Comparing Factor Models with Price-Impact Costs*

Sicong Li Victor DeMiguel Alberto Martín-Utrera

November 8, 2023

Abstract

Prominent asset-pricing models include factors constructed using time-varying characteristics, and thus, their implied investment-opportunity set requires investors to execute sizable trades whenever characteristics change. The price impact of these trades affects the portfolio choices of large investors, and therefore, the overall achievable investment-opportunity set. We propose a formal statistical test to compare factor models with price impact and show that model performance depends on the investor's absolute risk aversion. Empirically, we find that the q-factor model, the Fama-French six-factor model, and a high-dimensional model best span the investment opportunities of investors with high, medium, and low absolute risk aversion, respectively.

Keywords: trading costs, mean-variance utility, statistical test.

*Li, London Business School, e-mail: ali@london.edu, DeMiguel, London Business School, e-mail: avmiguel@london.edu; Martín-Utrera, Mendoza College of Business at the University of Notre Dame, email: amart224@nd.edu and Iowa State University, e-mail: amutrera@iastate.edu. We are grateful for comments from Pedro Barroso, Svetlana Bryzgalova, Andrew Chen, Andrew Detzel, Raymond Kan, Nicos Savva, Stephen Schaefer, Raman Uppal, Mihail Velikov, Guofu Zhou, Paolo Zaffaroni and seminar participants at AFA PhD poster session, FMA Annual Meeting, INFORMS Annual Meeting, International Young Finance Scholars' Conference, and Trans-Atlantic Doctoral Conference.

1 Introduction

Prominent asset-pricing models include factors constructed using time-varying firm characteristics such as profitability and momentum. The investment opportunity set implied by these models requires investors to execute sizable trades whenever the conditioning information in firm characteristics changes. For the large institutional investors that manage most of the capital in financial markets, the price impact of these trades affects their optimal portfolio choices, and thus, it also affects the overall achievable investment opportunity set. We propose a formal statistical test to compare asset-pricing models in the presence of price impact. In contrast to the cases without transaction costs and with proportional costs, we show that in the presence of price-impact costs different models may be best at spanning the investment opportunities of different investors depending on their absolute risk aversion. Empirically, we find that the q-factor model, the Fama-French six-factor model, and a high-dimensional model are best at spanning the investment opportunities of investors with high, medium, and low absolute-risk aversion, respectively.

A popular approach to compare asset-pricing models is the GRS test of Gibbons, Ross, and Shanken (1989), which evaluates the ability of the factors in a model to span the investment opportunity set generated by certain test assets. Specifically, the GRS statistic is a quadratic form of the time-series intercept (alpha) obtained from the regression of the test-asset returns on the factor returns. Gibbons et al. (1989) show that this quadratic form measures the squared Sharpe ratio improvement that an investor can achieve by having access to the test assets, in addition to the factors in the model. Moreover, Barillas and Shanken (2017) show that the asset-pricing model whose factors generate the highest squared Sharpe ratio is also the model that best spans the investment opportunity set. Thus, test assets are irrelevant and it suffices to compare factor models in terms of their squared Sharpe ratio.

Detzel, Novy-Marx, and Velikov (2023), however, point out that one has to account for *trading costs* when comparing factor models. In particular, they explain that the framework underpinning these models, the arbitrage pricing theory (APT) of Ross (1976), relies on the assumption that investment opportunities that deliver abnormal returns attract arbitrage capital until such opportunities vanish. However, arbitrageurs allocate capital only to investment opportunities that are profitable after trading costs, and thus, Detzel et al. (2023) propose comparing factor models in terms of their squared Sharpe ratio of returns net of *proportional* transaction costs, which measures the ability of the factor model to span the *achievable* investment opportunity set.

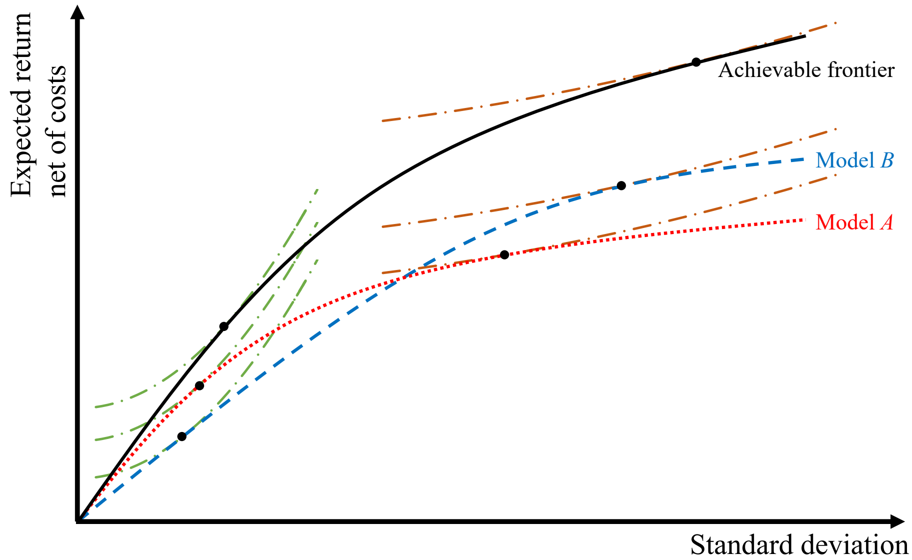
Proportional transaction costs capture the trading costs of retail investors, but price-impact costs are more relevant for the large institutional investors that manage most of the capital in financial markets. For instance, Gârleanu and Pedersen (2022) show that institutional investors held around 50% of the US equity market in 2017, and Edelen, Evans, and Kadlec (2007) show that price-impact costs represent 65% of the total trading costs of mutual funds, whereas proportional (bid-ask spread) costs represent only 17%.

Despite the importance of price impact for the large investors that dominate financial markets, price impact would not affect the achievable investment opportunity set if large investors did not have to trade or they had to execute only small trades. However, prominent asset-pricing models include factors constructed using firm characteristics such as profitability and momentum that vary substantially over time. These characteristics encapsulate conditioning information that investors optimally exploit when choosing their portfolios (Cochrane, 2009, p. 134). As a result, the investment opportunity set implied by these factor models requires investors to execute *sizable* trades to reach the equilibrium at regular intervals—whenever the conditioning information in firm characteristics varies. For large investors, the price impact of these sizable trades affects their optimal portfolio choices, and thus, price impact affects the overall achievable investment opportunity set, which includes the optimal portfolio of every investor. In this paper, we propose a methodological framework to compare factor models in terms of their ability to span the achievable investment opportunity set in the presence of price impact.

Our contribution to the literature is threefold. Our first contribution is to propose comparing factor models in terms of *mean-variance utility net of price-impact costs* and to show that different models may be better at spanning the investment opportunities of investors with different absolute risk aversion. In particular, we prove that the achievable efficient frontier in the presence of price impact is strictly concave, and thus, the squared Sharpe ratio criterion is no longer sufficient to compare factor models because each efficient portfolio has a different Sharpe ratio of returns net of price-impact costs. Moreover, the objective of investors is not to maximize Sharpe ratio, but rather their utility of returns net of price-impact costs, which is therefore the economically meaningful criterion to compare

Figure 1: Achievable efficient frontier and frontiers spanned by two factor models

This figure illustrates the achievable efficient frontier in the presence of price impact (black solid line) as well as the efficient frontiers spanned by the factors in models A (red dotted line) and B (blue dashed line). The figure also depicts the indifference curves of an investor with low absolute risk aversion (brown dash-dotted lines) and an investor with high absolute risk aversion (green dash-dotted lines).



the ability of different factor models to span the achievable investment opportunity set. We show that our proposed criterion is equivalent to the squared Sharpe ratio in the cases without costs and with proportional costs. In addition, we generalize the result of Gibbons et al. (1989) to show that the increase in the mean-variance utility net of price-impact costs of an investor when she has access to a set of test assets in addition to the factors in a model is a quadratic form of the alpha (net for price impact). Finally, we also generalize the result of Barillas and Shanken (2017) to show that test assets are irrelevant for model comparison *also* in the presence of price impact.

