included (column (17)), the adjusted characteristics' coefficients have the correct signs, albeit mostly insignificant, whereas the canceled parts' coefficients have the wrong signs.

Altogether, the results in Table 3 clearly show that our adjusted characteristics in fact capture and narrow down the raw characteristics' pricing information. The canceled parts contain, if anything, predictive information that goes in the opposite direction of the raw characteristics' predictive information.

4 Informative versus Standard Factors

4.1 Factor Construction

Having confirmed that our adjusted characteristics capture the raw characteristics' pricing information, we construct new versions of the Fama-French (2015) factors—to which we refer as informative factors—based on our adjusted characteristics.[12] For this purpose, we follow Fama and French's (2015) methodology. First, the market factor (MP) is the value-weighted return on the market portfolio in excess of the one-month T-bill rate. The market portfolio includes each month all stocks that are listed on the NYSE, AMEX, or NASDAQ, have a CRSP share code of 10 or 11, and have good market equity data at the beginning of the month.

To construct our informative value factor, we sort stocks at the end of each June into two groups according to their size at the end of June and into three groups according to their adjusted book-to-market. The sorting breakpoints are the median market equity and the 30th and 70th percentiles of the adjusted book-to-market of all NYSE stocks. Taking the intersections of the two size and three book-to-market groups yields six portfolios. The return on our informative value factor (HML^*) is the average of the value-weighted returns on the two high book-to-market portfolios minus the average of the value-weighted returns on the two low book-to-market portfolios.

The informative profitability and investment factors are constructed in the same way as the informative value factor, except that the second sort is with respect to our adjusted operating profitability and investment, respectively. The return on our informative profitability factor (RMW^*) is the average of the value-weighted returns on the two high profitability portfolios minus the average of the value-weighted returns on the two low profitability portfolios. The

[12] For comparison, we also reconstruct the standard Fama-French (2015) factors based on the raw characteristics.

return on our informative investment factor (CMA*) is the average of the value-weighted returns on the two low investment portfolios minus the average of the value-weighted returns on the two high investment portfolios. Finally, the return on our informative size factor (SMB*) is the average of the returns on the nine small portfolios resulting from the three bivariate sorts minus the average of the returns on the nine big portfolios.

4.2 Summary Statistics

Panel A of Table 4 presents summary statistics on the reconstructed standard Fama-French (2015) factors. All factors, except the size factor, have significant mean returns across the period from July 1968 to December 2019. They range between 0.53% per month for the market factor and 0.15% per month for the size factor. The investment factor exhibits the highest monthly Sharpe ratio (0.14), while the size factor exhibits the lowest Sharpe ratio (0.05).

Panel B presents summary statistics on our informative factors. All informative factors, even the size factor, have significant mean returns. Our informative factors' mean returns are unanimously higher than those of their standard counterparts, but the increases are, except for the size factor, not statistically significant. Moreover, our informative factors' volatilities are lower than those of their standard counterparts. Thus, narrowing down the characteristics' variation that is informative about expected returns leads to larger spreads in expected returns and identifies differences in expected returns with higher certainty.

The higher mean returns and lower volatilities result in substantial increases of around 50% in our informative factors' Sharpe ratios relative to their standard counterparts. These economically significant increases are also statistically significant or only marginally insignificant.[13] This result indicates that our informative factors capture more pricing information than the standard factors. Thereby, the higher Sharpe ratios are predominantly due to the informative factors' lower volatilities, which is well in line with our theoretical motivation discussed in Section 2. In particular, we argue that factors based on the characteristics' informative parts capture more pricing information. However, this conjecture does not suggest whether the increased pricing information emanates from higher mean returns or lower volatilities. For the factors' pricing power, it is in fact not relevant whether their increased Sharpe ratios stem from higher mean returns or lower volatilities.

Our informative size factor exhibits a particularly large improvement. This observation suggests that controlling for our adjusted characteristics rather than the raw characteristics—and thus for better indicators of expected returns—in the size factor's construction is beneficial for isolating the pricing information captured by size.

Table 4 also presents correlations between the factors. The correlations between our informative value, profitability, and investment factors are lower than those between their standard counterparts. Thus, our informative factors capture more independent pricing information than their standard counterparts.

[13] The standard errors of the Sharpe ratio differences are calculated based on the delta method.

4.3 Pricing Factors

Barillas and Shanken (2017) argue that a factor captures incremental pricing information relative to a set of factors if its alpha relative to these factors is significant. To verify that our informative factors capture incremental pricing information relative to the standard factors, we regress them on each other. Table 5 presents the results. In Panel A, we regress the informative factors on the standard five-factor model, and vice versa in Panel B. The results in Panel A show that each of our informative factors captures incremental pricing information relative to the standard factors. All informative factors exhibit significantly positive alphas, ranging between 0.06% and 0.21%. Thus, the standard factors fail to capture the full pricing information of any of our informative factors.

Conversely, Panel B shows that the standard value, profitability, and investment factors do not exhibit significant alphas relative to our informative factors. Hence, they do not capture significant incremental pricing information beyond our informative factors. Put differently, our informative factors capture the pricing information of the standard value, profitability, and investment factors, but not vice versa. The standard size factor exhibits a significantly negative alpha, meaning it captures incremental pricing information relative to our informative factors. However, the negative alpha implies that this pricing information goes in the opposite direction of the size effect.

We also examine whether our informative factors capture incremental pricing information relative to the four-factor model of Hou et al. (2015), which is the most common alternative to the Fama-French (2015) five-factor model in the current asset pricing literature.[14] Like the Fama-French (2015) model, this model includes market, size, profitability, and investment factors but constructs them differently. The different construction methodology results in higher mean returns and Sharpe ratios for their profitability and investment factors compared to Fama and French's (2015) profitability and investment factors. Nevertheless, despite its improved factors, Panel C of Table 5 shows that Hou et al.'s (2015) four-factor model also fails to price our informative factors, leaving sizable and highly significant alphas for all of them. Thus, our informative factors capture incremental pricing information relative to the factors of Hou et al. (2015).

The literature has proposed various other refinements to the Fama-French (2015) factors. The most prominent suggestions are adjusting book-to-market for intangibles when constructing the value factor (e.g., Eisfeldt et al., 2022) and using cash profitability rather than operating profitability when constructing the profitability factor (e.g., Ball et al., 2016; Fama and French, 2018). To verify whether our enhancements are robust to these refinements, we reconstruct Eisfeldt et al.'s (2022) intangible value factor and Fama and French's (2018) cash profitability factor.[15] Panel D of Table 5 presents the results from pricing our informative factors with a

[14] We obtain data on the factors from the authors' global-q website: `http://global-q.org/`.

[15] For the reconstruction of Eisfeldt et al.'s (2022) intangible value factor, we employ the intangibles-adjusted book-to-market data provided by Edward Kim on Github: `https://github.com/edwardtkim/intangiblevalue`. The data is available from 1975 onwards.

five-factor model that replaces the standard value and profitability factors in the Fama-French (2015) model with the intangible value and cash profitability factors. The informative size and value factors' alphas are halved compared to Panel A and turn insignificant. In contrast, our informative profitability and investment factors' alphas remain significant. Thus, while our enhancements in part overlap with these refinements, they are to a large degree incremental.

4.4 Maximum Sharpe Ratios

Table 5 shows that each of our informative factors captures incremental pricing information relative to the standard Fama-French (2015) model. A natural question is how much our informative factors improve the five-factor model's pricing performance if they replace the standard factors. Barillas and Shanken (2017) argue that the pricing performance of models should be compared based on their maximum Sharpe ratios they are able to attain. Barillas et al. (2020) propose a test to evaluate whether the maximum (squared) Sharpe ratios of two models are significantly different. Moreover, Fama and French (2018) suggest to compare models' maximum Sharpe ratios based on a bootstrap simulation approach.[16] This approach splits the sample period of T months into $T/2$ adjacent pairs of months. Each simulation run randomly draws $T/2$ pairs with replacement. From each pair, one month is allocated to an in-sample period and the other month to an out-of-sample period. The in-sample months are used to calculate the models' in-sample maximum Sharpe ratios and the factors' weights in the in-sample tangency portfolios. These in-sample weights are then used to calculate the models' maximum Sharpe ratios in the out-of-sample months. While the in-sample Sharpe ratios are upward-biased estimates of the models' true Sharpe ratios, Fama and French (2018) argue that the out-of-sample Sharpe ratios are unbiased estimates of the true Sharpe ratios.

We implement the test of Barillas et al. (2020) and the bootstrap approach of Fama and French (2018) to compare the maximum squared Sharpe ratios of the standard Fama-French (2015) five-factor model and a five-factor model using our informative factors. Panel A of Table 6 presents the results. First, our informative five-factor model exhibits a maximum squared Sharpe ratio of 0.155 across our sample period, which is much higher than the standard five-factor model's maximum squared Sharpe ratio of 0.093. In terms of annual Sharpe ratios, this represents an increase of around 30% (from 1.04 to 1.36). Barillas et al.'s (2020) test and Fama and French's (2018) bootstrap approach both indicate that the difference in the models' maximum squared Sharpe ratios is highly significant. Specifically, the test statistic of 2.63 from Barillas et al.'s (2020) test indicates that the difference is significant at the 1% level. Moreover, our informative model's out-of-sample Sharpe ratio is in 96.6% of 100,000 bootstrap simulation runs higher than the standard model's. Our informative model also outperforms the standard model in the full-sample and in-sample simulations, generating higher Sharpe ratios in 99.6% respectively 96.3% of simulation runs.[17] Given Barillas and Shanken's (2017)

[16]This bootstrap approach to compare models' maximum Sharpe ratios has also been adopted by other studies (e.g., Ehsani and Linnainmaa, 2022a; Detzel et al., 2023).

[17]The full-sample simulations randomly draw T months with replacement.

conclusion that models' pricing performance can be compared based on their maximum Sharpe ratios, these results indicate that our informative factors substantially improve the standard five-factor model's pricing performance.

To examine the relative contributions of our informative factors to this improvement, we explore factor models that swap the informative factors one at a time for the corresponding standard factors in the Fama-French (2015) model.[18] The results reveal that the informative profitability factor contributes the most and the informative value factor contributes the least to the improvement of our informative model relative to the Fama-French (2015) model. Nevertheless, we can also observe that the contributions are almost additive, meaning that the factors' enhancements are largely complementary.

Panel B presents the factors' weights in the models' tangency portfolios (i.e., in the portfolios attaining the maximum Sharpe ratios). Our informative value and profitability factors' weights are higher than the weights of their standard counterparts, meaning their contributions to the pricing performance increase. Notably, the value factor's contribution turns from negative to positive. Nevertheless, the ordering of the factors' contributions to the pricing performance remains the same. The investment factor has the highest weight while the value factor has the lowest weight.

4.5 Spanning Regressions

Finally, we examine the individual factors' incremental pricing power relative to each other. Given Barillas and Shanken's (2017) conclusion that a factor captures incremental pricing information if its alpha relative to the other factors is significant, we conduct spanning regressions that regress each factor on the respective other factors.

Panel A of Table 7 presents the results for the standard Fama-French (2015) factors. In line with the finding of Fama and French (2015), the standard value factor exhibits an insignificant alpha of -0.04%, and its positive mean return is primarily captured by the investment factor. Thus, the standard value factor does not capture incremental pricing power relative to the other factors and is therefore redundant. This result is consistent with the value factor's small and negative weight in the standard model's tangency portfolio (see Panel B of Table 6). Both observations suggest that the standard value factor hardly captures incremental pricing information and could be dropped. Unlike the value factor, the remaining standard factors exhibit significantly positive alphas. Thus, they possess incremental pricing power and significantly contribute to the Fama-French (2015) model's pricing performance.

Panel B presents the results for our informative factors. Like the standard value factor, our informative value factor exhibits an insignificant alpha and therefore possesses no significant incremental pricing power. The value factor is again primarily subsumed by the investment

[18] The size factor in each of these models is constructed based on the respective model's value, profitability, and investment factors. For example, in the model containing the informative value factor and the standard profitability and investment factors, the size factor is based on the portfolios resulting from the three double-sorts on market equity and any of adjusted book-to-market, standard operating profitability, and standard investment.

factor. Nevertheless, its alpha of 0.09% is non-negligible and much higher than the standard value factor's alpha. Hence, our informative value factor captures more incremental, albeit still insignificant, pricing information than its standard counterpart. This conclusion is supported by the considerable increase in the value factor's tangency weight from -3.0% to 5.9% in the informative model relative to the standard model (see Panel B of Table 6). In contrast to the value factor, the remaining informative factors exhibit highly significant alphas. Thus, they all possess significant incremental pricing power.

In sum, this section's results give rise to several conclusions. First, our informative factors substantially improve upon the standard factors on an individual basis, implying that our approach to narrow down the characteristics' expected return information is successful. The improvements emanate from higher mean returns as well as lower volatilities. Second, our informative factors capture nearly the full pricing information of the standard factors. Given that the opposite does not hold, our informative factors capture more pricing information than their standard counterparts. Third, our informative factors significantly enhance the five-factor model's pricing performance. Last, all of our informative factors, except the value factor, capture significant incremental pricing information. Our informative value factor's non-negligible tangency weight and alpha nevertheless indicate that it captures more incremental pricing information than its standard counterpart.

5 Factor Risk Prices

A well-specified factor model should comprise only genuine risk factors. A factor is a genuine risk factor if it captures a source of systematic risk for which investors demand compensation. Such a factor's risk price should be positive, meaning exposure to the factor should be associated with higher expected returns. Thus, a well-specified factor model should produce an upward-sloping multivariate security market line.

As discussed in Section 1, many factor models, including the Fama-French (2015) five-factor model, empirically fail to produce an upward-sloping multivariate security market line because estimates for their factors' risk prices are often not reliably positive or are even negative. However, it is critical for typical applications of factor models—for example, risk-adjusting returns or evaluating investment performance—that the model produces an upward-sloping multivariate security market line. This is because accounting for exposures to factors that earn positive mean returns does not make sense if higher exposures are not associated with higher returns.[19]

There are three prominent explanations of why factors may fail to exhibit positive risk prices. First, the factors are not true risk factors, meaning they do not capture a source of systematic risk. Second, the factors are imperfect proxies for the mean-variance efficient portfolio. Third, measurement errors in the factor exposures lead to an errors-in-variables bias in the risk price estimates. Table 7 documents that our informative factors generate a higher maximum Sharpe ratio than the standard factors. Our informative factors therefore address the second explanation: they are better proxies for the mean-variance efficient portfolio than the standard factors.

To evaluate whether our informative factors improve upon the standard factors in producing an upward-sloping multivariate security market line, we estimate their risk prices. For this purpose, we implement the two-stage procedure proposed by Fama and MacBeth (1973) using individual stocks as test assets.[20] In the first stage, we estimate stocks' betas on the factors at the end of each month from June 1969 to December 2019. Specifically, we regress their daily excess returns across the previous 12 months on the standard Fama-French (2015) and our informative five-factor models. We require at least 100 daily observations across the 12-month estimation window to estimate a stock's factor betas.

In the second stage, we regress stocks' compounded excess returns across the 12-month estimation window on their estimated betas; that is, we run at the end of each month from June 1969 to December 2019 the following cross-sectional regression:

$$r_{i,t}^{e} = \gamma_t^{ZB} + \gamma_t^{MP}\hat{\beta}_{i,t}^{MP} + \gamma_t^{SMB}\hat{\beta}_{i,t}^{SMB} + \gamma_t^{HML}\hat{\beta}_{i,t}^{HML} + \gamma_t^{RMW}\hat{\beta}_{i,t}^{RMW} + \gamma_t^{CMA}\hat{\beta}_{i,t}^{CMA} + \epsilon_{i,t} \quad (10)$$

where $r_{i,t}^{e}$ is the compounded return from the beginning of month $t-11$ to the end of month t in excess of the compounded one-month T-bill rate, and $\hat{\beta}_{i,t}^{MP}$, $\hat{\beta}_{i,t}^{SMB}$, $\hat{\beta}_{i,t}^{HML}$, $\hat{\beta}_{i,t}^{RMW}$, and $\hat{\beta}_{i,t}^{CMA}$ are the market, size, value, profitability, and investment betas estimated from the beginning of month $t-11$ to the end of month t. The coefficients γ_t^{MP}, γ_t^{SMB}, γ_t^{HML}, γ_t^{RMW}, and γ_t^{CMA}

[19] It is often argued that the best estimate of a factor's risk premium is its mean return rather than the average slope from cross-sectional regressions of assets' factor loadings on their returns (see, e.g., Lewellen et al., 2010). However, these two approaches yield different information. While a factor's mean return may be a good indicator of the total risk compensation the factor commands, it does not indicate how much of this compensation is for the incremental systematic risk the factor captures. Even if the factor's mean return reflects only compensation for risk, this compensation may be completely explained by other factors, and incremental exposure to this factor beyond exposures to other factors should not reflect compensated risk. By contrast, the average slope from multivariate cross-sectional regressions is informative about the price of the incremental risk the factor captures beyond the risks the other factors capture. If the price for the incremental risk is zero, the factor should not be considered an independent risk factor, and the determination of risk-adjusted returns and investment performance should not account for exposure to this factor. Thus, even though the factor's mean return may be completely due to risk, the factor may not necessarily be included in a factor model.

[20] Using individual stocks rather than portfolios as test assets is advocated, among others, by Jegadeesh et al. (2019), Pukthuanthong et al. (2019), and Ang et al. (2020). See Pukthuanthong et al. (2019) for an extensive discussion of the benefits of using individual stocks rather than portfolios as test assets.

are estimates for the factors' risk prices for the period from month $t-11$ to t.[21] γ_t^{ZB} is the zero-beta rate, which captures the model's common mispricing component and should be zero for a well-specified model. We winsorize all variables at the 1% and 99% levels and use weighted least squares with stocks' market capitalizations as weights. The final risk price estimates are obtained by averaging the monthly γ-estimates across our sample period. For comparison, we also estimate the factors' risk prices in a univariate setting, that is, by using only one of the betas at a time in the estimation of the monthly regressions in (10).