Our first contribution is illustrated in Figure 1, which depicts the achievable efficient frontier (black solid line) as well as the efficient frontiers spanned by the factors in models A (red dotted line) and B (blue dashed line) in the presence of price impact. Each portfolio in the achievable frontier maximizes the mean-variance utility net of price-impact costs of investors with a particular absolute risk aversion, which can be defined as the ratio of the investor's relative risk aversion to her endowment (Gârleanu and Pedersen, 2013). Intuitively, larger investors have lower absolute risk aversion, and thus, they are willing to take on larger investment positions to maximize their net mean return at the expense of higher return variance. The figure depicts the indifference curves of an investor with low absolute risk aversion (brown dash-dotted lines) and an investor with high absolute risk aversion (green dash-dotted lines). Each investor's optimal portfolio is at the tangent between the investor's indifference curve and the efficient frontier. The figure shows that model B spans better the investment opportunities of the low-absolute-risk-aversion investor and model A those of the high-absolute-risk-aversion investor. This is because the low-absolute-risk-aversion investor is willing to take on larger investment positions that incur higher price-impact costs. Because the price-impact costs from exploiting the factors in model B are much lower than those from exploiting the factors in model A, model B is better at spanning the achievable investment opportunities of the low-absolute-risk-aversion investor. On the other hand, model A is better at spanning the investment opportunities of the high-absolute-risk-aversion investor because she takes smaller investment positions that incur lower price-impact costs, and the factors of model A offer a better risk-return tradeoff when price-impact costs are lower.

Our second contribution is to develop a statistical methodology to test the significance of the difference between the mean-variance utilities net of price-impact costs of two factor models. In particular, we derive two asymptotic distributions that allow us to compare two factor models for the cases when they are nested or non-nested. Our approach extends the tests of Kan and Robotti (2009) and Barillas, Kan, Robotti, and Shanken (2020) to compare factor models with price impact. We also develop closed-form expressions for the variance of the asymptotic distribution and use them to show that it is easier to reject the null hypothesis that the mean-variance utilities net of price-impact costs of two models are equal not only when the mean-variance portfolio returns of the two models are positively correlated as shown by Barillas et al. (2020) for the case without trading costs, but also when the mean-variance portfolio return of each model is highly correlated with the rebalancing trades of the portfolio of the other model, and when the rebalancing trades of the two portfolios are highly correlated.

Our third contribution is to use our statistical test to compare the empirical performance of six factor models. We consider five prominent low-dimensional models: the CAPM model of Sharpe (1964) and Lintner (1965), the q-factor model of Hou, Xue, and Zhang (2015), HXZ4, the four-factor model of Fama and French (1993) and Carhart (1997), FFC4, the five-factor model of Fama and French (2015), FF5, and the six-factor model of Fama and French (2018), FF6. In addition, DeMiguel, Martin-Utrera, Nogales, and Uppal (2020) show that trading costs provide an economic rationale to consider high-dimensional factor models. In particular, they show that combining factors helps to reduce transaction costs because the trades required to rebalance different factor portfolios often cancel out, a phenomenon they term *trading diversification*. Moreover, they show that the benefits from trading diversification increase with the number of factors combined. For this reason, we consider a sixth factor model containing the 20 factors that DeMiguel et al. (2020) find statistically significant in the presence of price-impact costs, DMNU20.

We highlight two empirical findings. First, in the presence of price impact, model performance depends not only on the portfolio turnover required to trade the factors in the model, as pointed out by Detzel et al. (2023) for the case with proportional costs, but also on the liquidity of the stocks traded. In particular, we find that, compared to their FF6 counterparts, the HXZ4 investment and profitability factors not only involve higher portfolio turnover, but also require trading stocks with lower market capitalization, which are more illiquid and subject to higher price-impact costs. As a result, while in the absence of trading costs the four-factor model of Hou et al. (2015) outperforms the six-factor model of Fama and French (2018), in the presence of price-impact costs the six-factor model of Fama and French (2018) tends to perform better.

Second, the relative performance of factor models in the presence of price impact depends on the absolute risk aversion of the investor. For instance, the high-dimensional model of DeMiguel et al. (2020) significantly outperforms the low-dimensional models *only* when spanning the investment opportunities of large (low-absolute-risk-aversion) investors. This is because high-dimensional models provide larger trading-diversification benefits, and thus, they outperform low-dimensional models at spanning the investment opportunities of large investors for whom price-impact costs are relatively more important. Overall, accounting for price impact results in a nuanced comparison of the factor models we consider—the q-factor model of Hou et al. (2015), the six factor model of Fama and French (2018), and the high-dimensional model of DeMiguel et al. (2020) are best at spanning the investment opportunities of investors with high, medium, and low absolute risk aversion, respectively.

As a robustness check, we use the bootstrap test of Fama and French (2018) and Detzel et al. (2023) to show that the out-of-sample performance of the different models is consistent with the empirical findings from our statistical tests. We also show that our empirical findings are robust to considering factors constructed using the *banding* transaction-cost mitigation strategy used by Detzel et al. (2023).

An implication of our work is that different benchmark factor models should be used to evaluate the performance of investment strategies designed for different investors, depending on their absolute risk aversion. While the q-factor model may be appropriate to evaluate the performance of investment strategies designed for small investors with high absolute risk aversion, the six-factor model of Fama and French (2018) may be a better benchmark for strategies designed for large investors with low risk aversion, and a high-dimensional model may be appropriate only for the largest investors. Our proposed statistical test can be used not only to compare factor models, but also to evaluate the significance of the increase in mean-variance utility net of price impact-costs that an investor can achieve by having access to a particular investment strategy in addition to the factors in a benchmark model.

Our manuscript is closely related to Detzel et al. (2023), who compare prominent asset-pricing models in the presence of proportional transaction costs using the maximum squared Sharpe ratio criterion of Barillas and Shanken (2017). We formally prove that the squared Sharpe ratio criterion remains valid in the presence of proportional transaction costs, and thus, we provide theoretical support for the empirical analysis of Detzel et al. (2023). We also demonstrate that the squared Sharpe ratio criterion is no longer sufficient to characterize the investment opportunity set in the presence of price-impact costs and, instead, we propose comparing factor models in terms of the mean-variance utility of returns net of price-impact costs. The different comparison methodology and our focus on price-impact costs instead of proportional transaction costs are key distinctive elements of our work.

Our work is also related to Jensen, Kelly, Malamud, and Pedersen (2022), who generalize the dynamic portfolio framework of Gârleanu and Pedersen (2013) to integrate machine-learning return forecasts obtained from a large set of firm characteristics. Like us, Jensen et al. (2022) account for the price-impact costs that are relevant to "market participants with a substantial fraction of aggregate assets under management, such as large pension funds or other professional asset managers." A key distinctive feature of our work is that our focus is not to use machine learning to exploit a large number of characteristics, but rather to propose a rigorous methodology to compare existing asset-pricing models in terms of their ability to span the investment opportunity set in the presence of price impact.

There is a large literature that proposes statistical tests to compare asset-pricing models in the absence of transaction costs (Avramov and Chao, 2006; Kan and Robotti, 2009; Kan, Robotti, and Shanken, 2013; Barillas and Shanken, 2018; Goyal, He, and Huh, 2018; Fama and French, 2018; Ferson, Siegel, and Wang, 2019; Chib, Zeng, and Zhao, 2020; Kan, Wang, and Zheng, 2019). In contrast to these papers, we propose a statistical methodology that accounts for the effect of price-impact costs when comparing asset-pricing models.

Finally, our work is related to the literature on the profitability of factor strategies (Korajczyk and Sadka, 2004; Novy-Marx and Velikov, 2016; Frazzini, Israel, and Moskowitz, 2018; Chen and Velikov, 2022; Barroso and Detzel, 2021). Most of these papers study the profitability of individual-factor strategies. However, DeMiguel et al. (2020) show that the trades in the underlying stocks required to rebalance *different* factors often cancel out, and thus the trading cost of exploiting the factors in a model is lower when the factors are combined.[1] In this manuscript, instead of studying the profitability of the individual factor strategies, we explicitly account for the effect of trading diversification when we compare low- and high-dimensional factor models in the presence of price-impact costs.

The rest of the manuscript is organized as follows. Section 2 proposes mean-variance utility net of price-impact costs as a criterion to compare factor models. Section 3 develops a formal statistical test to compare factor models in the presence of price-impact costs. Section 4 describes our data and compares the empirical performance of six factor models from the literature. Section 5 concludes. Appendix A contains the proofs of all theoretical results with the exception of Proposition 4, which is proven and discussed in Appendix B. The Internet Appendix contains several robustness checks and additional information.

2 Comparing factor models with trading costs

In this section, we propose a novel criterion to compare factor models in the presence of price-impact costs. Section 2.1 gives the notation and assumptions. Section 2.2 reviews the squared Sharpe ratio criterion proposed by Barillas and Shanken (2017) to compare factor models in the absence of trading costs, and in Section 2.3 we prove that this criterion is also valid in the presence of *proportional* transaction costs. In Section 2.4, however, we show

[1] Other papers provide empirical evidence that combining factors can reduce trading costs (Barroso and Santa-Clara, 2015; Frazzini, Israel, and Moskowitz, 2015; Novy-Marx and Velikov, 2016).

that the squared Sharpe ratio criterion is no longer sufficient to characterize the achievable investment opportunity set in the presence of price-impact costs, and thus, in Section 2.5 we propose comparing factor models in terms of their mean-variance utility net of trading costs. Finally, Section 2.6 shows that there is a close relation between the mean-variance utility net of price-impact cost criterion and the alpha *net of price impact*.

2.1 Notation and assumptions

We first describe the notation we use in our analysis. We consider a market with N stocks whose return vector at time t is $r_t \in \mathbb{R}^N$ and a risk-free asset with return $r_{f,t} \in \mathbb{R}$. Let $X_t \in \mathbb{R}^{N \times K}$ be the matrix whose columns contain the weights of the K factor portfolios at time t. Then, the vector of returns of the K factors at time $t+1$ is

$$F_{t+1} = X_t^\top (r_{t+1} - r_{f,t+1} e) \in \mathbb{R}^K, \tag{1}$$

where e is the N-dimensional vector of ones. Every factor we consider is a return in excess of the risk-free rate. In particular, every factor (other than the market) is the return of a long-short portfolio of stocks with one dollar invested in the long leg and one dollar in the short leg, and thus, its returns equal its excess returns. The market factor is also a long-short portfolio because it is the market return in excess of the risk-free rate, and thus, its investment in the long leg is equal to that in the short leg once we account for its negative investment in the risk-free asset.

Let $\mu = E[F_t]$ and $\Sigma = \text{var}(F_t)$ be the mean and covariance matrix of factor returns. Then, the mean-variance factor portfolio, $\theta^* \in \mathbb{R}^K$, is the maximizer to the following problem:

$$\max_{\theta} \qquad \theta^\top \mu - f(\theta) - \frac{\gamma}{2} \theta^\top \Sigma \theta, \tag{2}$$

where the kth component of θ is the *dollar-amount* allocated to the kth factor, $\theta^\top \mu$ is the expected portfolio return, $f(\theta)$ is the trading cost associated with the portfolio θ, $\theta^\top \Sigma \theta$ is the portfolio return variance, and γ is the absolute risk-aversion parameter. Note that because the factors are returns in excess of the risk-free rate, we do not need to impose a budget constraint on the mean-variance factor portfolio weights. Thus, like the portfolio proposed by Gârleanu and Pedersen (2013), our mean-variance factor portfolio depends on the investor's endowment only through her absolute risk aversion, which is the ratio of the investor's relative risk-aversion parameter to her endowment.

A few comments are in order. First, we give specific examples of proportional transaction costs and price-impact costs in Sections 2.3 and 2.4, respectively. Second, although in the main body of the manuscript we consider factors that are constructed as in the original papers in which they were proposed, Section IA.6 of the Internet Appendix shows that our findings are robust to considering factors that are constructed using the *banding* transaction-cost mitigation strategy used by Detzel et al. (2023). Third, consistent with the asset-pricing literature on factor model comparison (Gibbons et al., 1989; Kan and Robotti, 2008; Barillas and Shanken, 2017, 2018; Barillas et al., 2020; Detzel et al., 2023), we consider an *unconditional* mean-variance portfolio of the factors in a model. This is not a limitation because prominent asset-pricing models include factors constructed using time-varying characteristics such as profitability and momentum that encapsulate conditioning information. Moreover, one can also use conditioning variables to generate *managed* versions of popular asset-pricing factors and include them as additional factors in the unconditional mean-variance portfolio.[2]

We now state the assumptions required in our theoretical analysis. First, we require that the factor returns are not perfectly colinear.

Assumption 2.1 *The covariance matrix of the factor returns* Σ *is positive definite.*

Second, we make the following assumption for the functional form of trading costs.

Assumption 2.2 *The trading-cost function* $f(\theta)$ *is continuous in* θ *and such that* $f(0) = 0$ *and* $f(\theta) > 0$ *for all* $\theta \neq 0$.

Assumption 2.2 is satisfied by most popular trading-cost models, such as proportional and quadratic trading-cost models. In particular, prominent asset-pricing models include factors that are constructed using time-varying firm characteristics, and thus, investing in these factors requires the investor to rebalance her portfolio regularly, incurring strictly positive trading costs. Finally, the following assumption rules out the trivial case in which it is not optimal to invest in any of the factors.

Assumption 2.3 *The set* $S = \{\theta | \theta^\top \mu - f(\theta) > 0\}$ *is non-empty.*

[2]For instance, Moreira and Muir (2017) consider volatility-managed factors and DeMiguel, Martin-Utrera, and Uppal (2022) incorporate the volatility-managed factors together with the unmanaged factors in an unconditional mean-variance portfolio.

2.2 The case without trading costs

In the absence of trading costs, the mean-variance portfolio θ^* of the factors is the solution to problem (2) for the case with $f(\theta) = 0$. One can recover all portfolios on the efficient frontier by solving the problem for different values of γ. The following proposition reviews a well-known property of the efficient frontier; see, for instance, Campbell (2017, Section 2.2.6).

Proposition 1 *Let Assumption 2.1 hold and consider an investor with absolute risk aversion $\gamma > 0$. Then, the unique maximizer to the mean-variance problem (2) in the absence of transaction costs is*

$$\theta^* = \Sigma^{-1}\mu/\gamma, \tag{3}$$

the mean-variance utility is $MVU^\gamma = \mu^\top\Sigma^{-1}\mu/(2\gamma)$, and the squared Sharpe ratio is $SR^2 = 2\gamma MVU^\gamma = \mu^\top\Sigma^{-1}\mu$. Thus, the efficient frontier is a straight line in the mean-standard-deviation diagram because every mean-variance portfolio delivers the same maximum Sharpe ratio, $SR = \sqrt{\mu^\top\Sigma^{-1}\mu}$.

Proposition 1 shows that, in the absence of trading costs, the Sharpe ratio of any mean-variance portfolio of the factors in the model is a sufficient statistic to characterize the investment opportunity set spanned by the model. Thus, the model that best spans the investment opportunity set is the one whose factors attain the highest squared Sharpe ratio as noted by Barillas and Shanken (2017).

2.3 The case with proportional trading costs

We first provide a general definition of proportional-trading-cost function.