The independent variables (i.e., the betas) in regression (10) are estimated and therefore measure stocks' true factor exposures with error. Hence, the coefficient estimates (i.e., the risk prices) may suffer from an errors-in-variables bias. To account for this bias, we additionally estimate the regressions in (10) with the instrumental variables approach of Jegadeesh et al. (2019) that aims to eliminate the bias. This approach splits each 12-month estimation window for estimating the betas into two parts and estimates the betas in both parts separately. The first set of betas is then used as instruments for the second set of betas, and vice versa. Appendix B describes our implementation of Jegadeesh et al.'s (2019) instrumental variables approach in detail.

Table 8 presents the annualized risk price estimates for the Fama-French (2015) and our informative factors. t-statistics are based on Newey-West (1987) standard errors with 12 lags. Panel A shows that the standard market and size factors carry significantly positive risk prices, no matter whether estimated in a univariate or multivariate setting and whether estimated with weighted least squares or the instrumental variables approach. Thus, stocks with high exposures to the market and size factors earn higher returns than stocks with low exposures. The market and size factors therefore satisfy the theoretically predicted positive relation between factor exposures and returns.

By contrast, the standard value, profitability, and investment factors do not satisfy the positive relation. These factors' estimated risk prices are negative, regardless of the setting and the estimation method. The value factor's risk price estimates are even significantly negative. Higher exposures to the standard value, profitability, and investment factors are thus associated with lower returns. Moreover, the estimated zero-beta rate is significantly positive for both estimation methods, meaning the Fama-French (2015) model creates substantial common mispricing across stocks. Altogether, these findings indicate that the standard factors are not valid risk factor candidates and that the Fama-French (2015) model is misspecified. Consequently, it is not reasonable to use the model for typical applications of factor models.

Panel B shows that our informative value, profitability, and investment factors do much better in satisfying the predicted positive relation between factor exposures and returns. The

[21] We follow the literature (e.g., Ang et al., 2006; Ang and Kristensen, 2012; Kim and Skoulakis, 2018; Jegadeesh et al., 2019; Raponi et al., 2019) and estimate the factors' risk prices in a contemporaneous setting, that is, from stocks' returns across the same period across which their factor betas are estimated. Contrary to the predictive setting, the contemporaneous setting is in line with the typical applications of factor models. In particular, risk-adjusting returns and evaluating investment performance involve regressing assets' returns on the factors' contemporaneous returns.

estimated risk prices for our informative profitability and investment factors are positive, in part even significantly. The estimates for the value factor's risk price are insignificantly negative in the univariate setting but slightly positive in the multivariate setting. Moreover, the estimates for the zero-beta rate are insignificant, suggesting the five-factor model with our informative factors produces little common mispricing across stocks. Hence, our informative factors adhere much more to the theoretical requirements for valid risk factors than the standard factors.

The most reasonable specification is arguably the multivariate setting estimated with the instrumental variables approach. It controls for stocks' exposures to all factors simultaneously and accounts for the potential errors-in-variables bias. This specification delivers the strongest results: all factors' estimated risk prices are positive and—except for the value factor—significant. Our informative factors thus generate an upward-sloping multivariate security market line. The factors' positive risk prices suggest that they are true risk factors capturing systematic sources of risk. This conjecture implies that the standard factors fail to produce positive risk prices because they are bad proxies for the mean-variance efficient portfolio. In addition, the zero-beta rate is small and insignificant, suggesting that our informative factors produce no common mispricing. Consequently, we cannot reject that the model with our informative factors is well-specified, making it a reasonable candidate to be used for typical applications of factor models.

To assess whether the improvements of our informative factors relative to the standard factors are significant, Panel C presents the differences between the factors' estimated risk prices.[22] The differences are unanimously positive and mostly statistically significant for the value, profitability, and investment factors, especially in the multivariate setting. Hence, the improvements of our informative factors in producing an upward-sloping multivariate security market line are significant. Moreover, even though the decrease is insignificant, the estimated zero-beta rates are considerably lower for our informative model than the standard model. Thus, our informative factors produce less common mispricing than the standard factors.[23]

Pukthuanthong et al. (2019) propose a protocol to examine whether a factor is a genuine risk factor. Beyond requiring the factors' risk prices to be positive, their protocol involves two further criteria. First, the factor must be materially related to the covariance matrix of returns. Second, it must exhibit a reward-to-risk ratio that is consistent with risk-based pricing. In the Internet Appendix, we show that all of our informative factors pass these two criteria. Contrary to the standard value, profitability, and investment factors given their negative risk prices, our informative profitability and investment factors are thus valid risk factor candidates according to Pukthuanthong et al.'s (2019) factor protocol.

[22]For the instrumental variables approach, the differences between the factors' risk prices do not add up to the factors' risk prices from Panels A and B. The reason is that the instrumental variables approach treats extreme monthly risk price estimates as missing and that the risk prices for the standard and informative factors may be missing in different months. See Appendix B for further details.

[23]Our informative model produces slightly lower R^2s than the standard model, indicating it does not improve in capturing stocks' idiosyncratic risks. This is not a major concern given that the goal of factor models like the Fama-French (2015) and our informative five-factor models is to capture the cross-section of expected returns, which can be assessed based on the zero-beta rate, rather than the cross-sectional variation in realized returns.

6 Predictive Power of Factor Betas

An upward-sloping multivariate security market line in contemporaneous returns is critical for typical applications of factor models to be sensible. The previous section establishes such an upward-sloping multivariate security market line for our informative factors. Nevertheless, it is also desirable that the betas on the factors capture predictive information, making them useful for investors. We verify this by conducting monthly cross-sectional Fama-MacBeth (1973) regressions that predict stocks' one-, three-, six-, and 12-month ahead excess returns based on the factor betas. As previously, we winsorize stocks' betas at the 1% and 99% levels, use weighted least squares with stocks' market capitalizations as weights, and run the regressions in univariate and multivariate settings.

Table 9 presents the average annualized coefficients. t-statistics are based on Newey-West (1987) standard errors with lags equal to the return horizon. Regarding the predictive power of betas on the standard factors, Panel A shows that only the profitability beta exhibits significantly positive predictive power for returns. Higher size and value betas are also associated with higher future returns, but their predictive power is insignificant. The investment beta is not unanimously positively related to future returns, and the market beta is negatively related to future returns. These results are consistent across the different settings and return horizons. Taken together, the predictive power of betas on the standard factors is weak and mostly not robustly positive.

By contrast, Panel B reveals that betas on our informative factors almost unanimously positively predict returns. While the relations between betas and future returns are insignificant in the univariate setting, they are predominantly significant or only marginally insignificant in the multivariate setting. Especially the significantly positive predictive power of betas on our informative size, value, and investment factors is novel compared to the standard factors. Hence, betas on our informative factors are more powerful in predicting returns than betas on the standard factors, corroborating the usefulness and superiority of our informative factors.

7 Comparison to other Enhancement Procedures

Several further procedures to enhance the Fama-French (2015) factors have been proposed in recent years. Most prominent are Daniel et al.'s (2020) hedging procedure, Fama and French's (2020) cross-section procedure, and Ehsani and Linnainmaa's (2022c) time-series efficiency procedure. This section evaluates how our enhancement procedure compares to these other procedures. For this purpose, we compare our informative factors to the factors obtained from these procedures in an analogous manner as we compare them to the Fama-French (2015) factors in Tables 4 to 9. Moreover, we also apply these other enhancement procedures to our informative factors to examine whether these procedures complement our procedure.

7.1 Hedging Procedure

Daniel et al.'s (2020) hedging procedure aims to hedge the factors' unpriced sources of variation to reduce their volatility without affecting their mean returns. The approach is to construct hedge portfolios with high exposures to the factors but close to zero mean returns. The factors' unpriced variation is then, roughly speaking, hedged by going long the factors and short their hedge portfolios. Appendix C details the exact procedure to construct the hedged versions of a given factor model's factors.

Panel A of Table 10 presents results on the hedged versions of the Fama-French (2015) factors. The hedged factors' mean returns are lower than the standard factors', but so are their volatilities (compare Panel A of Table 4). The latter effect mostly dominates the former, resulting in higher Sharpe ratios for the hedged factors. In line with Daniel et al. (2020), this result indicates that the hedging procedure improves the Fama-French (2015) factors. Like the standard and informative value factors, the hedged value factor is redundant, exhibiting a marginally insignificant spanning regression alpha relative to the other hedged factors.

The last five columns of Panel A compare the hedged factors to our informative factors. Our informative factors' mean returns and Sharpe ratios are unanimously and in part significantly higher than those of the hedged factors. Nevertheless, our informative model produces significant alphas for the hedged market, value, and investment factors, meaning it cannot price them. Conversely, the hedged model cannot price our informative value, profitability, and investment factors. Thus, the two sets of factors capture complementary pricing information.

Panel C reveals that the informative model's maximum squared Sharpe ratio is somewhat lower than the hedged model's (0.155 vs. 0.171). Thus, the hedged factors are more mean-variance efficient than our informative factors. Nevertheless, Barillas et al.'s (2020) test and the results from Fama and French's (2018) bootstrap approach indicate that the maximum Sharpe ratios are not significantly different, and their difference actually decreases in the bootstrap simulations. Hence, the hedged and informative factors exhibit a similar pricing performance.

Panel D of Table 10 presents risk price estimates for the hedged factors and compares them to the informative factors' risk prices from Panel B of Table 8. Amid the result from Panel C that the hedged factors are somewhat better proxies for the mean-variance efficient portfolio than our informative factors, one might expect them to produce a reasonable upward-sloping multivariate security market line as well. This is, however, not the case: the risk price estimates for the hedged value and profitability factors are negative, and those for the hedged investment factor are insignificant, regardless of the estimation method. Our informative factors' risk prices are unanimously and mostly significantly higher. Hence, despite being somewhat better proxies for the mean-variance efficient portfolio, the hedged factors seem to capture less risk of hedging concern for investors than our informative factors and fail to qualify as valid risk factor candidates.

Next, we apply the hedging procedure to our informative factors. We refer to the resulting factors as informative hedged factors. Panel B of Table 10 shows that they exhibit slightly lower mean returns and considerably lower volatilities than our usual informative factors, resulting in unanimously higher Sharpe ratios (compare Panel B of Table 4). Notably, all of the informative hedged factors exhibit significant spanning regression alphas. Thus, neither of them is redundant, not even the value factor. These findings suggest that applying the hedging procedure is, on an individual basis, beneficial for our informative factors.

Panel B further documents that our informative hedged factors exhibit uniformly, and in part significantly, higher mean returns and Sharpe ratios than the standard hedged factors. Importantly, given their significant alphas, our informative hedged factors cannot be priced by the hedged model. Conversely, the informative hedged model also fails to fully price the standard hedged factors, producing significant alphas for the value and investment factors. Thus, the standard hedged factors still capture incremental pricing information beyond the informative hedged factors. Nevertheless, the latter seem to capture more independent pricing information.

The results in Panel C corroborate this conjecture: the informative hedged factors produce a higher maximum squared Sharpe ratio than the standard hedged factors (0.231 vs. 0.171). Barillas et al.'s (2020) test indicates that the difference is marginally significant. Moreover, the informative hedged model achieves higher maximum Sharpe ratios than the standard hedged model in the vast majority of bootstrap simulation runs. Overall, the informative hedged factors should therefore achieve a better pricing performance than the standard hedged factors.

Furthermore, Panel D reveals that the informative hedged factors improve upon the standard hedged factors in producing a positive relation between factor exposures and returns. Specifically, the informative hedged value, profitability, and investment factors exhibit higher risk prices, in part significantly, than their standard hedged counterparts. Nevertheless, the informative hedged factors still fail to produce an upward-sloping multivariate security market line. The profitability factor's risk price estimates remain negative, and the size and value factors' risk price estimates are insignificant, regardless of the estimation method. Moreover, the estimated zero-beta rates are large. Hence, the hedging procedure harms our informative factors' ability to generate an upward-sloping multivariate security market line and turns them thus into invalid risk factor candidates.

Finally, Panel E examines the predictive power of betas on the hedged factors. Betas on the standard and informative hedged factors are mostly positively related to future returns. However, contrary to the betas on our usual informative factors (compare Panel B of Table 9), their predictive power is insignificant throughout. Betas on the hedged factors are thus less informative about future returns than betas on our usual informative factors, meaning the hedging procedure also impairs the predictive power of factor betas on our informative factors.

Overall, this subsection's results show that the informative and hedged factors compete well with each other on an individual basis and capture complementary pricing information. Applying the hedging procedure to our informative factors improves their mean-variance efficiency, implying that the hedging procedure complements our enhancement procedure in this regard. Nevertheless, the standard and informative hedged factors fail to generate an upward-sloping multivariate security market line. Unlike our informative factors, the hedged factors therefore do not satisfy the requirements for valid risk factor candidates. The hedging procedure is detrimental rather than complementary to our enhancement procedure in this regard, overshadowing the improvements in the factors' mean-variance efficiency.

7.2 Cross-Section Procedure

Fama and French's (2020) cross-section procedure constructs factors from Fama-MacBeth (1973) regressions. Specifically, the factors are obtained as the regression slopes from cross-sectional regressions that regress the returns of the factor portfolios of a given factor model on the characteristics based on which the factor portfolios are constructed (e.g., market equity, book-to-market, operating profitability, and investment in the case of the Fama-French (2015) factors). The regression slopes can be interpreted as zero-investment portfolios that have only exposure to the respective characteristic but zero exposure to the other characteristics. Fama and French (2020) argue that these factors are expected to outperform the standard factors as they are obtained from an optimization rather than ad hoc portfolio sorts. Appendix D details the exact procedure to construct the cross-section versions of a given factor model's factors.

Panel A of Table 11 presents results on the cross-section versions of the Fama-French (2015) factors. Their mean returns are, apart from the size factor, significantly positive, and their Sharpe ratios are mostly somewhat higher than those of the standard factors (compare Panel A of Table 4). Hence, the cross-section procedure in fact improves the standard factors. Notably, the size and value factors are redundant in the cross-section model, exhibiting insignificant spanning regression alphas.

Comparing the cross-section factors to our informative factors reveals that our informative factors have uniformly, and partly significantly, higher Sharpe ratios than their cross-section counterparts. Moreover, our informative model prices the cross-section factors well, producing only for the investment factor a significant alpha. By contrast, the cross-section model performs badly in pricing our informative factors, leaving significant alphas for the value, profitability, and investment factors. Thus, our informative factors capture the cross-section factors' pricing information much better than vice versa. Nevertheless, the two sets of factors still contain some complementary pricing information.

Panel C documents that the informative model's maximum squared Sharpe ratio is considerably higher than the cross-section model's (0.155 vs. 0.113). Barillas et al.'s (2020) test indicates that the difference is marginally significant. Moreover, the results from the bootstrap approach show that our informative model achieves higher maximum Sharpe ratios than the cross-section model in more than 90% of the simulation runs. Thus, our informative factors achieve a better pricing performance than the cross-section factors.

Panel D shows that the risk price estimates for the cross-section value, profitability, and investment factors are predominately negative. Thus, the cross-section factors fail to produce an upward-sloping multivariate security market line. Our informative factors' risk prices are unanimously and mostly significantly higher than those of the cross-section factors. Additionally, the estimates for the cross-section model's zero-beta rate are sizable. Contrary to our informative factors, the cross-section factors are invalid risk factor candidates.

Next, we apply the cross-section procedure to our informative factors. We refer to the resulting factors as informative cross-section factors. Panel B shows that, unlike the standard cross-section factors, all informative cross-section factors exhibit significantly positive mean returns and spanning regression alphas. Thus, all of them capture significant incremental pricing information. However, the informative cross-section factors exhibit mostly lower Sharpe ratios than our usual informative factors (compare Table 4). Hence, the cross-section procedure is, on an individual basis, detrimental to our informative factors.

Comparing the informative and standard cross-section factors shows that the formers' mean returns are, albeit mostly insignificantly, lower than the latters'. The picture improves somewhat regarding the Sharpe ratios: only the informative cross-section investment factor's Sharpe ratio decreases relative to its standard cross-section counterpart. Moreover, the standard cross-section model leaves significant alphas for the informative cross-section size and profitability factors. Conversely, the standard cross-section investment factor has a significant alpha relative to the informative cross-section model. Thus, the informative cross-section factors hardly improve upon the standard cross-section factors, but both sets of factors capture some complementary pricing information.

Panel C reveals that the informative cross-section factors produce a higher maximum squared Sharpe ratio than the standard cross-section factors (0.135 vs. 0.113). However, Barillas et al.'s (2020) test indicates that the increase is insignificant, suggesting their pricing performance is similar. Furthermore, the informative cross-section model's Sharpe ratio of 0.135 represents a decline compared to our usual informative model's Sharpe ratio of 0.155. These observations corroborate our conjectures that combining the cross-section procedure with our enhancement procedure is mildly beneficial relative to the standard cross-section factors but detrimental relative to our usual informative factors.

Panel D shows that the informative cross-section factors improve upon the standard cross-section factors in producing a positive relation between factor exposures and returns. The informative cross-section value, profitability, and investment factors' risk prices increase uniformly and mostly significantly relative to the corresponding standard cross-section factors' risk prices. These increases result in positive risk price estimates for the informative cross-section profitability and investment factors. In addition, the estimates for the informative cross-section model's zero-beta rate are insignificant. However, amid its value factor's negative risk price estimates, the informative cross-section model can still not keep up with our usual informative model in producing an upward-sloping multivariate security market line.

Finally, Panel E examines the predictive power of betas on the cross-section factors. The results show that only the beta on the standard cross-section profitability factor predicts returns significantly positively. Betas on the standard cross-section size and value factors are positively but insignificantly associated with future returns. The beta on the standard cross-section investment factor is negatively associated with future returns. The picture is similar for the informative cross-section factors. Hence, the predictive power of betas on the cross-section factors is overall weaker than of betas on our usual informative factors (compare Panel B of Table 9). Hence, the cross-section procedure results in less informative factor betas.