Definition 1 (Proportional-trading-cost function) *A proportional-trading-cost function $f(\theta)$ is one that satisfies Assumption 2.2 and is homogeneous of degree one, that is,*

$$f(c\theta) = cf(\theta) \quad \text{for all } \theta \text{ and } c \geq 0. \tag{4}$$

We now give a popular example of proportional-trading-cost function used (among others) by DeMiguel et al. (2020) and Detzel et al. (2023). We start by defining the rebalancing-trade matrix of the K factors at time t as

$$\tilde{X}_t = X_t - \text{diag}(e + r_t)X_{t-1}, \tag{5}$$

where e is the N-dimensional vector of ones and $\text{diag}(v)$ is a diagonal matrix whose diagonal contains the elements in vector v. Note that the element in the nth row and kth column of $\tilde{X}_t$ is the rebalancing trade on stock n required at time t to hold the kth factor portfolio. To see this, note that the kth factor portfolio weight on stock n changes from $x_{n,k,t-1}(1 + r_{n,t})$ before rebalancing at time t to $x_{n,k,t}$ after rebalancing, where $x_{n,k,t}$ is the kth factor portfolio weight on the nth stock at time t, that is, the element in the nth row and kth column of X_t. Then, the rebalancing trade required at time t to hold the factor portfolio θ can be written as $\Delta w = \tilde{X}_t\theta$, and thus, the proportional-trading-cost function can be defined as

$$f(\theta) = E\left[\|K_t\tilde{X}_t\theta\|_1\right], \tag{6}$$

where $\|v\|_1 = \sum_{i=1}^{N} |v_i|$ is the 1-norm of vector $v \in \mathbb{R}^N$ and $K_t \in \mathbb{R}^{N\times N}$ is a diagonal matrix whose nth element, $\kappa_{n,t} > 0$, is the transaction-cost parameter of stock n at time t.[3]

Solving problem (2) with a proportional-trading-cost function for different values of the risk-aversion parameter γ, one can recover the efficient frontier in the presence of proportional trading costs. In the following proposition, we prove that this efficient frontier is a straight line in the mean-standard-deviation diagram.[4]

Proposition 2 *Let $f(\theta)$ be a proportional-trading-cost function. Then, the efficient frontier in the presence of proportional trading costs is a straight line in the mean-standard-deviation diagram, and all portfolios on the efficient frontier deliver the same maximum Sharpe ratio of returns net of proportional trading costs, $SR_{PTC} < SR = \sqrt{\mu^\top\Sigma^{-1}\mu}$, where SR is the maximum Sharpe ratio in the absence of trading costs.*

[3] Detzel et al. (2023) consider the proportional-trading-cost function (6) as a robustness check in Section 6.2 of their manuscript. In their main analysis, Detzel et al. (2023) use the following proportional-trading-cost function:

$$f(\theta) = E\left[\sum_{n=1}^{N} \kappa_{n,t} \sum_{k=1}^{K} |\tilde{x}_{n,k,t}\theta_k|\right], \tag{7}$$

where $\tilde{x}_{n,k,t}$ is the rebalancing trade of factor k on stock n at time t, which is the element in the nth row and kth column of the rebalancing-trade matrix $\tilde{X}_t$. An advantage of the proportional-trading-cost function (6) compared to (7) is that it aggregates the rebalancing trades across the K factors and thus accounts for the trading-diversification benefits from combining multiple factors. DeMiguel et al. (2020) find that the trades in the underlying stocks required to rebalance different factors often net out, and therefore exploiting multiple factors simultaneously reduces trading costs.

[4] The mean-standard-deviation diagram for the case with proportional trading costs depicts in the horizontal axis the standard deviation of portfolio returns, and in the vertical axis the mean of portfolio returns net of proportional trading costs.

Proposition 2 shows that, similar to the case without trading costs, the investment opportunity set spanned by the factors in the presence of proportional trading costs is fully characterized by the Sharpe ratio of returns net of costs. Thus, Proposition 2 demonstrates that the maximum squared Sharpe ratio criterion remains valid to compare factor models in the presence of proportional costs, and thus, it provides theoretical support for the empirical analysis in Detzel et al. (2023). However, proportional costs ignore the price impact of large trades, which affects the portfolio choices of large investors and thus the overall achievable investment opportunity set. In the next section, we show that the squared Sharpe ratio criterion is no longer sufficient to characterize the investment opportunity set in the presence of price-impact costs.

2.4 The case with price-impact costs

We now consider the case with price-impact costs. First, we provide a general definition of price-impact-cost function.

Definition 2 (Price-impact-cost function) *A price-impact-cost function $f(\theta)$ satisfies Assumption 2.2 and the following inequality:*

$$f(c\theta) > cf(\theta) \quad \text{for all } \theta \neq 0 \text{ and } c > 1. \tag{8}$$

We now specify the price-impact-cost function that we use in our analysis. A common assumption in the literature is that the impact on prices from large trades is linear in the amount traded (Korajczyk and Sadka, 2004; Novy-Marx and Velikov, 2016). Under this assumption, the *price impact* of rebalancing the factor portfolio at time t is:

$$\mathrm{PI}_t = D_t \Delta w_t = D_t \tilde{X}_t \theta, \tag{9}$$

where $\theta \in \mathbb{R}^K$ is the factor portfolio in dollars, $\tilde{X}_t$ is the rebalancing-trade matrix defined in (5), $\Delta w_t = \tilde{X}_t \theta$ is the rebalancing trade required to rebalance the factor portfolio θ at time t, and $D_t \in \mathbb{R}^{N \times N}$ is a diagonal matrix whose nth element, $d_{n,t} > 0$, is the price-impact-cost parameter (i.e., Kyle's lambda) of stock n at time t. Then, the price-impact *cost*, in dollars, required to rebalance the factor portfolio θ at time t is half of the scalar product of the price impact $\mathrm{PI}_t = D_t \tilde{X}_t \theta$ and the rebalancing trade $\Delta w_t = \tilde{X}_t \theta$:

$$f_t(\theta) = \frac{1}{2} \theta^\top \tilde{X}_t^\top D_t \tilde{X}_t \theta. \tag{10}$$

To simplify notation, let

$$\Lambda_t = \tilde{X}_t^\top D_t \tilde{X}_t \in \mathbb{R}^{K\times K} \tag{11}$$

be the price-impact matrix at time t, and $\Lambda = E[\Lambda_t]$ the expected price-impact matrix, which is assumed to be positive definite. Then, the quadratic price-impact-cost function is

$$f(\theta) = E\left[\frac{\theta^\top \Lambda_t \theta}{2}\right] = \frac{\theta^\top \Lambda \theta}{2}, \tag{12}$$

which gives the expected price-impact costs from trading the factor portfolio θ. It is straightforward to show that this function satisfies Definition 2 and accounts for trading diversification across factors.

The mean-variance problem (2) for the case with quadratic price-impact costs can then be rewritten as

$$\max_{\theta} \quad \theta^\top \mu - \frac{1}{2}\theta^\top \Lambda \theta - \frac{\gamma}{2}\theta^\top \Sigma \theta,$$

where θ is the factor portfolio in dollars, $\theta^\top \mu$ is the expected factor portfolio return, $\theta^\top \Sigma \theta$ is the portfolio variance, and $\theta^\top \Lambda \theta / 2$ is the quadratic price-impact cost. Thus, the mean-variance portfolio is

$$\theta^* = \frac{1}{\gamma}(\Sigma + \Lambda/\gamma)^{-1}\mu, \tag{13}$$

and the investor's mean-variance utility net of price-impact costs is

$$\text{MVU}^\gamma = \frac{\mu^\top (\Sigma + \Lambda/\gamma)^{-1}\mu}{2\gamma}, \tag{14}$$

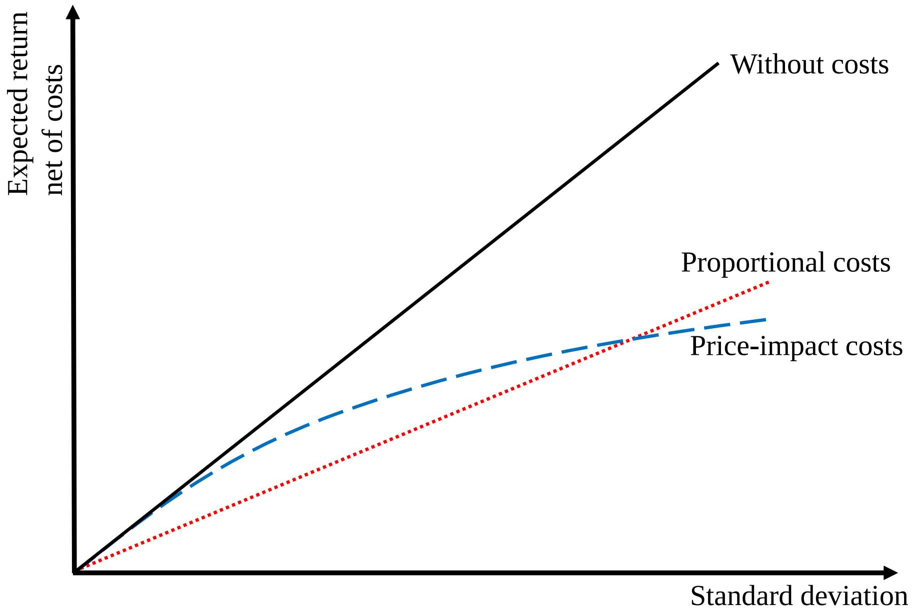
which is *not* proportional to the squared Sharpe ratio in the absence of costs. More precisely, price-impact costs affect the investor's portfolio choice and utility nonlinearly, by replacing the matrix Σ in (3) with the matrix $(\Sigma + \Lambda/\gamma)$, which depends on γ.