In sum, this subsection's results show that our enhancement procedure clearly outperforms the cross-section procedure in all aspects. Moreover, while applying the cross-section procedure improves the Fama-French (2015) factors, it is detrimental to our informative factors.

7.3 Time-Series Efficiency Procedure

Ehsani and Linnainmaa's (2022c) time-series efficiency procedure builds on the finding of Ehsani and Linnainmaa (2022b) that factors exhibit positive time-series momentum. The time-series efficient versions of factors condition on their time-series momentum by, roughly speaking, scaling up (down) the investment in the factors if their momentum is positive (negative). Appendix E details the exact procedure to construct the time-series efficient versions of a given factor model's factors.

Panel A of Table 12 presents results on the time-series efficient versions of the Fama-French (2015) factors. The time-series efficient factors have, in general, similar mean returns as their standard counterparts but exhibit lower volatilities, resulting in higher Sharpe ratios (compare Panel A of Table 4). In line with Ehsani and Linnainmaa's (2022c) conclusion, the time-series efficiency procedure thus improves the standard factors.

The time-series efficient factors' mean returns and Sharpe ratios are generally similar to our informative factors'. Moreover, the time-series efficient factors and our informative factors can hardly price each other, producing mostly significant alphas for the factors of the other model. Thus, the pricing information captured by the two models is largely complementary.

While the time-series efficient factors compare well with our informative factors on an individual basis, Panel C shows that the time-series efficient factors generate a lower maximum squared Sharpe ratio than our informative factors (0.110 vs. 0.155). Yet, Barillas et al.'s (2020) test indicates that the difference is not statistically significant. The difference increases in the bootstrap simulations, suggesting that our informative model's outperformance may in truth be more pronounced, but it still remains just short of significant.

Panel D documents that the ability of the time-series efficient factors to produce an upward-sloping multivariate security market line is underwhelming. The estimates for the time-series efficient value, profitability, and investment factors' risk prices are almost unanimously and mostly significantly negative. Our informative factors' risk prices are throughout and often significantly higher than those of the corresponding time-series efficient factors. Moreover, the estimates for the time-series efficient model's zero-beta rate are large and significant. Hence, the time-series efficient factors do not qualify as valid risk factor candidates.

Next, we apply the time-series efficiency procedure to our informative factors. We refer to the resulting factors as informative time-series efficient factors. Panel B shows that all informative time-series efficient factors exhibit significantly positive mean returns. Their mean returns are lower than those of the informative factors, but so are their volatilities, resulting in similar Sharpe ratios (compare Panel B of Table 4). Thus, the time-series efficiency procedure hardly improves our informative factors, suggesting they are already time-series efficient.

Comparing the informative time-series efficient factors to the standard time-series efficient factors reveals that their mean returns and Sharpe ratios are generally quite similar. Thus, the informative time-series efficient factors hardly improve upon the standard time-series efficient factors as well. Nevertheless, the informative time-series efficient factors seem to capture more independent pricing information as the standard time-series efficient factors have more problems in pricing them than vice versa.

In line with this conjecture, Panel C shows that the informative time-series efficient factors produce a higher maximum squared Sharpe ratio than the standard time-series efficient factors (0.157 vs. 0.110). Barillas et al.'s (2020) test indicates that the difference is marginally significant. Moreover, the results from the bootstrap approach reveal that the informative time-series efficient model outperforms the standard time-series efficient model in the vast majority of simulation runs. Thus, the informative time-series efficient factors achieve a better pricing performance than the standard time-series efficient factors. However, they hardly improve upon our usual informative factors given their similar maximum squared Sharpe ratios (0.157 vs. 0.155), corroborating that the time-series efficiency procedure hardly enhances our informative factors.

Panel D shows that the informative time-series efficient factors perform better than the standard time-series efficient factors in producing a positive relation between factor exposures and returns. The risk price estimates for the informative time-series efficient factors are unanimously and mostly significantly higher than those of the standard time-series efficient factors. Nevertheless, the informative time-series efficient value, profitability, and investment factors' risk price estimates are still insignificant and partly negative. Additionally, the estimated zero-beta rates are large and significant. Thus, the time-series efficiency procedure is detrimental to our informative factors' ability to generate an upward-sloping multivariate security market line.

Finally, regarding the predictive power of betas on the standard time-series efficient factors, Panel E shows that only the profitability beta predicts returns significantly positively. The picture improves for the informative time-series efficient factors. Specifically, betas on the value and investment factors are now more positively, albeit mostly insignificantly, associated with future returns. However, betas on both types of time-series efficient factors are less informative about future returns than betas on our usual informative factors (compare Panel B of Table 9). Hence, the time-series efficiency procedure harms the predictive power of factor betas on our informative factors.

Overall, this subsection's results reveal that the time-series efficiency procedure leads to similar improvements in the individual Fama-French (2015) factors as our enhancement procedure. However, combined in a model, our informative factors outperform the time-series efficient factors regarding their mean-variance efficiency and in generating an upward-sloping multivariate security market line. Moreover, the time-series efficiency procedure hardly improves our informative factors, suggesting they are already time-series efficient. The time-series efficiency procedure rather harms our informative factors' ability to generate an upward-sloping multivariate security market line and therefore turns them into invalid risk factor candidates.

8 Pricing of Characteristics-Sorted Portfolios

The ultimate purpose of factors and factor models is to price assets. In the empirical asset pricing literature, portfolios sorted according to characteristics are of particular interest. To verify the practical usefulness of our informative factors, we compare their performance in pricing characteristics-sorted portfolios to the standard Fama-French (2015) factors as well as the factors obtained from the alternative enhancement procedures discussed in Section 7. For this purpose, we obtain 113 sets of characteristics-sorted decile portfolios from the global-q website of Kewei Hou, Chen Xue, and Lu Zhang.[24,25] The characteristics cover the six broad anomaly categories defined by Hou et al. (2020): momentum (21), value-versus-growth (24), investment (26), profitability (18), intangibles (15), and frictions (9). Table A1 lists and briefly describes the anomaly characteristics.

We use five metrics to assess models' pricing performance: (1) the average absolute alpha, (2) the fraction of significant alphas, (3) the ratio of the average absolute alpha to the average absolute deviation from the mean returns' mean, reflecting the unexplained proportion of the mean returns' dispersion, (4) the cross-sectional R^2, measuring the explained proportion of the mean returns' variance,[26] and (5) the average time-series R^2.

Table 13 presents the results. Panel A documents that our informative factors strongly outperform the Fama-French (2015) factors in pricing the full set of portfolios based on almost all metrics. In particular, our informative factors produce a lower average absolute alpha (0.098% vs. 0.108%) and a lower fraction of significant alphas (17.9% vs. 26.0%), indicating that they price the individual portfolios better. Moreover, the informative factors also improve in explaining the portfolios' cross-section, leaving a lower proportion of the dispersion in mean returns unexplained (93.2% vs. 102.6%) and producing a higher cross-sectional R^2 (13.6% vs. -1.5%) than the Fama-French (2015) factors. The only metric based on which our informative factors do not outperform the Fama-French (2015) factors is the portfolios' average time-series R^2 (86.4% vs. 87.4%). The results for the different anomaly categories show that our informative factors consistently outperform the Fama-French (2015) factors in every category except the frictions category. Overall, these findings demonstrate that our informative factors achieve a

[24] We use only those decile portfolio sets with available data for the full sample period from July 1968 to December 2019.

[25] `http://global-q.org/testingportfolios.html`

[26] The cross-sectional R^2 is calculated as one minus the ratio of the alphas' variance to the mean returns' variance.

better pricing performance in practical applications.

The remaining panels of Table 13 present the results for the factors' hedged, cross-section, and time-series efficient versions. First, the cross-section versions perform similarly to the standard versions, the hedged versions perform somewhat worse than the standard versions, and the time-series efficient versions perform considerably worse than the standard versions. Thus, unlike the improvements in the factors generated by our enhancement procedure, the improvements generated by the hedging, cross-section, and time-series efficiency procedures do not result in an improved pricing of the characteristics-sorted portfolios.

Second, the hedged, cross-section, and time-series efficient versions of our informative factors outperform, in general, the corresponding versions of the Fama-French (2015) factors. Hence, the improvements of the informative factors' hedged, cross-section, and time-series efficient versions relative to the standard factors' hedged, cross-section, and time-series efficient versions as observed in Tables 10, 11, and 12 translate to an improved pricing of characteristics-sorted portfolios.

Finally, our informative factors strongly outperform not only the standard Fama-French (2015) factors but also the Fama-French (2015) factors' hedged, cross-section, and time-series efficient versions. This observation corroborates that our enhancement procedure is superior to the hedging, cross-section, and time-series efficiency procedures.

9 Conclusion

In this study, we propose a procedure to enhance characteristics-based factors. In particular, we argue that variation in factors' underlying characteristics reflects not only differences in expected returns but also in other dimensions. Canceling the part of the characteristics' variation unrelated to expected returns should yield better indicators of expected returns. Hence, factors built from the characteristics' informative parts should improve upon the original factors. We apply this idea to the value, profitability, and investment factors of Fama and French (2015) by canceling the cash flow information embedded in book-to-market, profitability, and investment. Our approach successfully narrows down the characteristics' information about expected returns, as only our adjusted characteristics predict returns, whereas the canceled parts do not.

Our informative factors based on the characteristics' informative parts substantially improve upon the standard factors. First, our informative value, profitability, and investment factors exhibit higher individual Sharpe ratios than their standard counterparts. The higher Sharpe ratios emanate from higher mean returns as well as lower volatilities. Thus, the characteristics' informative parts identify differences in expected returns more accurately and with higher certainty than the raw characteristics. Our informative factors also generate a significantly higher maximum Sharpe ratio than the Fama-French (2015) factors. Following the arguments of Barillas and Shanken (2017), this result implies that our informative factors achieve a better pricing performance than the Fama-French (2015) factors. We further show that our informative factors largely subsume the pricing information of the standard Fama-French (2015) factors. By contrast, our informative factors capture significant pricing information relative to the standard Fama-French (2015) factors as well as the four-factor model of Hou et al. (2015). They are also largely robust to alternative refinements of the factors proposed in the literature.

An essential requirement genuine risk factors need to satisfy is that returns are positively related to factor exposures. The Fama-French (2015) value, profitability, and investment factors exhibit negative risk prices and thus fail to satisfy this requirement. By contrast, our informative value, profitability, and investment factors achieve positive risk prices and thus generate an upward-sloping multivariate security market line. Moreover, unlike the standard factors, our informative factors generate an insignificant zero-beta rate and thus produce little common mispricing for the cross-section of stocks. Additionally, besides positive risk prices, they also satisfy the two further criteria of Pukthuanthong et al.'s (2019) factor protocol as they are materially related to the return covariance matrix and generate reasonable reward-to-risk ratios. Hence, our informative factors meet the requirements for genuine risk factors. Thus, they are more appropriate than the Fama-French (2015) factors for typical applications of factor models, such as risk-adjusting returns or evaluating investment performance.

We also compare our enhancement procedure to the hedging procedure of Daniel et al. (2020), the cross-section procedure of Fama and French (2020), and the time-series efficiency procedure of Ehsani and Linnainmaa (2022c). Overall, our procedure outperforms these procedures. First, it produces larger increases in the individual factors' Sharpe ratios. In addition, our informative factors achieve a higher maximum Sharpe ratio than the cross-section and time-series efficient factors. Importantly, unlike our informative factors, the factors obtained from the other procedures fail to generate an upward-sloping multivariate security market line and therefore do not qualify as valid risk factor candidates. Furthermore, applying these alternative enhancement procedures to our informative factors hardly improves them, implying that they do not suffer from the shortcomings addressed by these procedures. Applying these procedures is even detrimental as they impair our informative factors' ability to generate positive risk prices and thus turn them into invalid risk factor candidates.

We further demonstrate the superiority of our informative factors by employing them to price a large cross-section of characteristics-sorted portfolios. They outperform the Fama-French (2015) factors as well as the factors obtained from the hedging, cross-section, and time-series efficiency procedures across almost all metrics and anomaly categories. The factors from the alternative procedures even underperform the standard Fama-French (2015) factors, questioning whether these procedures actually enhance the factors' pricing power.

The empirical asset pricing literature proposed many factors. Most studies are content with identifying an indicator of expected returns—no matter whether theoretically motivated—and constructing a factor based on the indicator. This approach is insufficient to obtain efficient factors with good pricing performance, even for theoretically motivated factors like those of Fama and French (2015). Instead, indicators' variation that is informative about expected returns should be narrowed down before constructing factors. Our findings show that factors following this principle have the potential to be more efficient, achieve better pricing power, and adhere more to the requirements of genuine risk factors than factors that fail to do so.

Finally, our findings have valuable practical implications. Investment strategies based on factors detected in academic studies have been widely adopted in the investment management industry because of their attractive historical risk-return profiles. Our results show how the performance of such factor investing strategies can be further boosted. In particular, strategies based on expected return indicators whose variation that is uninformative about expected returns is canceled harvest the factor premia more successfully. Thereby, the improvements emanate from higher mean returns as well as lower volatilities.

A Variable Definitions

Market Equity (ME): A stock's market equity for the end of month t is calculated as its price times its shares outstanding at the end of month t. To reduce the skewness in ME, we take its natural logarithm. ME is considered missing if it is non-positive.

Book-to-Market (BM): A stock's book-to-market for the end of June of year y is calculated as book equity from the firm's last fiscal year ending in year $y-1$, divided by market equity at the end of the month of the fiscal year ending.[27] Following Davis et al. (2000), book equity (BE) is the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock (depending on availability, the redemption, liquidation, or par value of preferred stock is used, in that order); if the book value of stockholders' equity is not directly available, it is measured as the book value of common equity plus the par value of preferred stock or as the difference between total assets and total liabilities (in that order). To reduce the skewness in BM, we take its natural logarithm. BM is considered missing if market equity or book equity is non-positive.

Operating Profitability (OP): A stock's operating profitability for the end of June of year y is calculated as revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses, divided by book equity, all from the firm's last fiscal year ending in year $y-1$. OP is considered missing if revenues are missing, if each of cost of goods sold, interest expense, and selling, general, and administrative expenses are missing, or if book equity is non-positive.

Investment (INV): A stock's investment for the end of June of year y is calculated as total assets from the firm's last fiscal year ending in year $y-1$, divided by total assets from the firm's last fiscal year ending in year $y-2$, minus 1. To reduce the skewness in INV, we take its natural logarithm. INV is considered missing if total assets are non-positive.

Cash Profitability (CP): A stock's cash profitability for the end of June of year y is calculated as revenues minus cost of goods sold, interest expense, selling, general, and administrative expenses, and accruals, divided by book equity, all from the firm's last fiscal year ending in year $y-1$. Following Ball et al. (2016), accruals are calculated as the changes in accounts receivable, prepaid expenses, and inventory minus the changes in accounts payable, deferred revenue, and accrued expenses. Changes are calculated based on the firm's last fiscal year endings in year $y-2$ and year $y-1$. Missing changes are set to zero. CP is considered missing if revenues are missing, if each of cost of goods sold, interest expense, and selling, general, and administrative expenses are missing, or if book equity is non-positive.

[27] This construction of book-to-market slightly differs from Fama and French (2015), who divide book equity by market equity from the end of December of year $y-1$.

B Instrumental Variables Approach

Our implementation of the instrumental variables approach proposed by Jegadeesh et al. (2019) is as follows: we first split every 12-month estimation window into two subsets on a daily basis; that is, the first, third, fifth, ... day of the estimation window is assigned to the first subset, and the second, fourth, sixth, ... day of the estimation window is assigned to the second subset. Within each of these two subsets, we estimate stocks' betas on the factors of a given factor model. Then, we regress each beta estimated based on the first subset on all betas estimated based on the second subset. Formally, we run for each beta the following cross-sectional regression:

$$\beta_{i,t}^{k,1} = \delta_{0,t} + \sum_{k=1}^{5} \delta_{k,t} \beta_{i,t}^{k,2} + \epsilon_{i,t} \tag{B1}$$

where $\beta_{i,t}^{k,1}$ ($\beta_{i,t}^{k,2}$) is the beta on factor k estimated based on the first (second) subset of the estimation window from month $t-11$ to t. We use weighted least squares with stocks' market capitalizations as weights and winsorize the betas on the 1% and 99% levels.

In the second stage, we use the betas' fitted values from (B1) as explanatory variables in monthly cross-sectional Fama-MacBeth (1973) regressions like in (10). The dependent variable is the compounded return across the days in the first subset in excess of the compounded T-bill rate. Again, we use weighted least squares with stocks' market capitalizations as weights and winsorize the variables on the 1% and 99% levels. From these regressions, we obtain monthly estimates for the factors' risk prices.

We repeat the procedure by switching the roles of the first and second subsets; that is, the $\beta_{i,t}^{k,1}$ are now the instruments for the $\beta_{i,t}^{k,2}$ in (B1), and the dependent variable in the monthly cross-sectional Fama-MacBeth (1973) regressions is the compounded return across the days in the second subset in excess of the compounded T-bill rate. Thereby, we obtain a second set of monthly estimates for the factors' risk prices. We calculate the final risk price estimates by taking the average of the two risk price estimates each month and then averaging the monthly average risk prices across the full sample period.

Jegadeesh et al. (2019) highlight the possibility that the cross-product of the dependent and independent betas in the estimation of the regression in (B1) may be close to singular. This would lead to unreasonably large risk price estimates. Following Jegadeesh et al. (2019), we address this issue by treating monthly risk price estimates deviating six or more standard deviations of the corresponding factor from their mean as missing.

C Hedged Factors

We construct hedged versions for the factors of a given factor model following the methodology of Daniel et al. (2020); that is, we first construct hedge portfolios for the factors and then determine the optimal hedge ratios.

To construct the hedge portfolios, we estimate stocks' betas on the model's factors from multivariate regressions at the end of June from 1968 to 2019. As inspired by Frazzini and Pedersen (2014), two estimation windows are used. First, the stocks' and factors' volatilities are estimated from daily log returns across the previous 12 months (i.e., from the beginning of July of year $t-1$ until the end of June of year t). Second, stocks' correlations with the factors as well as the correlations between the factors are estimated from overlapping three-day cumulative log returns across the previous 60 months (i.e., from the beginning of July of year $t-5$ until the end of June of year t). We only consider daily returns for which the respective stock has non-missing prices for the same and the previous day. Moreover, we do not use the actual factor returns across the previous 12 respectively 60 months but rather hypothetical factor returns. These hypothetical factor returns are obtained by assuming that the factor portfolios on each day across the previous 12 respectively 60 months consisted of the same stocks with the same weights as the factor portfolios at the end of June of year t. Following Daniel et al. (2020), we include a dummy variable that equals one if the return observation is from the period between January and June of year t when estimating stocks' factor betas. Finally, we require stocks to have at least 100 daily return observations across the previous 12 months and at least 15 daily return observations across the previous six months.

To construct the hedge portfolio for the model's value factor, stocks are at the end of June sorted into terciles according to their size at the end of June and into terciles according to their book-to-market from the last fiscal year ending in the previous year.[28] Breakpoints are based only on NYSE stocks. Intersecting the size and book-to-market terciles yields nine portfolios. Within these nine portfolios, stocks are sorted into terciles according to their estimated beta on the value factor. The stocks in the 27 resulting portfolios are value-weighted. The value factor's hedge portfolio is obtained by going long the equal-weighted combination of the nine high value beta portfolios and short the equal-weighted combination of the nine low value beta portfolios.

The hedge portfolio for the model's profitability (investment) factor is constructed in the same way as the value factor's hedge portfolio, except that the respective measure of operating profitability (investment) and the beta on the profitability (investment) factor are used.

The construction of the hedge portfolios for the model's market and size factors uses the 27 portfolios from the bivariate sorts on size and any of book-to-market, operating profitability, and investment. To construct the market (size) factor's hedge portfolio, the stocks within each of the 27 portfolios are sorted into terciles according to their estimated betas on the market (size) factor. The stocks in the 81 resulting portfolios are value-weighted. The market (size)

[28]The book-to-market used for the sort is the one used for constructing the respective value factor. That is, the book-to-market used to construct the hedge portfolio of our informative value factor is the adjusted book-to-market defined in Section 3.3.

factor's hedge portfolio is obtained by going long the equal-weighted combination of the 27 high market (size) beta portfolios and short the equal-weighted combination of the 27 low market (size) beta portfolios.

The factors' hedged versions are obtained as follows:

$$r_t^{f,H} = r_t^f - r_t^h \delta_t^f \tag{C1}$$

where $r_t^{f,H}$ is the return on factor f's hedged version, r_t^f is the return on factor f's unhedged version, r_t^h is the vector of returns on the factors' hedge portfolios, and δ_t^f is the vector of factor f's hedge ratios. The hedge ratios are the betas of the unhedged factors on the hedge portfolios and are determined at the end of June. To determine the hedge ratios, volatilities are calculated from daily log returns across the previous 12 months, and correlations are calculated from overlapping three-day cumulative log returns across the previous 60 months. As previously, hypothetical factor and hedge portfolio returns obtained by assuming constant portfolio compositions and weights are used.

D Cross-Section Factors

We construct cross-section versions for the factors of a given factor model following the methodology of Fama and French (2020). The construction of the cross-section factors is based on the 18 portfolios used to construct the factors' usual versions (see Section 4.1); that is, based on the portfolios resulting from the bivariate sorts on size and any of book-to-market, profitability, and investment (respectively their adjusted versions). The cross-section factors are obtained from monthly cross-sectional Fama-MacBeth (1973) regressions that regress the factor portfolios' returns on their characteristics:[29]

$$r_{p,t} = r_{Z,t} + r_{ME,t}ME_{p,t} + r_{BM,t}BM_{p,t} + r_{OP,t}OP_{p,t} + r_{INV,t}INV_{p,t} + \epsilon_{p,t} \tag{D1}$$

where $r_{p,t}$ is portfolio p's return in month t and $ME_{p,t}$, $BM_{p,t}$, $OP_{p,t}$, and $INV_{p,t}$ are the portfolio's market equity, book-to-market, operating profitability, and investment. The cross-section size, value, profitability, and investment factors are the regression slopes $r_{ME,t}$, $r_{BM,t}$, $r_{OP,t}$, and $r_{INV,t}$, respectively. Since the relation of market equity and investment with returns is negative, we multiply $r_{ME,t}$ and $r_{INV,t}$ by -1 to obtain the usual positive factor mean returns.

In the case of the Fama-French (2015) five-factor model, book-to-market, operating profitability, and investment are the raw versions described in Appendix A; in the case of our informative five-factor model, book-to-market, operating profitability, and investment are the adjusted versions described in Section 3. The portfolios' characteristics are calculated as the value-weighted averages of their constituent stocks' characteristics. Market equity is from the beginning of month t; book-to-market, operating profitability, and investment are from the last fiscal year ending in the previous year if t is between July and December and from the last fiscal year ending in the year before the previous year if t is between January and June. The portfolios' characteristics are standardized to have means of zero and standard deviations of one.

[29] The cross-section factors' daily returns are obtained from daily Fama-MacBeth (1973) regressions.

E Time-Series Efficient Factors

We construct time-series efficient versions for the factors of a given factor model following the methodology of Ehsani and Linnainmaa (2022c). Specifically, we construct the real-time implementable versions of the time-series efficient factors as follows:

$$\begin{aligned} r_t^{f,TE} &= w_t^f r_t^f \\ w_t^f &= \mu_t \frac{SR_t^2 + 1}{SR_t^2 + \rho_t^2} \frac{\mu_t(1-\rho_t) + \rho_t r_{t-1}^f}{(\mu_t(1-\rho_t) + \rho_t r_{t-1}^f)^2 + (1-\rho_t^2)\sigma_t^2} \end{aligned} \tag{E1}$$

where $r_t^{f,TE}$ is the return on factor f's time-series efficient version in month t, r_t^f is the return on the factor's usual version in month t, and μ_t, σ_t, SR_t, and ρ_t are estimates for the factor's expected return, standard deviation, Sharpe ratio, and first-order autocorrelation, respectively, in month t.[30] The weight w_t^f is constrained to be in the interval [0,1].

μ_t, σ_t, SR_t, and ρ_t are re-estimated each month based on data across the past 120 months. Thereby, each factor's parameters are estimated not only based on the factor's data but rather based on pooled data across all factors. Consequently, the same parameter estimates are used for all factors of the model. Following Ehsani and Linnainmaa (2022c), we require at least two months of data to estimate the parameters; that is, the time-series efficient factors are first calculated in September 1968.

[30] The time-series efficient factors' daily returns are calculated by conditioning in (E1) on the previous day's return rather than the previous month's return.

References

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2006, Downside risk, *Review of Financial Studies* 19, 1191–1239.

Ang, Andrew, and Dennis Kristensen, 2012, Testing conditional factor models, *Journal of Financial Economics* 106, 132–156.

Ang, Andrew, Jun Liu, and Krista Schwarz, 2020, Using stocks or portfolios in tests of factor models, *Journal of Financial and Quantitative Analysis* 55, 1–72.

Asness, Clifford, and Andrea Frazzini, 2013, The devil in HML's details, *Journal of Portfolio Management* 39, 49–68.

Ball, Ray, Joseph Gerakos, Juhani T. Linnainmaa, and Valeri Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, *Journal of Financial Economics* 121, 28–45.

Ball, Ray, Joseph Gerakos, Juhani T. Linnainmaa, and Valeri Nikolaev, 2020, Earnings, retained earnings, and book-to-market in the cross section of expected returns, *Journal of Financial Economics* 135, 231–254.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barillas, Francisco, Raymond Kan, Cesare Robotti, and Jay Shanken, 2020, Model comparison with Sharpe ratios, *Journal of Financial and Quantitative Analysis* 55, 1840–1874.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha?, *Review of Financial Studies* 30, 1316–1338.

Barillas, Francisco, and Jay Shanken, 2018, Comparing asset pricing models, *Journal of Finance* 73, 715–754.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael C. Jensen, ed., *Studies in the Theory of Capital Markets*, 79–124 (Praeger Publishers Inc., New York).

Chen, Long, Zhi Da, and Xinlei Zhao, 2013, What drives stock price movements?, *Review of Financial Studies* 26, 841–876.

Chen, Long, and Xinlei Zhao, 2009, Return decomposition, *Review of Financial Studies* 22, 5213–5249.

Chen, Zhuo, Bibo Liu, Huijun Wang, Zhengwei Wang, and Jianfeng Yu, 2023, Characteristics-based factors, Unpublished working paper.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Cooper, Ilan, Liang Ma, Paulo Maio, and Dennis Philip, 2021, Multifactor models and their consistency with the APT, *Review of Asset Pricing Studies* 11, 402–444.

Cooper, Michael, Huseyin Gulen, and Mihai Ion, 2024, The use of asset growth in empirical asset pricing models, *Journal of Financial Economics* 151.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020, The cross-section of risk and returns, *Review of Financial Studies* 33, 1927–1979.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389–406.

Detzel, Andrew, Robert Novy-Marx, and Mihail Velikov, 2023, Model comparison with transaction costs, *Journal of Finance* 78, 1743–1775.

Ehsani, Sina, and Juhani T. Linnainmaa, 2022a, Correcting asset pricing models, Unpublished working paper.

Ehsani, Sina, and Juhani T. Linnainmaa, 2022b, Factor momentum and the momentum factor, *Journal of Finance* 77, 1877–1919.

Ehsani, Sina, and Juhani T. Linnainmaa, 2022c, Time-series efficient factors, Unpublished working paper.

Eisfeldt, Andrea L., Edward T. Kim, and Dimitris Papanikolaou, 2022, Intangible value, *Critical Finance Review* 11, 299–332.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and Kenneth R. French, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

Fama, Eugene F., and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Fama, Eugene F., and Kenneth R. French, 2020, Comparing cross-section and time-series factor models, *Review of Financial Studies* 33, 1891–1926.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Frazzini, Andrea, and Lasse H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gerakos, Joseph, and Juhani T. Linnainmaa, 2018, Decomposing value, *Review of Financial Studies* 31, 1825–1854.

Golubov, Andrey, and Theodosia Konstantinidi, 2019, Where is the risk in value? Evidence from a market-to-book decomposition, *Journal of Finance* 74, 3135–3186.

Gonçalves, Andrei S., and Gregory Leonard, 2023, The fundamental-to-market ratio and the value premium decline, *Journal of Financial Economics* 147, 382–405.

Goyal, Amit, and Sunil Wahal, 2023, R&D, expected profitability, and expected returns, Unpublished working paper.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Hollstein, Fabian, and Marcel Prokopczuk, 2022, Testing factor models in the cross-section, *Journal of Banking and Finance* 145.

Hou, Kewei, and Mathijs A. van Dijk, 2019, Resurrecting the size effect: Firm size, profitability shocks, and expected stock returns, *Review of Financial Studies* 32, 2850–2889.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019–2133.

Jagannathan, Ravi, Robert A. Korajczyk, and Kai Wang, 2023, Intangibles investment and asset quality, Unpublished working paper.

Jegadeesh, Narasimhan, Joonki Noh, Kuntara Pukthuanthong, Richard Roll, and Junbo Wang, 2019, Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation, *Journal of Financial Economics* 133, 273–298.

Kim, Soohun, and Georgios Skoulakis, 2018, Ex-post risk premia estimation and asset pricing tests using large cross sections: The regression-calibration approach, *Journal of Econometrics* 204, 159–188.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175 – 194.

Lin, Hsiou-wei, and Maureen F. McNichols, 1998, Underwriting relationships, analysts' earnings forecasts and investment recommendations, *Journal of Accounting and Economics* 25, 101–127.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

McNichols, Maureen F., and Patricia C. O'Brien, 1997, Self-selection and analyst coverage, *Journal of Accounting Research* 35, 167–199.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam, 2019, A protocol for factor identification, *Review of Financial Studies* 32, 1573–1607.

Raponi, Valentina, Cesare Robotti, and Paolo Zaffaroni, 2019, Testing beta-pricing models using large cross-sections, *Review of Financial Studies* 33, 2796–2842.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–16.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Shumway, Tyler, and Vincent A. Warther, 1999, The delisting bias in CRSP's Nasdaq data and its implications for the size effect, *Journal of Finance* 54, 2361–2379.

Titman, Sheridan, K.C. John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns?, *Journal of Finance* 57, 233–264.

Xing, Yuhang, 2008, Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Table 1
Profitability Shock Estimation

This table displays time-series averages of regression coefficients from the cross-sectional profitability model of Hou and van Dijk (2019). The regressions are estimated at the end of each June from 1964 to 2019 using common US stocks traded on the NYSE, AMEX, or NASDAQ with total assets above \$10 million and book equity above \$5 million. The dependent variable is operating income-to-total assets. The independent variables are the market-to-book value of assets (FV/AT), a dummy variable that equals one if the firm does not pay dividends (DD), the dividend-to-book equity ratio (D/BE), and operating income-to-total assets (OI/AT). The independent variables are lagged by one year relative to the dependent variable. The variables are measured at the end of June. R^2 is the average adjusted R-squared of the annual regressions. t-statistics are reported in parentheses and are based on Newey-West (1987) heteroskedasticity-robust standard errors with five lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Intercept	FV/AT	DD	D/BE	OI/AT	R^2
Coefficient	0.0155***	0.0064**	−0.0128***	0.0675***	0.7187***	0.613
	(7.37)	(2.14)	(−4.50)	(3.65)	(40.55)	

Table 2
Identification of the Characteristics' Pricing Information

This table displays time-series averages of regression coefficients from cross-sectional Fama-MacBeth (1973) regressions. The regressions are estimated at the end of each June from 1968 to 2019 (exception Panel D: 1964 to 2019) using all common US stocks traded on the NYSE, AMEX, or NASDAQ. The dependent variable in Panel A (C, E) is book-to-market (investment, operating profitability). In Panel B, the dependent variable is book-to-market's market equity-driven part, calculated as the fitted value from the regression in Panel A. In Panel D, the dependent variable is the profitability shock as calculated in Section 3.2. The independent variables are the log change in market equity (dME), the log change in investment (dINV), the profitability shock (PS), and the fitted profitability shock obtained from the regression in Panel D (PS-Fit). dME is the annual log change in the market equity used in the calculation of book-to-market. The variables are constructed as described in Appendix A, are measured at the end of June, and are winsorized at the 1% and 99% levels. The regressions are estimated with weighted least squares with the stocks' market capitalizations as weights. A subscript $t-l$ indicates that the respective variable is lagged by l years relative to the dependent variable. R^2 is the average adjusted R-squared of the annual regressions. t-statistics are reported in parentheses and are based on Newey-West (1987) heteroskedasticity-robust standard errors with five lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Market Equity-Driven Part of Book-to-Market

	Intercept	dME_t	dME_{t-1}	dME_{t-2}	dME_{t-3}	dME_{t-4}	R^2
Coefficient	−0.68***	−0.61***	−0.43***	−0.36***	−0.26***	−0.18***	0.488
	(−4.63)	(−8.90)	(−6.58)	(−5.71)	(−4.25)	(−3.24)	

Panel B: Orthogonalization of Book-to-Market's Market Equity-Driven Part to Profitability Shocks

	Intercept	PS_t	PS_{t-1}	PS_{t-2}	PS_{t-3}	PS_{t-4}	R^2
Coefficient	−0.14***	−1.43***	−1.02***	−0.90***	−0.92***	−0.53***	0.388
	(−3.20)	(−4.97)	(−5.51)	(−4.89)	(−3.51)	(−2.96)	

Panel C: Orthogonalization of Investment to Profitability Shocks

	Intercept	PS_t	PS_{t-1}	PS_{t-2}	PS_{t-3}	PS_{t-4}	R^2
Coefficient	0.09***	0.92***	0.29***	0.18***	0.17***	0.04**	0.213
	(9.89)	(11.29)	(9.68)	(6.40)	(6.14)	(1.96)	

Panel D: Regression of Profitability Shocks on Changes in Investment and Market Equity

	Intercept	$dINV_t$	dME_t	R^2
Coefficient	0.01*	0.09***	0.04***	0.298
	(1.85)	(7.03)	(10.40)	

Panel E: Orthogonalization of Operating Profitability to Fitted Profitability Shocks

	Intercept	$PS\text{-}Fit_t$	$PS\text{-}Fit_{t-1}$	$PS\text{-}Fit_{t-2}$	$PS\text{-}Fit_{t-3}$	$PS\text{-}Fit_{t-4}$	R^2
Coefficient	0.34***	0.68***	0.83***	0.74***	0.45**	0.19	0.404
	(18.39)	(4.05)	(5.09)	(3.77)	(2.03)	(1.16)	