Solving problem (2) with a price-impact-cost function for different values of γ, one can recover the efficient frontier in the presence of of price-impact costs. The following proposition shows that the efficient frontier in the presence of price-impact costs is strictly concave in the mean-standard-deviation diagram.

Proposition 3 *Let $f(\theta)$ be a price-impact-cost function. Then, the efficient frontier in the presence of price-impact costs is strictly concave. In addition, the Sharpe ratio of returns net of price-impact costs of any portfolio on the efficient frontier, SR^γ_{PIC}, is lower than the maximum Sharpe ratio in the absence of trading costs, $SR^\gamma_{PIC} < SR = \sqrt{\mu^\top \Sigma^{-1}\mu}$.*

Figure 2: Efficient frontiers for different trading-cost functions

This figure illustrates the efficient frontiers of a factor model in the presence of different trading-cost functions. The black solid, red dotted, and blue dashed lines depict the efficient frontiers in the absence of trading costs, presence of proportional costs, and presence of price-impact costs, respectively.



The intuition behind Proposition 3 is that, while the mean and standard deviation of the portfolio returns grow proportionally with the dollar amount invested, the price-impact costs grow faster than linearly, and thus, the efficient frontier in the presence of price-impact costs is *strictly concave*. Consequently, the squared Sharpe ratio is no longer a sufficient criterion to compare factor models in the presence of price-impact costs because the achievable investment opportunity set of a factor model is not fully characterized by a *single* slope in the mean-standard-deviation diagram as in the absence of trading costs or the presence of proportional trading costs.

Figure 2 illustrates the efficient frontiers attained by the factors of a model for the cases without trading costs, with proportional trading costs, and with price-impact costs. The frontiers for the cases with proportional costs and with price-impact costs are below that for the case without costs. Moreover, while the efficient frontier is a straight line in the cases without costs and with proportional trading costs, in the presence of price-impact costs, the efficient frontier is strictly concave, and thus the investment opportunity set in this case cannot be summarized by a single Sharpe ratio.

2.5 Mean-variance utility as a comparison criterion

In the previous section we showed that, in the presence of price-impact costs, the efficient frontier is strictly concave and thus a single Sharpe ratio no longer characterizes the achievable investment opportunity set as in the cases without costs or with proportional transaction costs. Thus, we cannot compare asset-pricing models in the presence of price-impact costs using the squared Sharpe ratio criterion because this metric is no longer a sufficient statistic to describe the extent to which the factors of a model span the achievable investment opportunity set. Instead, in this section we propose comparing factor models in terms of mean-variance utility net of price-impact costs.

Barillas and Shanken (2017) posit that when comparing two factor models, the better model should be able to span not only the investment opportunity set offered by the tests assets, but also by the factors in the other model. In particular, let us consider two models with factors F_A and F_B and a set of test assets Π. In the absence of price-impact costs, Barillas and Shanken (2017) show that model A is better than model B if

$$SR^2([\Pi, F_A, F_B]) - SR^2(F_A) < SR^2([\Pi, F_A, F_B]) - SR^2(F_B), \tag{15}$$

where $SR^2(x)$ is the squared Sharpe ratio delivered by the assets in vector x. In particular, they explain that the two sides of Inequality (15) measure the *misspecification* of models A and B, and thus, model A is considered better (less misspecified) than model B because an investor with access to the factors in model A obtains a lower Sharpe ratio improvement by having access to the test assets and the factors in the other model than an investor with access to the factors in model B. This inequality is equivalent to

$$SR^2(F_A) > SR^2(F_B), \tag{16}$$

and thus Barillas and Shanken (2017) show that the test assets Π are irrelevant for model comparison, and it is sufficient to compare models in terms of squared Sharpe ratio, which measures the ability of factor models to span the investment opportunity set.

In the absence of trading costs or in the presence of proportional transaction costs, the efficient frontier is a straight line in the mean-standard-deviation diagram, as shown in Propositions 1 and 2. Thus, the portfolios in the efficient frontier that maximize the investor's mean-variance utility are equivalent to those that maximize the Sharpe ratio. In contrast, in the presence of price-impact costs the efficient frontier is strictly concave and hence the portfolios that maximize mean-variance utility are not equivalent to those that maximize Sharpe ratio. Thus, model comparison via the squared Sharpe ratio as in (15) is no longer consistent with the optimal choices of investors that determine the asset-pricing equilibrium. To address this issue, we propose measuring model misspecification in terms of mean-variance utility net of price-impact costs. Thus, applying the logic of Barillas and Shanken (2017), model A is better than model B if

$$\mathrm{MVU}^{\gamma}([\Pi, F_A, F_B]) - \mathrm{MVU}^{\gamma}(F_A) < \mathrm{MVU}^{\gamma}([\Pi, F_A, F_B]) - \mathrm{MVU}^{\gamma}(F_B), \tag{17}$$

where $\mathrm{MVU}^{\gamma}(x)$ is the maximum mean-variance utility net of price-impact costs of an investor with absolute risk aversion γ who has access to the assets in x.[5] Therefore, we have that model A is better than model B if

$$\mathrm{MVU}^{\gamma}(F_A) > \mathrm{MVU}^{\gamma}(F_B), \tag{18}$$

which shows that test assets are irrelevant *also* when comparing factor models in terms of mean-variance utility net of price-impact costs. Consequently, the best model is the one whose factors generate the highest mean-variance utility net of price-impact costs, and thus, is best at spanning the achievable investment opportunity set.

2.6 Relation between mean-variance utility and alpha

In the absence of trading costs, the squared Sharpe ratio criterion proposed by Barillas and Shanken (2017) to compare factor models is closely related to the traditional alpha criterion. In particular, Gibbons et al. (1989) show that a quadratic form of the alpha measures the increase in the squared Sharpe ratio that an investor can achieve by having access to the test assets, in addition to the factors in the model. In this section, we show that the mean-variance utility net of price-impact cost criterion that we propose is also closely related to the alpha *net of price impact*. To do this, in the following proposition, which we prove and discuss in Appendix B, we generalize the result by Gibbons et al. (1989) to the case with quadratic price-impact costs.

Proposition 4 *Consider an investor with absolute risk aversion γ who faces the quadratic price-impact costs defined in (12). Then, the increase in the mean-variance utility net of*

[5] Note that the same argument can be made for investor utility functions other than the mean-variance utility that we consider for our empirical work.

price impact costs of the investor when she has access to a set of test assets R in addition to the factors in the model F is:

$$MVU^{\gamma}([F,R]) - MVU^{\gamma}(F) = \left(\alpha^{net}\right)^{\top} H_{\gamma}^{-1}\alpha^{net}, \tag{19}$$

where H_{γ} is a positive-definite matrix that depends on the investor's absolute risk aversion, and α^{net} is the net alpha of the test assets with respect to the factors in the model:

$$\alpha^{net} = \underbrace{\alpha}_{\substack{gross\\ alpha}} - \underbrace{\left(\Lambda_{R,F} - \beta^{\top}\Lambda_{F,F}\right)\theta^{*}}_{\substack{price\text{-}impact\\ adjustment}}, \tag{20}$$

where α and β are the intercept and slope vectors obtained from an OLS regression of the test asset returns on the factors in the model, θ^{*} is the investor's mean-variance portfolio of the factors in the model, $\Lambda_{F,F} = E[(\tilde{X}_t^F)^{\top} D_t \tilde{X}_t^F]$ is the expected price-impact matrix for the factors in the model, and $\Lambda_{R,F} = E[(\tilde{X}_t^R)^{\top} D_t \tilde{X}_t^F]$ is the expected price-impact matrix for the test assets when the investor is also holding the factors in the model.

A couple of comments are in order. First, Appendix B.2 shows that for the case with no trading costs, Proposition 4 implies the result in equation (23) of Gibbons et al. (1989), which shows that in the absence of trading costs the increase in the squared Sharpe ratio of the investor when she has access to the test assets in addition to the factors in the model is a quadratic form of the gross alpha.