Table 3
Return Prediction with Characteristics

This table displays time-series averages of regression coefficients from cross-sectional Fama-MacBeth (1973) regressions. The regressions are estimated at the end of each month from June 1968 to November 2019 using all common US stocks traded on the NYSE, AMEX, or NASDAQ. The dependent variable is the one-month ahead return in excess of the one-month T-bill rate. The independent variables are log market equity (ME), log book-to-market (BM), operating profitability (OP), and log investment (INV); the adjusted book-to-market (BM^*), operating profitability (OP^*), and investment (INV^*); and the canceled parts of book-to-market (BM^d), operating profitability (OP^d), and investment (INV^d) obtained by subtracting the adjusted characteristics from the raw characteristics. The independent variables are constructed as described in Section 3 and Appendix A, are measured at the end of the most recent June (except ME, which is measured at the end of each month), and are winsorized at the 1% and 99% levels. The regressions are estimated with weighted least squares with stocks' market capitalizations as weights. R^2 is the average adjusted R-squared of the monthly regressions. t-statistics are reported in parentheses and are computed using Newey-West (1987) heteroskedasticity-robust standard errors with one lag. Boldface indicates significance at the 10% level.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Intercept	**0.0122**	**0.0102**	**0.0091**	**0.0097**	**0.0119**	**0.0096**	**0.0099**	**0.0093**	**0.0119**	**0.0103**	**0.0090**	**0.0096**	**0.0123**	**0.0106**	**0.0096**	**0.0096**	**0.0120**
	(4.99)	(5.70)	(4.95)	(5.60)	(4.89)	(5.35)	(5.67)	(5.31)	(4.90)	(5.99)	(4.92)	(5.52)	(5.03)	(5.88)	(5.28)	(5.48)	(4.93)
ME	**−0.0015**				**−0.0012**				**−0.0012**				**−0.0013**				**−0.0012**
	(−2.07)				(−1.81)				(−1.73)				(−1.82)				(−1.69)
BM		0.0007			0.0005												
		(1.32)			(0.84)												
OP			0.0005		**0.0010**												
			(1.06)		(1.67)												
INV				**−0.0010**	−0.0005												
				(−2.27)	(−1.55)												
BM^d						0.0001			−0.0001					0.0001			0.0000
						(0.30)			(−0.30)					(0.29)			(−0.07)
OP^d							−0.0005		**−0.0012**						−0.0006		−0.0001
							(−0.92)		(−2.16)						(−1.10)		(−0.19)
INV^d								0.0005	**0.0012**							0.0005	0.0004
								(0.80)	(2.40)							(0.93)	(0.76)
BM^*										**0.0020**			**0.0015**	**0.0019**			**0.0012**
										(3.16)			(2.37)	(2.94)			(1.80)
OP^*											0.0007		0.0007		**0.0007**		0.0005
											(1.58)		(1.57)		(1.68)		(0.86)
INV^*												**−0.0012**	−0.0004			**−0.0012**	−0.0001
												(−3.21)	(−1.36)			(−3.27)	(−0.35)
R^2	0.103	0.110	0.095	0.093	0.143	0.102	0.097	0.099	0.142	0.101	0.092	0.090	0.130	0.120	0.105	0.106	0.165

Table 4
Summary Statistics of Factors

Panel A of this table displays monthly mean returns (in percent), volatilities (in percent), and Sharpe ratios for the standard Fama-French (2015) market (MP), size (SMB), value (HML), profitability (RMW), and investment (CMA) factors. It also displays the correlations between the factors' monthly returns. Panel B displays the same statistics for the standard market factor (MP) as well as the informative size (SMB*), value (HML*), profitability (RMW*), and investment (CMA*) factors. Panel B also compares the informative factors to the corresponding Fama-French (2015) factors: "Diff" is the difference between the factors' mean returns; "Corr" is the correlation between the factors' monthly returns; "dSR" is the difference between the factors' Sharpe ratios. The sample period is from July 1968 to December 2019. t-statistics are reported in parentheses. Standard errors of Sharpe ratio differences are calculated based on the delta method. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Fama-French Factors

				Correlations			
Factor	Mean	Std	SR	SMB	HML	RMW	CMA
MP	0.53***	4.49	0.12	0.26	−0.27	−0.26	−0.39
	(2.93)						
SMB	0.15	2.97	0.05		−0.10	−0.37	−0.06
	(1.22)						
HML	0.31***	2.81	0.11			0.15	0.68
	(2.76)						
RMW	0.26***	2.28	0.12				−0.01
	(2.88)						
CMA	0.25***	1.80	0.14				
	(3.51)						

Panel B: Informative Factors

				Correlations				Comparison to Fama-French		
Factor	Mean	Std	SR	SMB*	HML*	RMW*	CMA*	Diff	Corr	dSR
MP	0.53***	4.49	0.12	0.21	−0.17	−0.22	−0.39	0.00	1.00	0.00
	(2.93)							(0.00)		(0.00)
SMB*	0.26**	2.91	0.09		0.18	−0.38	0.02	0.11***	0.97	0.04***
	(2.21)							(4.07)		(4.23)
HML*	0.34***	2.11	0.16			−0.04	0.51	0.03	0.67	0.05
	(4.05)							(0.37)		(1.56)
RMW*	0.28***	1.60	0.18				0.00	0.02	0.76	0.06**
	(4.40)							(0.34)		(2.15)
CMA*	0.28***	1.42	0.20					0.02	0.75	0.05*
	(4.86)							(0.46)		(1.89)

Table 5
Pricing Factors

This table displays results from factor model regressions. In Panel A, the Fama-French (2015) five-factor model is used to explain the informative factors. In Panel B, the informative five-factor model is used to explain the Fama-French (2015) factors. In Panel C, the Hou-Xue-Zhang (2015) four-factor model, consisting of market (rMKT), size (rME), profitability (rROE), and investment (rIA) factors, is used to explain the informative factors. In Panel D, the Fama-French (2015) five-factor model with an intangible value factor (HMLi) and a cash profitability factor (RMWc) is used to explain the informative factors. The sample period is from July 1968 to December 2019 in Panels A, B, and C, and from July 1976 to December 2019 in Panel D. α is in percent. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Pricing Informative Factors with Fama-French Factors

Factor	α	β^{MP}	β^{SMB}	β^{HML}	β^{RMW}	β^{CMA}	R^2
SMB^*	0.06***	0.00	0.97***	0.13***	0.02**	0.04**	0.966
	(2.63)	(0.57)	(121.76)	(11.73)	(2.00)	(2.39)	
HML^*	0.14**	0.01	0.11***	0.40***	−0.04	0.28***	0.514
	(2.21)	(0.61)	(5.03)	(13.55)	(−1.42)	(5.82)	
RMW^*	0.21***	−0.02**	−0.06***	−0.11***	0.51***	−0.01	0.620
	(4.94)	(−2.01)	(−4.33)	(−5.56)	(25.78)	(−0.31)	
CMA^*	0.12***	−0.03***	0.04***	−0.07***	0.09***	0.64***	0.597
	(3.23)	(−3.28)	(2.83)	(−4.14)	(5.11)	(21.71)	

Panel B: Pricing Fama-French Factors with Informative Factors

Factor	α	β^{MP}	β^{SMB^*}	β^{HML^*}	β^{RMW^*}	β^{CMA^*}	R^2
SMB	−0.08***	0.02***	0.99***	−0.07***	−0.02	−0.05**	0.952
	(−2.76)	(3.08)	(98.81)	(−4.56)	(−1.27)	(−2.20)	
HML	0.10	−0.09***	−0.07**	0.82***	−0.17***	0.16**	0.489
	(1.15)	(−4.35)	(−2.25)	(17.92)	(−3.05)	(2.20)	
RMW	−0.04	−0.03*	−0.03	−0.07**	1.03***	0.19***	0.591
	(−0.56)	(−1.92)	(−1.40)	(−2.00)	(25.61)	(3.65)	
CMA	0.07	−0.06***	−0.03*	0.24***	−0.19***	0.69***	0.651
	(1.58)	(−5.66)	(−1.89)	(10.03)	(−6.31)	(18.19)	

Panel C: Pricing Informative Factors with Hou-Xue-Zhang Factors

Factor	α	β^{rMKT}	β^{rME}	β^{rROE}	β^{rIA}	R^2
SMB^*	0.07**	−0.02**	0.91***	−0.08***	0.11***	0.933
	(2.28)	(−2.57)	(84.13)	(−6.48)	(6.34)	
HML^*	0.17**	−0.02	0.11***	−0.16***	0.63***	0.360
	(2.37)	(−1.12)	(4.42)	(−5.40)	(15.82)	
RMW^*	0.23***	−0.04***	−0.13***	0.27***	−0.10***	0.349
	(4.07)	(−2.93)	(−7.14)	(12.19)	(−3.36)	
CMA^*	0.10**	−0.05***	0.05***	0.01	0.49***	0.517
	(2.38)	(−5.48)	(3.23)	(0.74)	(21.32)	

Panel D: Pricing Informative Factors with Intangible Value and Cash Profitability Factors

Factor	α	β^{MP}	β^{SMB}	β^{HMLi}	β^{RMWc}	β^{CMA}	R^2
SMB^*	0.03	−0.01	0.96***	0.08***	0.01	0.09***	0.968
	(1.42)	(−1.50)	(106.12)	(6.62)	(0.97)	(5.82)	
HML^*	0.06	0.00	0.08***	0.41***	−0.06	0.30***	0.484
	(0.89)	(−0.10)	(3.18)	(12.43)	(−1.30)	(6.45)	
RMW^*	0.15***	0.00	−0.03*	0.05**	0.58***	−0.28***	0.577
	(3.22)	(0.23)	(−1.81)	(2.30)	(19.24)	(−8.46)	
CMA^*	0.08**	−0.01	0.08***	−0.10***	0.22***	0.59***	0.655
	(2.07)	(−0.63)	(5.45)	(−5.57)	(9.45)	(23.32)	

Table 6
Maximum Sharpe Ratios

Panel A of this table displays results on the maximum Sharpe ratios of the Fama-French (2015) and informative five-factor models as well as five-factor models using four of the Fama-French (2015) factors and one of the informative factors. "SR^2" is the maximum monthly squared Sharpe ratio across the sample period from July 1968 to December 2019. "BKRS" is the test statistic from testing whether the models' maximum squared Sharpe ratios are equal to the Fama-French (2015) model's maximum squared Sharpe ratio. The remaining columns display mean and median maximum squared Sharpe ratios from 100,000 full-sample, in-sample, and out-of-sample bootstrap simulation runs. A full-sample simulation run calculates the models' maximum squared Sharpe ratios from a randomly drawn sample of 618 months from the 618 months between July 1968 and December 2019. For the in- and out-of-sample simulations, the 618 months between July 1968 and December 2019 are split into 309 pairs of adjacent months, from which 309 pairs are randomly drawn with replacement. The in-sample simulation randomly selects one month of each pair to calculate the models' maximum squared Sharpe ratios and the factors' weights in the models' tangency portfolios. The out-of-sample simulation calculates the models' maximum squared Sharpe ratios based on (1) the factors' tangency weights from the in-sample months and (2) the factors' mean returns and covariance matrices from the unused months of the adjacent months. Panel A also displays the percentage of simulation runs in which the models exhibit higher maximum squared Sharpe ratios than the Fama-French (2015) model. Panel B displays the factors' weights in the models' tangency portfolios across the sample period from July 1968 to December 2019.

Panel A: Maximum Squared Sharpe Ratios

			Full-Sample			In-Sample			Out-of-Sample		
Model	SR^2	BKRS	Mean	Median	Percent	Mean	Median	Percent	Mean	Median	Percent
Fama-French	0.093		0.135	0.133		0.151	0.145		0.113	0.107	
Fama-French + HML^*	0.098	0.643	0.137	0.134	0.596	0.152	0.147	0.548	0.113	0.108	0.517
Fama-French + RMW^*	0.140	2.593	0.211	0.207	1.000	0.228	0.221	0.991	0.185	0.179	0.990
Fama-French + CMA^*	0.110	1.225	0.148	0.145	0.819	0.163	0.157	0.745	0.126	0.120	0.758
Informative	0.155	2.631	0.210	0.206	0.996	0.226	0.220	0.963	0.184	0.177	0.966

Panel B: Tangency Weights

Model	MP	SMB	HML	RMW	CMA
Fama-French	0.176	0.086	−0.030	0.304	0.465
Fama-French + HML^*	0.173	0.071	0.138	0.290	0.329
Fama-French + RMW^*	0.128	0.098	0.023	0.431	0.320
Fama-French + CMA^*	0.164	0.062	0.042	0.212	0.520
Informative	0.121	0.085	0.059	0.357	0.378

Table 7
Spanning Regressions

Panel A of this table displays the results from spanning regressions aiming to explain each of the five Fama-French (2015) factors based on the respective other four factors. Panel B displays the same results for the informative factors. The sample period is from July 1968 to December 2019. α is in percent. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Fama-French Factors

Factor	α	β^{MP}	β^{SMB}	β^{HML}	β^{RMW}	β^{CMA}	R^2
MP	0.85***		0.23***	0.11	−0.44***	−1.07***	0.240
	(5.26)		(4.03)	(1.40)	(−5.72)	(−8.84)	
SMB	0.20*	0.11***		0.00	−0.43***	−0.01	0.159
	(1.76)	(4.03)		(0.08)	(−8.27)	(−0.11)	
HML	−0.04	0.03	0.00		0.22***	1.09***	0.485
	(−0.46)	(1.40)	(0.08)		(5.48)	(22.12)	
RMW	0.39***	−0.12***	−0.24***	0.22***		−0.38***	0.213
	(4.67)	(−5.72)	(−8.27)	(5.48)		(−5.97)	
CMA	0.22***	−0.11***	0.00	0.41***	−0.14***		0.533
	(4.35)	(−8.84)	(−0.11)	(22.12)	(−5.97)		

Panel B: Informative Factors

Factor	α	β^{MP}	β^{SMB^*}	β^{HML^*}	β^{RMW^*}	β^{CMA^*}	R^2
MP	0.94***		0.25***	0.00	−0.45***	−1.26***	0.220
	(5.62)		(4.08)	(0.01)	(−4.15)	(−9.55)	
SMB*	0.29**	0.11***		0.28***	−0.60***	−0.03	0.191
	(2.54)	(4.08)		(4.76)	(−8.88)	(−0.31)	
HML*	0.09	0.00	0.13***		0.03	0.75***	0.286
	(1.21)	(0.01)	(4.76)		(0.58)	(13.59)	
RMW*	0.38***	−0.06***	−0.19***	0.02		−0.08	0.161
	(6.18)	(−4.15)	(−8.88)	(0.58)		(−1.52)	
CMA*	0.24***	−0.10***	−0.01	0.31***	−0.05		0.357
	(5.00)	(−9.55)	(−0.31)	(13.59)	(−1.52)		

Table 8
Factor Risk Prices

This table displays average annualized risk price estimates (in percent) from monthly cross-sectional Fama-MacBeth (1973) regressions. The regressions are estimated at the end of each month from June 1969 to December 2019 using all common US stocks traded on the NYSE, AMEX, or NASDAQ. The dependent variable is the compounded return across the previous 12 months in excess of the compounded one-month T-bill rate. The independent variables are a constant and betas on the factors of the Fama-French (2015) five-factor model (Panel A) and the informative five-factor model (Panel B). Betas are estimated at the end of each month from multivariate time-series regressions that regress stocks' daily excess returns across the previous 12 months on the models' factors. We require at least 100 daily observations to estimate the betas. Variables are in each month winsorized at the 1% and 99% levels. Rows labeled "UV" display risk price estimates from univariate Fama-MacBeth (1973) regressions using only the constant and one of the betas as explanatory variables; Rows labeled "MV" display risk price estimates from multivariate Fama-MacBeth (1973) regressions using the constant and all betas as explanatory variables. "Method" displays the estimation method; "WLS" estimates the regressions using weighted least squares with stocks' market capitalizations as weights; "IV" estimates the regressions using the instrumental variables approach described in Appendix B, which is also implemented using weighted least squares with stocks' market capitalizations as weights. Panel C displays the differences between the informative and Fama-French (2015) factors' risk price estimates. t-statistics are reported in parentheses and are computed using Newey-West (1987) heteroskedasticity-robust standard errors with 12 lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Risk Prices of Fama-French Factors

Setting	Method	γ^{ZB}	γ^{MP}	γ^{SMB}	γ^{HML}	γ^{RMW}	γ^{CMA}	R^2
UV	WLS		7.91***	5.96***	−3.05***	−2.22	−0.86	
			(3.38)	(3.81)	(−2.59)	(−1.40)	(−0.86)	
UV	IV		9.65***	7.90***	−3.43*	−1.02	−0.44	
			(3.03)	(3.99)	(−1.93)	(−0.80)	(−0.34)	
MV	WLS	4.41***	9.36***	5.71***	−3.98***	−1.92	−0.95	0.322
		(2.70)	(3.57)	(3.90)	(−2.63)	(−1.30)	(−1.04)	
MV	IV	3.67*	8.96***	7.37***	−4.32***	−0.19	−0.02	0.215
		(1.88)	(3.40)	(4.14)	(−2.80)	(−0.14)	(−0.02)	

Panel B: Risk Prices of Informative Factors

Setting	Method	γ^{ZB}	γ^{MP}	γ^{SMB^*}	γ^{HML^*}	γ^{RMW^*}	γ^{CMA^*}	R^2
UV	WLS		11.84***	7.56***	−0.61	1.55	1.04	
			(3.70)	(4.03)	(−0.57)	(1.32)	(1.28)	
UV	IV		13.14***	8.53***	−1.39	−0.20	1.91	
			(3.79)	(4.56)	(−0.73)	(−0.14)	(1.37)	
MV	WLS	2.95	10.87***	5.06***	0.11	0.37	1.37**	0.308
		(1.47)	(3.62)	(3.45)	(0.10)	(0.43)	(2.31)	
MV	IV	0.81	12.13***	6.94***	0.17	2.57**	3.96***	0.206
		(0.37)	(4.14)	(3.87)	(0.11)	(2.31)	(3.81)	

Panel C: Differences in Risk Prices

Setting	Method	γ^{ZB}	γ^{MP}	γ^{SMB^*}	γ^{HML^*}	γ^{RMW^*}	γ^{CMA^*}	R^2
UV	WLS		3.93	1.60	2.44**	3.77*	1.90***	
			(1.53)	(1.33)	(2.05)	(1.66)	(3.07)	
UV	IV		2.75	0.88	2.00	0.14	2.23*	
			(1.12)	(0.63)	(1.63)	(0.12)	(1.70)	
MV	WLS	−1.45	1.51	−0.66*	4.09***	2.29	2.32***	−0.015
		(−1.35)	(1.41)	(−1.80)	(2.63)	(1.63)	(3.69)	
MV	IV	−2.86	2.64	−0.31	4.37***	2.34*	4.04***	−0.009
		(−1.28)	(1.35)	(−0.42)	(3.03)	(1.84)	(3.92)	