Second, Appendix B.3 shows that the net alpha (α^{net}) defined in (20) is the incremental return net of price-impact costs that an investor with absolute risk-aversion γ can achieve by making a marginal investment in the test assets when she is already holding the mean-variance portfolio of the factors in the model. In other words, the net alpha is a generalization of the traditional alpha to the case with price-impact costs.

2.7 Model performance and absolute risk aversion

Note that the net alpha α^{net} depends on the investor's absolute risk aversion via her mean-variance portfolio $\theta^{*} = (\Sigma_{F,F} + \Lambda_{F,F}/\gamma)^{-1}\mu_F/\gamma$, where μ_F and Σ_F are the mean and covariance matrix of the factors in the model. Thus, the net alpha is different for each investor. Moreover, the matrix H_{γ} also depends on γ. Consequently, Equations (19) and (20) in Proposition 4 show that the relative performance of two factor models in the presence of price impact may depend on the investor's absolute risk aversion, which determines the importance of portfolio risk relative to the average portfolio return net of price-impact costs.

This is illustrated in Figure 1 in the introduction, which depicts the investment opportunity set spanned by two different factor models A and B, where the factors in model A generate a higher Sharpe ratio in the absence of trading costs, but also generate higher price-impact costs as the amount traded increases, compared to the factors in model B. Then, model B is better at spanning the investment opportunities of large investors with low absolute risk aversion, while model A is better at spanning the investment opportunities of investors with high absolute risk aversion. This is because investors with low absolute risk aversion are willing to take on larger investment positions to increase their mean return at the expense of higher return variance. However, by increasing their positions, they also increase the amount they trade, and thus, face higher price-impact costs. Consequently, model B is better at spanning their investment opportunities because its factors generate lower price-impact costs.

3 Statistical tests

We now develop a formal statistical methodology to compare factor models in the presence of price-impact costs. In Section 3.1, we derive two asymptotic distributions for the difference in mean-variance utility net of price-impact costs of two factor models. In Section 3.2, we describe how these two asymptotic distributions can be used to compare two factors models for the cases where they are nested, non-nested without overlapping factors, and non-nested with overlapping factors. Finally, in Section 3.3, we develop closed-form expressions for the variance of the asymptotic distribution and use them to study how the statistical properties of factor models affect the power of our proposed test.

3.1 Two asymptotic distributions

We assume price-impact costs are quadratic as in (12). Also, for simplicity we make Assumption 3.1, but it can be relaxed by adjusting the variance of the asymptotic distribution.

Assumption 3.1 *The factor returns F_t, the matrix $\Sigma_t = (F_t - \mu)(F_t - \mu)^\top$, and the price-impact matrix Λ_t in Equation (11) are serially uncorrelated.*

We now derive two asymptotic distributions in Propositions 5 and 6 for the difference between the sample mean-variance utilities net of price-impact costs of two factor models.

Proposition 5 *Let Assumptions 2.1–2.3 and 3.1 hold. Then, the asymptotic distribution of the sample estimator of the mean-variance utility net of price-impact costs in (14) is*

$$\sqrt{T}(\widehat{MVU^{\gamma}} - MVU^{\gamma}) \overset{A}{\sim} N(0, \frac{E[h_t^2]}{4\gamma^2}), \tag{21}$$

provided that $E[h_t^2] > 0$, *where*

$$\begin{aligned} h_t &= 2\mu^\top(\Sigma + \Lambda/\gamma)^{-1}(F_t - \mu) - \mu^\top(\Sigma + \Lambda/\gamma)^{-1}(\Sigma_t + \Lambda_t/\gamma)(\Sigma + \Lambda/\gamma)^{-1}\mu \\ &\quad + \mu^\top(\Sigma + \Lambda/\gamma)^{-1}\mu. \end{aligned} \tag{22}$$

In addition, the asymptotic distribution of the difference between the sample mean-variance utilities net of price-impact costs of two factor models A and B is

$$\sqrt{T}([\widehat{MVU^{\gamma}_A} - \widehat{MVU^{\gamma}_B}] - [MVU^{\gamma}_A - MVU^{\gamma}_B]) \overset{A}{\sim} N(0, \frac{E\big[(h_{t,A} - h_{t,B})^2\big]}{4\gamma^2}), \tag{23}$$

provided that $E\big[(h_{t,A} - h_{t,B})^2\big] > 0$, *where $h_{t,A}$ and $h_{t,B}$ are given by Equation (22) applied to models A and B.*

A couple of comments are in order. First, Proposition 5 generalizes the analysis of Barillas et al. (2020), who provide an asymptotic distribution for the difference in squared Sharpe ratios in the absence of costs. Second, Proposition 5 shows that the distribution in (23) can be used to compare factor models provided that the variance of the asymptotic distribution is strictly greater than zero. However, the variance is zero under the null hypothesis, $\text{MVU}^{\gamma}_A = \text{MVU}^{\gamma}_B$, in two cases.[6] First, when model A nests model B and the extra factors of model A are redundant, and second, when models A and B overlap (share common factors) and the extra factors of both models are redundant. In both cases, one cannot apply Proposition 5 to test whether two models generate the same maximum mean-variance utility net of price-impact costs. Therefore, we provide in Proposition 6 another asymptotic distribution to deal with these two cases. Section 3.2 discusses how Propositions 5 and 6 can be used to compare nested or non-nested factor models.

[6] Barillas et al. (2020) discuss a similar issue for the case without transaction costs.

Proposition 6 *Let Assumptions 2.1–2.3 and 3.1 hold. Consider two nested models A and B containing factors $F_A = [F_1, F_2]$ and $F_B = F_1$, where F_1 and F_2 contain K_1 and K_2 mutually exclusive factors. Partition the matrix $\Sigma_A + \Lambda_A/\gamma$ as*

$$\Sigma_A + \Lambda_A/\gamma = \begin{bmatrix} \Sigma_{11} + \Lambda_{11}/\gamma & \Sigma_{12} + \Lambda_{12}/\gamma \\ \Sigma_{21} + \Lambda_{21}/\gamma & \Sigma_{22} + \Lambda_{22}/\gamma \end{bmatrix},$$

where $\Sigma_{22} + \Lambda_{22}/\gamma \in \mathbb{R}^{K_2 \times K_2}$. Then, under the null hypothesis that $MVU_A^\gamma = MVU_B^\gamma$, the asymptotic distribution of the difference between the sample mean-variance utilities net of price-impact costs of the two models A and B is given by

$$T(\widehat{MVU_A^\gamma} - \widehat{MVU_B^\gamma}) \overset{A}{\sim} \sum_{i=1}^{K_2} \xi_i x_i, \tag{24}$$

where x_i for $i = 1, \ldots, K_2$ are independent chi-square random variables with one degree of freedom, and ξ_i for $i = 1, \ldots, K_2$ are the eigenvalues of matrix

$$\frac{E[l_t l_t^\top]_{22} W}{2\gamma}, \tag{25}$$

where

$$W = (\Sigma_{22} + \Lambda_{22}/\gamma) - (\Sigma_{21} + \Lambda_{21}/\gamma)(\Sigma_{11} + \Lambda_{11}/\gamma)^{-1}(\Sigma_{12} + \Lambda_{12}/\gamma) \quad \textit{and} \tag{26}$$

$$l_t = (\Sigma_A + \Lambda_A/\gamma)^{-1} R_{A,t} - (\Sigma_A + \Lambda_A/\gamma)^{-1}(\Sigma_{A,t} + \Lambda_{A,t}/\gamma)(\Sigma_A + \Lambda_A/\gamma)^{-1}\mu_A. \tag{27}$$

This proposition is related to Proposition 2 of Kan and Robotti (2009), which compares nested factor models in terms of their Hansen-Jagannathan distance in the absence of trading costs. We extend their result to compare nested factor models in terms of mean-variance utility net of price-impact costs.[7]

3.2 Comparing models with any nesting structure

We now show how to compare two factor models with any nesting structure using Propositions 5 and 6. We consider three cases: (i) non-nested factor models without overlapping factors, (ii) nested factor models, and (iii) non-nested factor models with overlapping factors.