Table 9
Predictive Power of Factor Betas

This table displays average annualized coefficients (in percent) from monthly cross-sectional Fama-MacBeth (1973) regressions. The regressions are estimated at the end of each month from June 1969 to November 2019 using all common US stocks traded on the NYSE, AMEX, or NASDAQ. The dependent variables are the compounded returns across the next one, three, six, or 12 months in excess of the compounded one-month T-bill rate. The independent variables are a constant and betas on the factors of the Fama-French (2015) five-factor model (Panel A) and the informative five-factor model (Panel B). Betas are estimated at the end of each month from multivariate time-series regressions that regress stocks' daily excess returns across the previous 12 months on the models' factors. We require at least 100 daily observations to estimate the betas. Betas are in each month winsorized at the 1% and 99% levels. Rows labeled "UV" display coefficients from univariate Fama-MacBeth (1973) regressions using only the constant and one of the betas as explanatory variables; Rows labeled "MV" display coefficients from multivariate Fama-MacBeth (1973) regressions using the constant and all betas as explanatory variables. "Horizon" displays the return horizon. The regressions are estimated with weighted least squares with stocks' market capitalizations as weights. t-statistics are reported in parentheses and are computed using Newey-West (1987) heteroskedasticity-robust standard errors with lags equal to the return horizon. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Betas on Fama-French Factors

Setting	Horizon	Intercept	β^{MP}	β^{SMB}	β^{HML}	β^{RMW}	β^{CMA}	R^2
UV	1		−0.10	1.42	1.01	1.94**	0.05	
			(−0.05)	(1.32)	(1.10)	(2.09)	(0.06)	
UV	3		−0.08	1.49	1.13	1.88**	−0.18	
			(−0.05)	(1.56)	(1.36)	(2.37)	(−0.29)	
UV	6		−0.50	1.37	0.99	1.77**	−0.12	
			(−0.33)	(1.49)	(1.23)	(2.28)	(−0.21)	
UV	12		−0.90	1.40	0.78	1.49**	−0.14	
			(−0.67)	(1.47)	(1.01)	(2.11)	(−0.24)	
MV	1	6.67***	−0.14	1.11	1.03	1.83**	0.36	0.197
		(4.29)	(−0.08)	(1.03)	(0.90)	(2.00)	(0.44)	
MV	3	6.77***	−0.30	1.16	1.12	1.67**	0.18	0.228
		(4.68)	(−0.18)	(1.22)	(1.06)	(2.13)	(0.24)	
MV	6	7.25***	−0.76	1.18	1.09	1.50**	0.18	0.225
		(4.97)	(−0.50)	(1.25)	(1.06)	(1.96)	(0.25)	
MV	12	7.93***	−1.19	1.25	0.91	1.21*	0.22	0.219
		(5.01)	(−0.85)	(1.25)	(0.89)	(1.65)	(0.31)	

Panel B: Betas on Informative Factors

Setting	Horizon	Intercept	β^{MP}	β^{SMB^*}	β^{HML^*}	β^{RMW^*}	β^{CMA^*}	R^2
UV	1		−1.32	1.32	1.43	1.18	0.52	
			(−0.58)	(1.02)	(1.50)	(1.45)	(0.71)	
UV	3		−1.35	1.37	1.17	1.06	0.32	
			(−0.66)	(1.25)	(1.39)	(1.38)	(0.52)	
UV	6		−1.67	1.42	0.91	1.02	0.15	
			(−0.86)	(1.30)	(1.17)	(1.29)	(0.28)	
UV	12		−1.82	1.66	0.61	1.07	−0.15	
			(−1.04)	(1.48)	(0.86)	(1.41)	(−0.26)	
MV	1	7.04***	−0.49	1.43	1.58*	1.65**	1.13*	0.191
		(4.32)	(−0.25)	(1.32)	(1.69)	(2.24)	(1.73)	
MV	3	6.99***	−0.49	1.50	1.46*	1.48**	0.87	0.221
		(4.56)	(−0.27)	(1.61)	(1.79)	(2.24)	(1.63)	
MV	6	7.35***	−0.84	1.60*	1.26*	1.35**	0.66	0.218
		(4.71)	(−0.47)	(1.71)	(1.71)	(2.07)	(1.45)	
MV	12	7.97***	−1.23	1.72*	0.93	1.37**	0.37	0.212
		(4.75)	(−0.74)	(1.67)	(1.39)	(2.16)	(0.74)	

Table 10
Comparison to Hedging Procedure

Panel A of this table displays results for the hedged market ($\mathrm{MP^H}$), size ($\mathrm{SMB^H}$), value ($\mathrm{HML^H}$), profitability ($\mathrm{RMW^H}$), and investment ($\mathrm{CMA^H}$) factors. Mean returns, volatilities, and alphas are in percent. α^{SR} denotes the intercepts from spanning regressions. The panel also compares the informative factors to the corresponding hedged factors: "Diff" is the difference between the factors' mean returns; "Corr" is the factors' correlation; "dSR" is the difference between the factors' Sharpe ratios; α^H (α^*) denotes the intercepts from regressing the respective hedged (informative) factor on the informative (hedged) five-factor model. Panel B displays the same results for the informative hedged market ($\mathrm{MP^{H*}}$), size ($\mathrm{SMB^{H*}}$), value ($\mathrm{HML^{H*}}$), profitability ($\mathrm{RMW^{H*}}$), and investment ($\mathrm{CMA^{H*}}$) factors and compares them to the hedged factors. α^{H*} (α^H) denotes intercepts from regressing the respective informative hedged (standard hedged) factor on the standard hedged (informative hedged) five-factor model. Panel C displays results on the maximum squared Sharpe ratios of the hedged, informative, and informative hedged models. "SR^2" is the maximum monthly squared Sharpe ratio across the sample period. "BKRS" is the test statistic from testing whether the models' maximum squared Sharpe ratios are equal to the hedged model's maximum squared Sharpe ratio. The remaining columns display mean and median maximum squared Sharpe ratios from 100,000 bootstrap simulation runs conducted as described in Table 6. The panel also displays the percentage of simulation runs in which the models exhibit higher maximum squared Sharpe ratios than the hedged model. Panel D displays average annualized risk price estimates (in percent) for the hedged and informative hedged factors. The risk prices are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 8. The panel also displays the differences between the informative and informative hedged factors' risk price estimates and the hedged factors' risk price estimates. Panel E displays average annualized coefficients (in percent) for the betas on the hedged and informative hedged factors. The coefficients are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 9. The sample period is from July 1968 to December 2019. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Hedged Factors

					Correlations				Comparison to Informative				
Factor	Mean	Std	SR	α^{SR}	$\mathrm{SMB^H}$	$\mathrm{HML^H}$	$\mathrm{RMW^H}$	$\mathrm{CMA^H}$	Diff	Corr	dSR	α^H	α^*
$\mathrm{MP^H}$	0.52***	3.03	0.17	0.72***	−0.29	−0.18	0.17	−0.26	0.01	0.68	−0.06	0.19**	0.06
	(4.29)			(6.08)					(0.05)		(−1.57)	(2.09)	(0.45)
$\mathrm{SMB^H}$	0.12	1.95	0.06	0.30***		0.15	−0.31	0.14	0.13*	0.74	0.03	0.07	0.08
	(1.58)			(3.91)					(1.71)		(0.89)	(1.29)	(0.99)
$\mathrm{HML^H}$	0.25***	1.75	0.14	0.09			−0.48	0.67	0.09	0.42	0.02	0.27***	0.21**
	(3.57)			(1.59)					(1.09)		(0.45)	(4.26)	(2.53)
$\mathrm{RMW^H}$	0.16***	1.55	0.10	0.33***				−0.49	0.12**	0.62	0.07**	0.01	0.22***
	(2.59)			(6.23)					(2.22)		(2.03)	(0.11)	(4.18)
$\mathrm{CMA^H}$	0.25***	1.28	0.19	0.21***					0.03	0.43	0.00	0.21***	0.14**
	(4.76)			(5.51)					(0.55)		(0.09)	(4.51)	(2.53)

Panel B: Informative Hedged Factors

					Correlations				Comparison to Hedged				
Factor	Mean	Std	SR	α^{SR}	$\mathrm{SMB^{H*}}$	$\mathrm{HML^{H*}}$	$\mathrm{RMW^{H*}}$	$\mathrm{CMA^{H*}}$	Diff	Corr	dSR	α^{H*}	α^H
$\mathrm{MP^{H*}}$	0.56***	3.10	0.18	0.86***	−0.32	−0.31	0.13	−0.19	0.04	0.93	0.01	0.00	−0.02
	(4.50)			(7.01)					(0.86)		(0.52)	(−0.05)	(−0.49)
$\mathrm{SMB^{H*}}$	0.21***	2.03	0.11	0.39***		0.26	−0.39	0.06	0.09**	0.89	0.04**	0.07*	−0.06
	(2.63)			(5.04)					(2.38)		(2.17)	(1.85)	(−1.53)
$\mathrm{HML^{H*}}$	0.29***	1.42	0.21	0.24***			−0.24	0.44	0.04	0.28	0.06	0.29***	0.21***
	(5.13)			(4.52)					(0.53)		(1.34)	(5.19)	(3.07)
$\mathrm{RMW^{H*}}$	0.23***	1.22	0.19	0.29***				0.04	0.07	0.68	0.08**	0.15***	0.04
	(4.66)			(6.07)					(1.47)		(2.56)	(3.93)	(0.75)
$\mathrm{CMA^{H*}}$	0.25***	1.06	0.24	0.14***					0.01	0.48	0.05	0.11***	0.15***
	(5.92)			(3.33)					(0.13)		(1.12)	(2.91)	(3.22)

- Continued on next page -

Panel C: Maximum Squared Sharpe Ratios

Model	SR^2	BKRS	Full-Sample Mean	Full-Sample Median	Full-Sample Percent	In-Sample Mean	In-Sample Median	In-Sample Percent	Out-of-Sample Mean	Out-of-Sample Median	Out-of-Sample Percent
Hedged	0.171		0.215	0.211		0.235	0.226		0.190	0.182	
Informative	0.155	−0.434	0.210	0.206	0.456	0.226	0.220	0.456	0.184	0.177	0.466
Informative Hedged	0.231	1.692	0.244	0.242	0.950	0.257	0.252	0.868	0.224	0.220	0.878

Panel D: Risk Prices

Model	Method	γ^{ZB}	$\gamma^{MP^{\mathrm{H}}}$	$\gamma^{SMB^{\mathrm{H}}}$	$\gamma^{HML^{\mathrm{H}}}$	$\gamma^{RMW^{\mathrm{H}}}$	$\gamma^{CMA^{\mathrm{H}}}$	R^2
Hedged	WLS	5.05***	5.48***	0.71	−1.18	−2.13*	1.03	0.288
		(2.98)	(3.27)	(0.67)	(−1.24)	(−1.82)	(1.49)	
Hedged	IV	0.13	8.78***	3.31**	−2.51*	−1.46	0.67	0.189
		(0.05)	(3.20)	(2.43)	(−1.92)	(−1.33)	(0.55)	
Informative Hedged	WLS	6.10***	5.47***	0.13	1.82	−0.75	1.04	0.274
		(3.53)	(2.92)	(0.10)	(1.32)	(−0.90)	(1.62)	
Informative Hedged	IV	5.02	7.41***	1.31	1.42	−1.43	3.29***	0.183
		(1.45)	(2.85)	(0.89)	(1.12)	(−1.25)	(3.13)	
Inf. vs. Hedged	WLS	−2.09	5.39*	4.35***	1.30	2.50**	0.34	0.020
		(−1.11)	(1.81)	(4.06)	(1.08)	(2.08)	(0.44)	
Inf. vs. Hedged	IV	0.67	2.17	3.49***	2.71	4.47***	2.82**	0.017
		(0.22)	(0.75)	(2.67)	(1.45)	(3.72)	(2.20)	
Inf. Hedged vs. Hedged	WLS	1.05	−0.01	−0.58	3.01**	1.39	0.01	−0.013
		(1.45)	(−0.02)	(−1.16)	(1.98)	(1.20)	(0.01)	
Inf. Hedged vs. Hedged	IV	4.89	−0.93	−1.83*	3.63**	0.13	3.27**	−0.006
		(1.56)	(−0.46)	(−1.80)	(1.96)	(0.13)	(2.14)	

Panel E: Betas' Predictive Power

Model	Horizon	Intercept	$\beta^{MP^{\mathrm{H}}}$	$\beta^{SMB^{\mathrm{H}}}$	$\beta^{HML^{\mathrm{H}}}$	$\beta^{RMW^{\mathrm{H}}}$	$\beta^{CMA^{\mathrm{H}}}$	R^2
Hedged	1	6.00***	1.23	−0.12	−0.01	0.81	0.54	0.173
		(3.20)	(0.93)	(−0.15)	(−0.01)	(1.16)	(0.95)	
Hedged	12	7.26***	0.40	0.44	0.45	0.17	0.61	0.195
		(3.99)	(0.40)	(0.66)	(0.63)	(0.31)	(1.16)	
Informative Hedged	1	7.57***	0.34	0.21	0.68	0.33	0.78	0.167
		(4.08)	(0.24)	(0.23)	(0.86)	(0.53)	(1.39)	
Informative Hedged	12	7.77***	0.45	0.53	0.32	0.26	0.08	0.192
		(4.48)	(0.37)	(0.63)	(0.45)	(0.54)	(0.19)	

Table 11
Comparison to Cross-Section Procedure

Panel A of this table displays results for the cross-section size (SMB^{CS}), value (HML^{CS}), profitability (RMW^{CS}), and investment (CMA^{CS}) factors. Mean returns, volatilities, and alphas are in percent. α^{SR} denotes the intercepts from spanning regressions. The panel also compares the informative factors to the corresponding cross-section factors: "Diff" is the difference between the factors' mean returns; "Corr" is the factors' correlation; "dSR" is the difference between the factors' Sharpe ratios; α^{CS} (α^{*}) denotes the intercepts from regressing the respective cross-section (informative) factor on the informative (cross-section) five-factor model. Panel B displays the same results for the informative cross-section size ($\text{SMB}^{\text{CS*}}$), value ($\text{HML}^{\text{CS*}}$), profitability ($\text{RMW}^{\text{CS*}}$), and investment ($\text{CMA}^{\text{CS*}}$) factors and compares them to the cross-section factors. α^{CS*} (α^{CS}) denotes intercepts from regressing the respective informative cross-section (standard cross-section) factor on the standard cross-section (informative cross-section) five-factor model. Panel C displays results on the maximum squared Sharpe ratios of the cross-section, informative, and informative cross-section models. "SR^2" is the maximum monthly squared Sharpe ratio across the sample period. "BKRS" is the test statistic from testing whether the models' maximum squared Sharpe ratios are equal to the cross-section model's maximum squared Sharpe ratio. The remaining columns display mean and median maximum squared Sharpe ratios from 100,000 bootstrap simulation runs conducted as described in Table 6. The panel also displays the percentage of simulation runs in which the models exhibit higher maximum squared Sharpe ratios than the cross-section model. Panel D displays average annualized risk price estimates (in percent) for the cross-section and informative cross-section factors. The risk prices are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 8. The panel also displays the differences between the informative and informative cross-section factors' risk price estimates and the cross-section factors' risk price estimates. Panel E displays average annualized coefficients (in percent) for the betas on the cross-section and informative cross-section factors. The coefficients are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 9. The sample period is from July 1968 to December 2019. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Cross-Section Factors

					Correlations				Comparison to Informative				
Factor	Mean	Std	SR	α^{SR}	SMB^{CS}	HML^{CS}	RMW^{CS}	CMA^{CS}	Diff	Corr	dSR	α^{CS}	α^{*}
MP	0.53***	4.49	0.12	0.91***	0.27	−0.27	−0.27	−0.45	0.00	1.00	0.00	0.00	0.00
	(2.93)			(5.79)					(0.00)		(0.00)	(0.00)	(0.00)
SMB^{CS}	0.08	1.50	0.05	0.08		−0.35	−0.29	−0.10	0.18***	0.94	0.04***	−0.03	0.03
	(1.29)			(1.32)					(2.84)		(2.71)	(−1.57)	(1.22)
HML^{CS}	0.09**	1.02	0.09	0.00			0.71	0.16	0.25***	0.59	0.07*	0.03	0.12*
	(2.26)			(−0.07)					(3.63)		(1.93)	(0.84)	(1.93)
RMW^{CS}	0.10***	0.73	0.14	0.07***				0.12	0.18***	0.50	0.03	0.01	0.19***
	(3.54)			(3.41)					(3.21)		(0.81)	(0.33)	(4.27)
CMA^{CS}	0.07***	0.41	0.17	0.09***					0.21***	0.62	0.03	0.03**	0.13***
	(4.22)			(6.01)					(4.27)		(0.76)	(2.16)	(2.90)

Panel B: Informative Cross-Section Factors

					Correlations				Comparison to Cross-Section				
Factor	Mean	Std	SR	α^{SR}	$\text{SMB}^{\text{CS*}}$	$\text{HML}^{\text{CS*}}$	$\text{RMW}^{\text{CS*}}$	$\text{CMA}^{\text{CS*}}$	Diff	Corr	dSR	α^{CS*}	α^{CS}
MP	0.53***	4.49	0.12	0.90***	0.21	−0.14	−0.12	−0.33	0.00	1.00	0.00	0.00	0.00
	(2.93)			(5.46)					(0.00)		(0.00)	(0.00)	(0.00)
$\text{SMB}^{\text{CS*}}$	0.14**	1.45	0.10	0.12**		−0.06	0.01	−0.10	0.07***	0.94	0.05***	0.03**	−0.02
	(2.49)			(1.97)					(3.37)		(3.59)	(1.99)	(−1.20)
$\text{HML}^{\text{CS*}}$	0.06***	0.60	0.11	0.11***			−0.06	−0.13	−0.03	0.60	0.02	0.01	0.03
	(2.65)			(4.64)					(−0.88)		(0.42)	(0.76)	(0.77)
$\text{RMW}^{\text{CS*}}$	0.08***	0.51	0.17	0.13***				−0.28	−0.02	0.53	0.02	0.05***	0.01
	(4.16)			(6.48)					(−0.79)		(0.61)	(3.48)	(0.28)
$\text{CMA}^{\text{CS*}}$	0.03**	0.36	0.09	0.08***					−0.04***	0.55	−0.08**	0.01	0.03***
	(2.14)			(5.98)					(−2.67)		(−2.23)	(0.85)	(2.59)

- Continued on next page -

Panel C: Maximum Squared Sharpe Ratios

Model	SR^2	BKRS	Full-Sample Mean	Full-Sample Median	Full-Sample Percent	In-Sample Mean	In-Sample Median	In-Sample Percent	Out-of-Sample Mean	Out-of-Sample Median	Out-of-Sample Percent
Cross-Section	0.113		0.147	0.145		0.163	0.157		0.123	0.118	
Informative	0.155	1.661	0.210	0.206	0.977	0.226	0.220	0.913	0.184	0.177	0.921
Informative CS	0.135	0.937	0.147	0.144	0.832	0.159	0.154	0.752	0.128	0.123	0.747

Panel D: Risk Prices

Model	Method	γ^{ZB}	γ^{MP}	$\gamma^{SMB^{CS}}$	$\gamma^{HML^{CS}}$	$\gamma^{RMW^{CS}}$	$\gamma^{CMA^{CS}}$	R^2
Cross-Section	WLS	4.12**	9.63***	2.70***	−1.48***	−0.86	−0.28	0.320
		(2.52)	(3.68)	(3.78)	(−2.73)	(−1.59)	(−1.40)	
Cross-Section	IV	2.73	9.94***	3.40***	−1.44***	−0.10	0.12	0.213
		(1.33)	(3.63)	(3.82)	(−2.85)	(−0.21)	(0.44)	
Informative CS	WLS	2.62	11.20***	2.53***	−0.07	0.38	0.06	0.308
		(1.28)	(3.69)	(3.61)	(−0.23)	(1.38)	(0.34)	
Informative CS	IV	1.08	11.41***	3.41***	−0.76*	0.68*	0.77**	0.206
		(0.46)	(3.76)	(3.73)	(−1.72)	(1.83)	(2.45)	
Inf. vs. CS	WLS	−1.17	1.23	2.35***	1.59	1.23	1.65***	−0.012
		(−1.09)	(1.17)	(2.91)	(1.39)	(1.33)	(3.36)	
Inf. vs. CS	IV	−1.92	1.81	3.64***	1.62	2.42**	3.73***	−0.007
		(−0.86)	(0.91)	(3.49)	(1.16)	(2.31)	(3.89)	
Inf. CS vs. CS	WLS	−1.50	1.56	−0.18	1.40**	1.24**	0.35**	−0.012
		(−1.36)	(1.44)	(−0.68)	(2.31)	(2.32)	(2.01)	
Inf. CS vs. CS	IV	−1.65	1.22	0.08	0.70	0.62	0.68**	−0.007
		(−0.72)	(0.63)	(0.18)	(1.28)	(1.26)	(2.12)	

Panel E: Betas' Predictive Power

Model	Horizon	Intercept	β^{MP}	$\beta^{SMB^{CS}}$	$\beta^{HML^{CS}}$	$\beta^{RMW^{CS}}$	$\beta^{CMA^{CS}}$	R^2
Cross-Section	1	6.47***	0.08	0.44	0.54	0.66**	−0.07	0.196
		(4.10)	(0.04)	(0.80)	(1.36)	(2.37)	(−0.36)	
Cross-Section	12	7.69***	−0.95	0.58	0.46	0.62***	−0.13	0.218
		(5.00)	(−0.68)	(1.20)	(1.33)	(2.82)	(−0.91)	
Informative CS	1	7.01***	−0.48	0.87	0.39	0.51**	0.04	0.191
		(4.27)	(−0.24)	(1.54)	(1.38)	(2.19)	(0.24)	
Informative CS	12	8.12***	−1.40	1.07**	0.21	0.57***	−0.11	0.212
		(4.86)	(−0.84)	(2.09)	(1.10)	(2.63)	(−0.75)	

Table 12
Comparison to Time-Series Efficiency Procedure

Panel A of this table displays results for the time-series efficient market ($\mathrm{MP^{TE}}$), size ($\mathrm{SMB^{TE}}$), value ($\mathrm{HML^{TE}}$), profitability ($\mathrm{RMW^{TE}}$), and investment ($\mathrm{CMA^{TE}}$) factors. Mean returns, volatilities, and alphas are in percent. α^{SR} denotes the intercepts from spanning regressions. The panel also compares the informative factors to the corresponding time-series efficient factors: "Diff" is the difference between the factors' mean returns; "Corr" is the factors' correlation; "dSR" is the difference between the factors' Sharpe ratios; α^{TE} (α^{*}) denotes the intercepts from regressing the respective time-series efficient (informative) factor on the informative (time-series efficient) five-factor model. Panel B displays the same results for the informative time-series efficient market ($\mathrm{MP^{TE*}}$), size ($\mathrm{SMB^{TE*}}$), value ($\mathrm{HML^{TE*}}$), profitability ($\mathrm{RMW^{TE*}}$), and investment ($\mathrm{CMA^{TE*}}$) factors and compares them to the time-series efficient factors. α^{TE*} (α^{TE}) denotes intercepts from regressing the respective informative time-series efficient (standard time-series efficient) factor on the standard time-series efficient (informative time-series efficient) five-factor model. Panel C displays results on the maximum squared Sharpe ratios of the time-series efficient, informative, and informative time-series efficient models. "SR^2" is the maximum monthly squared Sharpe ratio across the sample period. "BKRS" is the test statistic from testing whether the models' maximum squared Sharpe ratios are equal to the time-series efficient model's maximum squared Sharpe ratio. The remaining columns display mean and median maximum squared Sharpe ratios from 100,000 bootstrap simulation runs conducted as described in Table 6. The panel also displays the percentage of simulation runs in which the models exhibit higher maximum squared Sharpe ratios than the time-series efficient model. Panel D displays average annualized risk price estimates (in percent) for the time-series efficient and informative time-series efficient factors. The risk prices are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 8. The panel also displays the differences between the informative and informative time-series efficient factors' risk price estimates and the time-series efficient factors' risk price estimates. Panel E displays average annualized coefficients (in percent) for the betas on the time-series efficient and informative time-series efficient factors. The coefficients are estimated from multivariate cross-sectional Fama-MacBeth (1973) regressions as described in Table 9. The sample period is from July 1968 to December 2019. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Time-Series Efficient Factors

					Correlations				Comparison to Informative				
Factor	Mean	Std	SR	α^{SR}	$\mathrm{SMB^{TE}}$	$\mathrm{HML^{TE}}$	$\mathrm{RMW^{TE}}$	$\mathrm{CMA^{TE}}$	Diff	Corr	dSR	α^{TE}	α^{*}
$\mathrm{MP^{TE}}$	0.31***	2.61	0.12	0.50***	0.14	−0.22	−0.32	−0.23	0.22*	0.73	0.00	0.15*	0.36***
	(2.99)			(5.02)					(1.74)		(−0.04)	(1.96)	(2.85)
$\mathrm{SMB^{TE}}$	0.14**	1.82	0.08	0.14*		0.02	−0.16	0.00	0.11	0.75	0.01	0.07	0.11
	(1.96)			(1.87)					(1.45)		(0.35)	(1.45)	(1.39)
$\mathrm{HML^{TE}}$	0.33***	1.82	0.18	0.07			0.31	0.67	0.01	0.61	−0.02	0.20***	0.11
	(4.51)			(1.24)					(0.10)		(−0.59)	(3.26)	(1.61)
$\mathrm{RMW^{TE}}$	0.27***	1.49	0.18	0.27***				0.18	0.02	0.59	0.00	0.09*	0.20***
	(4.49)			(4.74)					(0.32)		(−0.03)	(1.85)	(3.91)
$\mathrm{CMA^{TE}}$	0.20***	1.19	0.17	0.09**					0.08*	0.63	0.03	0.09**	0.16***
	(4.21)			(2.35)					(1.65)		(0.74)	(2.49)	(3.47)

Panel B: Informative Time-Series Efficient Factors

					Correlations				Comparison to Time-Series Efficient				
Factor	Mean	Std	SR	α^{SR}	$\mathrm{SMB^{TE*}}$	$\mathrm{HML^{TE*}}$	$\mathrm{RMW^{TE*}}$	$\mathrm{CMA^{TE*}}$	Diff	Corr	dSR	α^{TE*}	α^{TE}
$\mathrm{MP^{TE*}}$	0.33***	3.13	0.11	0.69***	0.13	−0.20	−0.26	−0.40	0.02	0.94	−0.02	0.05	0.01
	(2.61)			(5.84)					(0.35)		(−1.05)	(1.26)	(0.37)
$\mathrm{SMB^{TE*}}$	0.17**	2.12	0.08	0.18**		0.16	−0.27	0.03	0.02	0.94	0.00	0.01	0.01
	(1.97)			(2.07)					(0.80)		(0.02)	(0.34)	(0.25)
$\mathrm{HML^{TE*}}$	0.29***	1.64	0.18	0.07			0.09	0.51	−0.04	0.67	−0.01	0.07	0.13**
	(4.37)			(1.17)					(−0.75)		(−0.18)	(1.44)	(2.36)
$\mathrm{RMW^{TE*}}$	0.24***	1.22	0.20	0.27***				0.13	−0.03	0.67	0.02	0.14***	0.06
	(4.97)			(5.66)					(−0.58)		(0.60)	(3.92)	(1.38)
$\mathrm{CMA^{TE*}}$	0.23***	1.06	0.22	0.18***					0.03	0.72	0.05*	0.12***	0.04
	(5.46)			(4.99)					(0.97)		(1.69)	(4.00)	(1.10)

- Continued on next page -

Panel C: Maximum Squared Sharpe Ratios

Model	SR^2	BKRS	Full-Sample Mean	Full-Sample Median	Full-Sample Percent	In-Sample Mean	In-Sample Median	In-Sample Percent	Out-of-Sample Mean	Out-of-Sample Median	Out-of-Sample Percent
Time-Series Efficient	0.110		0.139	0.137		0.153	0.148		0.115	0.110	
Informative	0.155	1.328	0.210	0.206	0.944	0.226	0.220	0.871	0.184	0.177	0.884
Informative TE	0.157	1.756	0.167	0.165	0.959	0.178	0.173	0.885	0.147	0.143	0.899

Panel D: Risk Prices

Model	Method	γ^{ZB}	$\gamma^{MP^{\mathrm{TE}}}$	$\gamma^{SMB^{\mathrm{TE}}}$	$\gamma^{HML^{\mathrm{TE}}}$	$\gamma^{RMW^{\mathrm{TE}}}$	$\gamma^{CMA^{\mathrm{TE}}}$	R^2
Time-Series Efficient	WLS	7.49***	1.83*	1.48**	−0.85**	−0.38	−0.37**	0.292
		(5.03)	(1.74)	(2.38)	(−2.09)	(−1.24)	(−1.97)	
Time-Series Efficient	IV	6.37**	2.60	1.80*	−1.02*	0.21	−0.60***	0.193
		(2.44)	(1.61)	(1.82)	(−1.75)	(0.41)	(−2.73)	
Informative TE	WLS	7.63***	3.52**	2.20**	0.32	0.20	−0.13	0.282
		(4.75)	(2.27)	(2.49)	(0.67)	(0.71)	(−0.57)	
Informative TE	IV	8.86***	2.99	2.59***	−0.15	0.40	0.33	0.187
		(3.19)	(1.54)	(2.67)	(−0.25)	(0.83)	(1.12)	
Inf. vs. TE	WLS	−4.54***	9.04***	3.58***	0.96	0.75	1.73***	0.015
		(−3.70)	(3.49)	(3.27)	(0.91)	(0.93)	(3.48)	
Inf. vs. TE	IV	−5.56**	9.05***	5.24***	1.03	2.45**	4.60***	0.013
		(−2.22)	(3.37)	(3.72)	(0.74)	(2.31)	(4.43)	
Inf. TE vs. TE	WLS	0.14	1.69	0.72	1.18**	0.58*	0.23*	−0.010
		(0.32)	(1.64)	(1.53)	(2.42)	(1.77)	(1.89)	
Inf. TE vs. TE	IV	2.50	0.69	1.02	0.83	0.54	0.98***	−0.006
		(0.99)	(0.41)	(1.52)	(1.36)	(0.99)	(3.03)	

Panel E: Betas' Predictive Power

Model	Horizon	Intercept	$\beta^{MP^{\mathrm{TE}}}$	$\beta^{SMB^{\mathrm{TE}}}$	$\beta^{HML^{\mathrm{TE}}}$	$\beta^{RMW^{\mathrm{TE}}}$	$\beta^{CMA^{\mathrm{TE}}}$	R^2
Time-Series Efficient	1	7.15***	0.97	0.44	0.33	0.74**	0.06	0.182
		(4.46)	(1.17)	(1.24)	(0.82)	(2.24)	(0.32)	
Time-Series Efficient	12	7.65***	0.63	0.44	0.07	0.50	−0.12	0.205
		(4.50)	(1.10)	(1.49)	(0.21)	(1.61)	(−0.69)	
Informative TE	1	7.01***	0.74	0.47	0.51	0.59**	0.43*	0.178
		(4.24)	(0.68)	(0.88)	(1.33)	(2.07)	(1.75)	
Informative TE	12	7.31***	0.25	0.47	0.25	0.35	0.00	0.201
		(4.41)	(0.33)	(1.15)	(1.26)	(1.35)	(0.01)	

Table 13
Pricing Characteristics-Sorted Portfolios

Panel A of this table displays results on the pricing performance of the Fama-French (2015) and informative five-factor models for 113 sets of characteristics-sorted decile portfolios from the global-q website of Kewei Hou, Chen Xue, and Lu Zhang. Panels B, C, and D display the results for the models' hedged, cross-section, and time-series efficient versions, respectively. The sample period is from July 1968 to December 2019. The metrics measuring the models' pricing performance are the portfolios' average absolute alpha (in percent); the fraction of the portfolios' alphas significant at the 5% level; the portfolios' average absolute alpha over the average absolute deviation of the portfolios' mean returns from their mean; the portfolios' cross-sectional R^2; and the portfolios' average time-series R^2. "N" is the number of portfolios.

Panel A: Standard Models

		Fama-French Model					Informative Model				
Category	N	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$
All Deciles	1130	0.108	0.260	1.026	−0.015	0.874	0.098	0.179	0.932	0.136	0.864
Value Deciles	240	0.075	0.163	0.767	0.491	0.867	0.069	0.050	0.703	0.543	0.847
Profitability Deciles	180	0.137	0.383	1.165	−0.222	0.888	0.122	0.306	1.038	−0.036	0.881
Investment Deciles	260	0.090	0.219	0.977	0.094	0.881	0.081	0.142	0.879	0.250	0.873
Momentum Deciles	210	0.134	0.367	1.264	−0.413	0.882	0.115	0.257	1.087	−0.051	0.877
Intangibles Deciles	150	0.136	0.307	1.171	−0.314	0.832	0.126	0.233	1.082	−0.110	0.824
Frictions Deciles	90	0.076	0.067	0.827	0.399	0.891	0.087	0.100	0.946	0.075	0.880

Panel B: Hedged Models

		Hedged Model					Informative Hedged Model				
Category	N	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$
All Deciles	1130	0.129	0.050	1.232	−0.260	0.429	0.133	0.028	1.271	0.066	0.413
Value Deciles	240	0.086	0.000	0.878	0.292	0.448	0.115	0.013	1.172	0.565	0.432
Profitability Deciles	180	0.151	0.089	1.284	−0.391	0.425	0.144	0.039	1.226	−0.183	0.406
Investment Deciles	260	0.117	0.008	1.261	0.148	0.428	0.130	0.008	1.410	0.157	0.411
Momentum Deciles	210	0.161	0.124	1.520	−1.080	0.426	0.139	0.048	1.319	−0.171	0.413
Intangibles Deciles	150	0.153	0.073	1.317	−0.536	0.415	0.145	0.047	1.250	−0.018	0.398
Frictions Deciles	90	0.125	0.011	1.349	0.174	0.421	0.135	0.033	1.460	−0.108	0.405

Panel C: Cross-Section Models

		Cross-Section Model					Informative Cross-Section Model				
Category	N	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$
All Deciles	1130	0.104	0.226	0.991	0.007	0.870	0.097	0.190	0.927	0.146	0.863
Value Deciles	240	0.066	0.050	0.667	0.527	0.861	0.065	0.063	0.662	0.584	0.844
Profitability Deciles	180	0.135	0.367	1.144	−0.205	0.886	0.123	0.317	1.044	−0.045	0.880
Investment Deciles	260	0.086	0.208	0.925	0.226	0.877	0.084	0.165	0.907	0.201	0.872
Momentum Deciles	210	0.136	0.376	1.283	−0.495	0.881	0.116	0.276	1.096	−0.038	0.878
Intangibles Deciles	150	0.132	0.253	1.141	−0.253	0.830	0.124	0.233	1.072	−0.093	0.825
Frictions Deciles	90	0.077	0.067	0.830	0.375	0.888	0.082	0.078	0.890	0.183	0.881

Panel D: Time-Series Efficient Models

		Time-Series Efficient Model					Informative Time-Series Efficient Model				
Category	N	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$	$A(\|\alpha\|)$	% Sig	$\frac{A(\|\alpha\|)}{A(\|\mu\|)}$	CSR^2	$A(R^2)$
All Deciles	1130	0.365	0.748	3.510	0.179	0.477	0.300	0.666	2.887	0.301	0.581
Value Deciles	240	0.354	0.746	3.671	0.524	0.446	0.302	0.663	3.127	0.573	0.556
Profitability Deciles	180	0.351	0.689	2.995	0.136	0.502	0.282	0.611	2.406	0.250	0.603
Investment Deciles	260	0.387	0.842	4.239	0.198	0.487	0.313	0.758	3.427	0.389	0.590
Momentum Deciles	210	0.355	0.738	3.385	0.123	0.489	0.292	0.662	2.782	0.309	0.593
Intangibles Deciles	150	0.363	0.653	3.132	−0.160	0.451	0.301	0.553	2.598	−0.068	0.551
Frictions Deciles	90	0.387	0.778	4.166	−0.030	0.501	0.316	0.722	3.401	0.061	0.599

Table A1
Characteristics-Sorted Portfolios

This table lists the characteristics underlying the characteristics-sorted decile portfolios from the global-q website of Kewei Hou, Chen Xue, and Lu Zhang.