[7]Note that to compare nested models in the *absence* of trading costs, one can either use Proposition 6 with $\Lambda = 0$, or run time-series regressions of the additional factors of the larger model on the common factors of the two models, and apply the GRS test to assess whether the non-common factors contribute to expand the investment opportunity set of the common factors. Section IA.1 of the Internet Appendix compares these two approaches in the absence of trading costs.

When models A and B are non-nested and have no overlapping factors, the variance of the asymptotic distribution in (23) is strictly greater than zero. Therefore, one can directly apply Proposition 5 and reject the null hypothesis $\text{MVU}_A^\gamma = \text{MVU}_B^\gamma$ when $\sqrt{T}(\widehat{\text{MVU}_A^\gamma} - \widehat{\text{MVU}_B^\gamma})$ is greater (less) than, for instance, the 97.5th (2.5th) percentile of the probability density function on the right-hand side of (23).

However, as explained in the previous section, one cannot use Proposition 5 to compare nested factor models because under the null hypothesis where the extra factors of the larger model are redundant, the variance of the distribution in (23) is zero. Instead, we use Proposition 6 and reject the null hypothesis $\text{MVU}_A^\gamma = \text{MVU}_B^\gamma$ when $T(\widehat{\text{MVU}_A^\gamma} - \widehat{\text{MVU}_B^\gamma})$ is greater than, for instance, the 95th percentile of the probability density function of the distribution on the right-hand side of (24), in which case the larger model A performs significantly better than the smaller model B.

Comparing two non-nested models with overlapping factors is more complicated because, as Barillas et al. (2020) point out, the null hypothesis may hold in two ways: (i) the two models have the same mean-variance utility net of price-impact costs as the common factors of the two models, and (ii) the two models have the same utility net of price-impact costs and it is *higher* than that of their common factors. In the first case, the extra factors of both models are redundant and Proposition 5 cannot be applied because the variance of the distribution in (23) is zero. Thus, we test whether the null hypothesis holds using Proposition 6 where we define a nesting model containing all factors of models A and B, and a nested model containing only the common factors of models A and B. If this test does not reject the null, the two models are statistically indistinguishable in the first way. However, if this test rejects its null, then the null hypothesis does not hold in the first way, but it may still hold in the second way, which can be tested using Proposition 5 because in this case the asymptotic variance in (23) is greater than zero.

Finally, to empirically characterize the asymptotic distribution in Proposition 5, one can replace h_t in (22) with its sample counterpart, $\hat{h}_t$, which guarantees that $\sum_{t=1}^{T}(\hat{h}_{t,A} - \hat{h}_{t,B})^2/T$ is a consistent estimator of $E[(h_{t,A} - h_{t,B})^2]$. Similarly, to empirically characterize the asymptotic distribution in Proposition 6, one can replace $E[l_t l_t^\top]_{22}$ and W in (25) with their sample counterparts to obtain consistent estimators of the eigenvalues ξ_i in (24).

3.3 The asymptotic variance

In this section, we derive closed-form expressions for the asymptotic variance in Proposition 5, and use them to study how the statistical properties of factor models affect the power of our proposed test. Our main finding is that it is easier to reject the null hypothesis that the mean-variance utilities net of price-impact costs of two models are equal not only when the mean-variance portfolio returns of the two models are positively correlated as shown by Barillas et al. (2020) for the case without trading costs, but also when the mean-variance portfolio return of each model is highly correlated with the rebalancing trades of the portfolio of the other model, and when the rebalancing trades of the two portfolios are highly correlated.

Let the *matrix of scaled rebalancing trades* at time t be

$$\tilde{Y}_t = \frac{D_t^{1/2}\tilde{X}_t}{\sqrt{\gamma}} \in \mathbb{R}^{N\times K},$$

where D_t, defined in (9), is the diagonal matrix whose nth element, $d_{n,t}$, is the price-impact parameter of stock n at time t. Note that

$$E[\tilde{Y}_t^\top \tilde{Y}_t] = E\left[\frac{\tilde{X}_t^\top D_t \tilde{X}_t}{\gamma}\right] = \frac{\Lambda}{\gamma}.$$

Let $\tilde{y}_{n,t} \in \mathbb{R}^K$ be the nth row of matrix $\tilde{Y}_t$, which contains the scaled rebalancing trades on the nth stock required by the K factors at time t.

For simplicity, we assume that the factor returns F_t and $\tilde{y}_{n,t}$ are normally distributed, but similar results can be derived for the case where they are elliptically distributed.

Assumption 3.2 *The factor returns F_t follow a multivariate normal distribution with mean μ and covariance matrix Σ. In addition, each vector $\tilde{y}_{n,t}$ for $n = 1, \ldots, N$ follows a multivariate normal distribution with zero mean and covariance matrix Λ_n/γ.*

The following proposition gives the closed-form expressions for the asymptotic variance of the sample mean-variance utility net of price-impact costs of a factor model and that of the difference between the sample mean-variance utilities of two models. For notational simplicity, we define $u_t = \mu^\top(\Sigma + \Lambda/\gamma)^{-1}F_t \in \mathbb{R}$, which is proportional to the mean-variance factor portfolio return at time t, and $v_{n,t} = \mu^\top(\Sigma + \Lambda/\gamma)^{-1}\tilde{y}_{n,t} \in \mathbb{R}$, which is proportional to the total scaled rebalancing trade on stock n at time t of the mean-variance factor portfolio.

Proposition 7 *Let Assumptions 2.1–2.3, 3.1, and 3.2 hold. Then,*

$$E[h_t^2] = 4\text{var}(u_t) + 2\left[\text{var}(u_t)\right]^2 + 4\sum_{n=1}^{N}\left[\text{cov}(u_t, v_{n,t})\right]^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\text{cov}(v_{i,t}, v_{j,t})\right]^2. \tag{28}$$

Moreover, given two factor models A and B, we have

$$E[(h_{t,A} - h_{t,B})^2] = E[h_{t,A}^2] + E[h_{t,B}^2] - 2E[h_{t,A}h_{t,B}], \tag{29}$$

where $E[h_{t,A}^2]$ *and* $E[h_{t,B}^2]$ *are given by applying (28) to models A and B, and*

$$E[h_{t,A}h_{t,B}] = 4\text{cov}(u_t^A, u_t^B) + 2\left[\text{cov}(u_t^A, u_t^B)\right]^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\text{cov}(v_{i,t}^A, v_{j,t}^B)\right]^2$$
$$+ 2\sum_{n=1}^{N}\left(\left[\text{cov}(u_t^A, v_{n,t}^B)\right]^2 + \left[\text{cov}(u_t^B, v_{n,t}^A)\right]^2\right). \tag{30}$$

Equation (28) shows that the asymptotic variance of the sample mean-variance utility net of price-impact costs increases not only with the variance of the mean-variance portfolio returns, $\text{var}(u_t)$, as shown by Barillas et al. (2020) for the case without trading costs, but also with the squared covariance between the mean-variance portfolio returns and the rebalancing trades for each stock in the mean-variance portfolio, $[\text{cov}(u_t, v_{n,t})]^2$, and with the squared covariance between the rebalancing trades for different firms in the mean-variance portfolio, $[\text{cov}(v_{i,t}, v_{j,t})]^2$.

Equations (29) and (30) show that, similar to the case without costs, the asymptotic variance of the difference between the estimated mean-variance utilities net of price-impact costs of two models increases with the variance of the mean-variance portfolio return for each of the two models, and decreases with the covariance of the mean-variance portfolio returns for the two models, provided that $\text{cov}(u_t^A, u_t^B) > -1$. In addition, the asymptotic variance of the difference decreases with the squared covariance between the mean-variance portfolio return of one model and the rebalacing trades for each stock in the mean-variance portfolio of the other model, $[\text{cov}(u_t^A, v_{n,t}^B)]^2$ and $[\text{cov}(u_t^B, v_{n,t}^A)]^2$, and with the squared covariance between the rebalacing trades of the stocks in the mean-variance portfolios of the two models, $[\text{cov}(v_{i,t}^A, v_{j,t}^B)]^2$.