Variable	Category	Description
beta1	Frictions	market beta, 1-month holding period
Dtv12	Frictions	dollar trading volume, 12-month holding period
Isff1	Frictions	idiosyncratic skewness estimated from the Fama-French 3-factor model, 1-month holding period
Isq1	Frictions	idiosyncratic skewness estimated from the q-factor model, 1-month holding period
Ivff1	Frictions	idiosyncratic volatility estimated from the Fama-French 3-factor model, 1-month holding period
Ivq1	Frictions	idiosyncratic volatility estimated from the q-factor model, 1-month holding period
Me	Frictions	market equity
Srev	Frictions	short-term reversal
Tv1	Frictions	total volatility, 1-month holding period
Eprd	Intangibles	earnings predictability
Etl	Intangibles	earnings timeliness
Etr	Intangibles	effective tax rate
Hs	Intangibles	industry concentration in sales
Ioca	Intangibles	industry-adjusted organizational capital-to-assets
Oca	Intangibles	organizational capital-to-assets
Ol	Intangibles	operating leverage
R[6,10]a	Intangibles	seasonality, average return across months t-72, t-84, t-96, t-108, t-120
R[6,10]n	Intangibles	seasonality, average return from month t-120 to t-61 except for months t-72, t-84, t-96, t-108, t-120
R[11,15]a	Intangibles	seasonality, average return across months t-132, t-144, t-156, t-168, t-180
R1a	Intangibles	seasonality, return in month t-12
R1n	Intangibles	seasonality, average return from month t-11 to t-1
R[16,20]a	Intangibles	seasonality, average return across months t-192, t-204, t-216, t-228, t-240
R[2,5]a	Intangibles	seasonality, average return across months t-24, t-36, t-48, t-60
R[2,5]n	Intangibles	seasonality, average return from month t-60 to t-13 except for months t-24, t-36, t-48, t-60
Aci	Investment	abnormal corporate investment
Cei	Investment	composite equity issuance
Dac	Investment	discretionary accruals
dBe	Investment	changes in book equity
dCoa	Investment	changes in current operating assets
dFin	Investment	changes in net financial assets
dFnl	Investment	changes in financial liabilities
dIi	Investment	percent changes in investment relative to industry
dLno	Investment	changes in long-term net operating assets
dNca	Investment	changes in non-current operating assets
dNco	Investment	changes in net non-current operating assets
dNoa	Investment	changes in net operating assets
dPia	Investment	changes in PPE and inventory scaled by lagged assets
dWc	Investment	changes in net non-cash working capital
I/A	Investment	investment-to-assets (asset growth)
2Ig	Investment	2-year investment growth
Ig	Investment	investment growth
Ivc	Investment	inventory changes
Ivg	Investment	inventory growth
Noa	Investment	net operating assets
Nsi	Investment	net stock issues
Oa	Investment	operating accruals
Pda	Investment	percent discretionary accruals
Poa	Investment	percent operating accruals
Pta	Investment	percent total accruals
Ta	Investment	total accruals

- Continued on next page -

Variable	Category	Description
Cim12	Momentum	customer industries momentum, 12-month holding period
Cim1	Momentum	customer industries momentum, 1-month holding period
Cim6	Momentum	customer industries momentum, 6-month holding period
52w12	Momentum	52-week high, 12-month holding period
52w6	Momentum	52-week high, 6-month holding period
R11.12	Momentum	prior 11-month returns, 12-month holding period
R11.1	Momentum	prior 11-month returns, 1-month holding period
R11.6	Momentum	prior 11-month returns, 6-month holding period
R6.12	Momentum	prior 6-month returns, 12-month holding period
R6.1	Momentum	prior 6-month returns, 1-month holding period
R6.6	Momentum	prior 6-month returns, 6-month holding period
Resid11.12	Momentum	11-month residual momentum, 12-month holding period
Resid11.1	Momentum	11-month residual momentum, 1-month holding period
Resid11.6	Momentum	11-month residual momentum, 6-month holding period
Resid6.12	Momentum	6-month residual momentum, 12-month holding period
Resid6.6	Momentum	6-month residual momentum, 6-month holding period
Rs1	Momentum	revenue surprises, 1-month holding period
Sim12	Momentum	supplier industries momentum, 12-month holding period
Sim1	Momentum	supplier industries momentum, 1-month holding period
Sue1	Momentum	standard unexpected earnings, 1-month holding period
Sue6	Momentum	standard unexpected earnings, 6-month holding period
Ato	Profitability	assets turnover
Cla	Profitability	cash-based operating profits-to-lagged assets
Cop	Profitability	operating cash flow-to-assets
Cto	Profitability	capital turnover
dRoe12	Profitability	4-quarter changes in return on equity, 12-month holding period
dRoe1	Profitability	4-quarter changes in return on equity, 1-month holding period
dRoe6	Profitability	4-quarter changes in return on equity, 6-month holding period
Eg12	Profitability	expected growth, 12-month holding period
Eg1	Profitability	expected growth, 1-month holding period
Eg6	Profitability	expected growth, 6-month holding period
Gpa	Profitability	gross profits-to-assets
Opa	Profitability	operating profits-to-assets
Ope	Profitability	operating profits-to-book equity
Roe1	Profitability	return on equity, 1-month holding period
Roe6	Profitability	return on equity, 6-month holding period
Sgq1	Profitability	quarterly sales growth, 1-month holding period
Tbiq12	Profitability	quarterly tax income-to-book income, 12-month holding period
Tbiq6	Profitability	quarterly tax income-to-book income, 6-month holding period
Bm	Value vs. Growth	book-to-market equity
Bmj	Value vs. Growth	book-to-June-end market equity
Bmq12	Value vs. Growth	quarterly book-to-market equity, 12-month holding period
Cp	Value vs. Growth	cash flow-to-price
Cpq12	Value vs. Growth	quarterly cash flow-to-price, 12-month holding period
Cpq1	Value vs. Growth	quarterly cash flow-to-price, 1-month holding period
Cpq6	Value vs. Growth	quarterly cash flow-to-price, 6-month holding period
Dp	Value vs. Growth	dividend yield
Dur	Value vs. Growth	equity duration
Ebp	Value vs. Growth	enterprise book-to-price
Em	Value vs. Growth	enterprise multiple
Ep	Value vs. Growth	earnings-to-price
Epq12	Value vs. Growth	quarterly earnings-to-price, 12-month holding period
Epq1	Value vs. Growth	quarterly earnings-to-price, 1-month holding period
Epq6	Value vs. Growth	quarterly earnings-to-price, 6-month holding period
Ir	Value vs. Growth	intangible return
Rev12	Value vs. Growth	long-term reversal, 12-month holding period
Rev1	Value vs. Growth	long-term reversal, 1-month holding period
Rev6	Value vs. Growth	long-term reversal, 6-month holding period
Sp	Value vs. Growth	sales-to-price
Spq12	Value vs. Growth	quarterly sales-to-price, 12-month holding period
Spq1	Value vs. Growth	quarterly sales-to-price, 1-month holding period
Spq6	Value vs. Growth	quarterly sales-to-price, 6-month holding period
Vhp	Value vs. Growth	ROE-based intrinsic value-to-market

Informative Value, Profitability, and Investment Factors: Internet Appendix

Manuel Ammann[*] Tobias Hemauer[†] Simon Straumann[‡]

December 28, 2023

Abstract

This Internet Appendix contains supplementary analyses and results for "Informative Value, Profitability, and Investment Factors."

[*]University of St.Gallen, School of Finance. Address: Unterer Graben 21, CH-9000 St. Gallen, Switzerland. Email: manuel.ammann@unisg.ch

[†]University of St.Gallen, School of Finance. Address: Unterer Graben 21, CH-9000 St. Gallen, Switzerland. Email: tobias.hemauer@unisg.ch

[‡]WHU - Otto Beisheim School of Management, Finance and Accounting Group. Address: Burgplatz 2, D-56179 Vallendar, Germany. Email: simon.straumann@whu.edu

G Factor Protocol

This section evaluates whether our informative factors pass the two additional criteria—besides exhibiting positive risk prices—of the factor protocol to identify genuine risk factors proposed by Pukthuanthong et al. (2019). The first additional criterion is that a genuine factor must be related to the covariance matrix of returns. Following Pukthuanthong et al. (2019), we examine this criterion by determining whether our factors are significantly related to the principal components of the return covariance matrix. For this purpose, we split our sample period into five decades (the first decade is from July 1968 to June 1978, the second decade is from July 1978 to June 1988, and so forth). Within each of the five decades, we extract the first ten eigenvectors of the matrix $(1/T)RR'$, where R is the TxN matrix containing the de-meaned monthly returns of all N stocks with non-missing data across the T=120 months in the decade. Connor and Korajczyk (1988) show that, for large N, the eigenvectors of this matrix are asymptotically equivalent to the principal components of the return covariance matrix. Then, we calculate the canonical correlations between the ten principal components and our five informative factors. Given that we use ten principal components but only five factors, there are five pairs of canonical variates and thus five canonical correlations within each decade. For each of the five canonical pairs, we compute the principal components' canonical variates as the weighted averages of the ten principal components. The principal components' weights are chosen to maximize the correlation between the respective canonical variates of the principal components and factors and to make the respective canonical variates orthogonal to the other canonical variates. The principal components' five canonical variates are then regressed on the five factors. In total, we conduct 25 such regressions (five regressions per decade).

The significance of the factors in explaining the return covariance matrix can be assessed based on their absolute t-statistics in these regressions. The first row of Panel B of Table IA1 displays our informative factors' average absolute t-statistics across the 25 regressions; the second row displays the average absolute t-statistics across those regressions for which the respective canonical correlation is significant at the 5% level. The next rows display the number of coefficients that are significant at the 5% level in each decade (the maximum is five) and the average across the decades. Pukthuanthong et al. (2019) deem a factor to be materially related to the return covariance matrix (i) if its average absolute t-statistic across those regressions for which the respective canonical correlation is significant exceeds 1.96 and (ii) if the average percentage of significant t-statistics across the decades is larger than 25% (i.e., in our case, if the average number of significant t-statistics is larger than 1.25). All of our informative factors pass these thresholds, indicating that they are materially related to the return covariance matrix.

For comparison, Panel A displays the same results for the standard Fama-French (2015) factors. The standard investment factor marginally fails to pass the thresholds as its average absolute t-statistic is below 1.96. This result indicates that the standard investment factor does not qualify as a valid risk factor candidate. By contrast, the other factors are all materially related to the return covariance matrix.

The second additional criterion suggested by Pukthuanthong et al. (2019) is that a genuine risk factor must generate a reasonable reward-to-risk ratio. This criterion is evaluated based on the Sharpe ratio of a portfolio that combines a long position in the market with a hedge portfolio that goes long stocks with high betas on the factor and short stocks with low betas on the factor. For this purpose, at the end of each month from June 1969 to November 2019, we estimate stocks' factor betas by regressing their daily excess returns across the previous 12 months on our informative factors.[1] Stocks are then sorted into deciles according to the betas using NYSE breakpoints. The portfolios are held for one month, and the stocks in the portfolios are value-weighted. We go long the top decile and short the bottom decile and combine this zero net investment portfolio with a long position in the value-weighted market portfolio.

Panel B of Table IA1 presents the mean, volatility, and Sharpe ratio of this strategy for each factor. Following Pukthuanthong et al. (2019), the table also displays the t-statistic from testing whether the monthly Sharpe ratio exceeds the bound of $0.6/\sqrt{12} \approx 0.173$ (corresponding to a bound on the annual Sharpe ratio of 0.6 as recommended by MacKinlay (1995)). The results show that the Sharpe ratios are below this bound for all factors, in part even significantly. The same applies when we go long the top 30% and short the bottom 30% of stocks. Thus, the factors generate risk-to-reward ratios that are consistent with risk-based pricing and therefore also pass this criterion. Panel A shows that the same holds for the standard Fama-French (2015) factors.

Overall, the results in Table IA1 show that all of our informative factors pass the two additional criteria of Pukthuanthong et al.'s (2019) factor protocol and thus qualify as potential risk factors. Except for the standard investment factor, the same also holds for the standard Fama-French (2015) factors. Hence, the main difference between the standard factors and our informative factors in passing the factor protocol emanates from requiring positive risk prices. As shown in Table 8, our informative factors, except our informative value factor, also pass this criterion. In contrast, the standard value, profitability, and investment factors fail to pass this criterion and can therefore not be considered valid risk factor candidates.

[1] We require at least 100 daily observations across the 12-month estimation window to estimate a stock's factor betas.

References

Connor, Gregory, and Robert A. Korajczyk, 1988, Risk and return in an equilibrium APT: Application of a new test methodology, *Journal of Financial Economics* 21, 255–289.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3–28.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam, 2019, A protocol for factor identification, *Review of Financial Studies* 32, 1573–1607.

Table IA1
Factor Protocol

This table displays results on the standard Fama-French (2015) (Panel A) and informative (Panel B) factors' association with the return covariance matrix and their risk-to-reward ratios. "Mean T-Stats" displays the average absolute t-statistics of the coefficients obtained from regressing the canonical variates of the covariance matrix's asymptotical principal components on the factors. "Mean T-Stats Sig Corr" displays the average absolute t-statistics from those regressions for which the respective canonical correlation is significant. "Decade x" displays the number of coefficients that are significant at the 5% level in the regressions in decade x. "Average" displays the average number of significant coefficients across the decades. The first decade is from July 1968 to June 1978, the second decade is from July 1978 to June 1988, and so forth. In each decade, there are five canonical correlations and thus five regressions; therefore, there are in total 25 regressions. The table further displays monthly means, volatilities, and Sharpe ratios of strategies that combine long positions in the market portfolio with long-short portfolios. For each factor, the long-short portfolio is formed by going long the top 10% (30%) of stocks with the highest betas on the factor and short the bottom 10% (30%) of stocks with the lowest betas on the factor. The portfolios are reformed at the end of each month from June 1969 to November 2019, are based on betas estimated across the previous 12 months, use NYSE breakpoints, are held for one month, and are value-weighted. Means and volatilities are in percent. t-statistics are reported in parentheses. Sharpe ratios are tested against a threshold of $0.6/\sqrt{12}$ using the delta method. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Fama-French Factors

	MP	SMB	HML	RMW	CMA
Mean T-Stats	10.96	7.11	3.95	2.51	1.90
Mean T-Stats Sig Corr	11.40	7.40	4.08	2.58	1.92
Decade 1	3.00	3.00	4.00	3.00	2.00
Decade 2	3.00	3.00	2.00	2.00	0.00
Decade 3	2.00	4.00	5.00	3.00	4.00
Decade 4	3.00	4.00	4.00	3.00	1.00
Decade 5	3.00	4.00	3.00	2.00	2.00
Average	2.80	3.60	3.60	2.60	1.80
Mean^{10}	0.77*	0.77*	0.74**	0.97***	0.59**
	(1.88)	(1.92)	(2.44)	(4.04)	(2.49)
Std^{10}	10.21	9.98	7.48	5.94	5.85
SR^{10}	0.08**	0.08**	0.10*	0.16	0.10*
	(−2.41)	(−2.39)	(−1.85)	(−0.25)	(−1.79)
Mean^{30}	0.66**	0.66**	0.67***	0.72***	0.65***
	(2.10)	(2.07)	(2.77)	(3.78)	(3.38)
Std^{30}	7.86	7.92	6.02	4.75	4.79
SR^{30}	0.08**	0.08**	0.11	0.15	0.14
	(−2.18)	(−2.22)	(−1.52)	(−0.50)	(−0.88)

Panel B: Informative Factors

	MP	SMB^*	HML^*	RMW^*	CMA^*
Mean T-Stats	10.36	6.13	2.97	2.94	2.31
Mean T-Stats Sig Corr	12.81	7.36	3.37	3.23	2.58
Decade 1	3.00	3.00	4.00	4.00	4.00
Decade 2	2.00	3.00	3.00	4.00	2.00
Decade 3	3.00	3.00	3.00	4.00	3.00
Decade 4	3.00	4.00	3.00	3.00	2.00
Decade 5	3.00	3.00	3.00	3.00	2.00
Average	2.80	3.20	3.20	3.60	2.60
Mean^{10}	0.50	0.76*	0.76***	0.96***	0.83***
	(1.09)	(1.77)	(2.59)	(3.88)	(3.60)
Std^{10}	11.28	10.63	7.32	6.14	5.73
SR^{10}	0.04***	0.07**	0.10*	0.16	0.15
	(−3.20)	(−2.55)	(−1.68)	(−0.41)	(−0.68)
Mean^{30}	0.55	0.67**	0.75***	0.71***	0.64***
	(1.59)	(1.97)	(3.37)	(3.52)	(3.44)
Std^{30}	8.65	8.47	5.56	4.99	4.64
SR^{30}	0.06***	0.08**	0.14	0.14	0.14
	(−2.69)	(−2.32)	(−0.91)	(−0.76)	(−0.83)