Consequently, it is easier to reject the null hypothesis that the mean-variance utilities net of price-impact costs of two models are equal when the mean-variance portfolio returns of the two models are highly positively correlated, the mean-variance portfolio return of each model is highly correlated with the rebalancing trades of the portfolio of the other model, and when the rebalancing trades of the two portfolios are highly correlated.[8]

4 Empirical results

In this section, we use the asymptotic distributions derived in Section 3 to compare the empirical performance of six factor models in the presence of price-impact costs. Section 4.1 lists the six factor models we consider and describes the data we use to construct their factors. Section 4.2 describes how we estimate the price-impact cost incurred by different stocks. Section 4.3 reports summary statistics for the factors. Section 4.4 compares the different models using the statistical tests introduced in Section 3. Finally, as a robustness check, Section 4.5 compares the out-of-sample performance of the different models using the bootstrap approach of Fama and French (2018).

4.1 Factor models and data

Table 1 lists the six factor models we consider. In particular, we consider five popular *low-dimensional* factor models: the CAPM model of Sharpe (1964) and Lintner (1965), the four-factor model of Hou et al. (2015), HXZ4, the four-factor model of Fama and French (1993) and Carhart (1997), FFC4, the five-factor model of Fama and French (2015), FF5, and the six-factor model of Fama and French (2018), FF6. In addition, to evaluate the trading-diversification benefits from combining a large number of factors, we consider a *high-dimensional* factor model containing the 20 factors that DeMiguel et al. (2020) find statistically significant in the presence of price-impact costs, DMNU20.

To construct the factors associated with the aforementioned six factor models, we download data for the 28 tradable factors listed in Table 2. Our sample spans the period from January 1980 to December 2020. We consider nine factors included in prominent low-dimensional asset-pricing models. In particular, we construct the market (MKT), size (SMB), value (HML), profitability (RMW) and investment (CMA) factors of Fama and French (2015), the momentum (UMD) factor of Carhart (1997), and the profitability (ROE),

[8] To estimate the asymptotic variances, one can plug the sample estimators $\hat{\mu}, \hat{\Sigma}$, and $\hat{\Lambda}_n$ into the closed-form expressions in Proposition 7.

Table 1: List of factor models considered

This table lists the factor models we consider, ordered in increasing number of factors. The first column gives the acronym of the model, the second column the number of factors in the model (K), the third and fourth columns the authors who proposed the model, and the date and journal of publication, respectively. The last column lists the acronyms of the factors in the model.

Acronym	K	Authors	Date, journal	Factor acronyms
CAPM	1	Sharpe and Lintner	1964, JOF and 1965, JOF	MKT
HXZ4	4	Hou, Xue & Zhang	2015, RFS	MKT, ROE, IA, ME
FFC4	4	Fama & French and Carhart	1993, JFE and 1997, JOF	MKT, SMB, HML, UMD
FF5	5	Fama & French	2015, JFE	MKT, SMB, HML, RMW, CME
FF6	6	Fama & French	2018, JFE	MKT, SMB, HML, RMW, CME, UMD
DMNU20	20	DeMiguel, Martin-Utrera, Nogales & Uppal	2020, RFS	MKT, agr, cashpr, chatoia, chcsho, convind, egr, ep, gma, idiovol, ind-mom, ps, rd_mve, retvol, roaq, sgr, std_turn, sue, turn, zerotrade

investment (IA), and size (ME) factors of Hou et al. (2015). We construct the market factor as the excess return of the value-weighted market portfolio and the rest of the factors as the returns of value-weighted long-short portfolios obtained from double or triple sorts on firm characteristics following the procedure in the papers that originally proposed the factors. Detzel et al. (2023) show that the transaction-cost mitigation strategy known as *banding*, which was proposed by Novy-Marx and Velikov (2016), can help to improve the performance of models that use factors whose portfolios are rebalanced monthly and quarterly. Section IA.6 of the Internet Appendix shows that our findings are robust to considering the case where the factors that are rebalanced monthly—momentum (UMD), profitability (ROE), investment (IA), and size (ME)—are constructed using the banding transaction-cost mitigation strategy employed by Detzel et al. (2023).

DeMiguel et al. (2020) provide an economic rationale based on trading costs to consider high-dimensional factor models. Moreover, in their Appendix IA.2, they propose a model containing 20 factors, including the market, that are statistically significant in the presence of price-impact costs. Therefore, we construct the 19 factors (other than the market) in the model of DeMiguel et al. (2020) as the returns on value-weighted long-short portfolios obtained from single sorts on 19 firm characteristics. In particular, we start with a database that contains every firm traded on the NYSE, AMEX, and NASDAQ exchanges. We then drop firms with negative book-to-market or with market capitalization below the

Table 2: List of characteristics considered

This table lists the 28 factors we consider. Panel A lists nine factors that replicate those in prominent asset-pricing models, including the market. Other than the market, each of these factors are constructed as value-weighted portfolios obtained from double or triple sorts on firm characteristics. Panel B lists 19 factors constructed using value-weighted portfolios from single sorts on characteristics that together with the market factor compose the 20-factor model of DeMiguel et al. (2020). The first column gives the factor number, the second column gives the factor's definition, the third column gives the acronym, and the fourth and fifth columns give the authors who analyzed them, and the date and journal of publication, respectively.

#	Definition	Acronym	Author(s)	Date and Journal
Panel A: Market factor and factors constructed from double and triple sorts				
1	Market: value-weighted portfolio of all tradable stocks in US equity markets.	MKT	Sharpe	1964, JF
2	Small-minus-big: value-neutral portfolio that is long stocks with small market capitalization and is short stocks with large market capitalization.	SMB	Fama & French	1993, JFE
3	High-minus-low: size-neutral portfolio that is long stocks with high book-to-market ratios and is short stocks with low book-to-market ratios.	HML	Fama & French	1993, JFE
4	Robust-minus-weak: size-neutral portfolio that is long stocks with high operating profitability and is short stocks with low operating profitability.	RMW	Fama & French	2015, JFE
5	Conservative-minus-aggressive: size-neutral portfolio that is long stocks with high investment and is short stocks with low investment.	CMA	Fama & French	2015, JFE
6	Momentum: portfolio that is long stocks with the largest return over the past 12 months, skipping the last month, and is short stocks with the lowest return over the past 12 months, skipping the last month.	UMD	Carhart	1997, JF
7	Return on equity: portfolio that is long stocks with high profitability and is short stocks with low profitability.	ROE	Hou, Xue & Zhang	2015, RFS
8	Investment: portfolio that is long stocks with high investment and is short stocks with low investment.	IA	Hou, Xue & Zhang	2015, RFS
9	Size: portfolio that is long stocks with low market capitalization and is short stocks with large market capitalization.	ME	Hou, Xue & Zhang	2015, RFS

Table 2 continued

#	Definition	Acronym	Author(s)	Date and journal
Panel B: Factors constructed from single sorts				
10	Asset growth: Annual percent change in total assets	agr	Cooper, Gulen & Schill	2008, JF
11	Cash productivity: Fiscal year-end market capitalization plus long term debt minus total assets divided by cash and equivalents	cashpr	Chandrashekar & Rao	2009 WP
12	Industry adjusted change in asset turnover: 2-digit SIC fiscal-year mean adjusted change in sales divided by average total assets	chatoia	Soliman	2008, TAR
13	Change in shares outstanding: Annual percent change in shares outstanding	chcsho	Pontiff & Woodgate	2008, JF
14	Convertible debt indicator: An indicator equal to 1 if company has convertible debt obligations	convind	Valta	2016 JFQA
15	Change in common shareholder equity: Annual percent change in equity book value	egr	Richardson, Sloan, Soliman & Tuna	2005, JAE
16	Earnings to price: Annual income before extraordinary items divided by end of fiscal year market cap	ep	Basu	1977, JF
17	Gross profitability: Revenues minus cost of goods sold divided by lagged total assets	gma	Novy-Marx	2013, JFE
18	Idiosyncratic return volatility: Standard deviation of residuals of weekly returns on weekly equal weighted market returns for 3 years prior to month-end	idiovol	Ali, Hwang & Trombley	2003, JFE
19	Industry momentum: Equal weighted average industry 12-month returns	indmom	Moskowitz & Grinblatt	1999, JF
20	Financial-statements score: Sum of 9 indicator variables to form fundamental health score	ps	Piotroski	2000, JAR
21	R&D to market cap: R&D expense divided by end-of-fiscal-year market cap	rd_mve	Guo, Lev & Shi	2006, JBFA
22	Return volatility: Standard deviation of daily returns from month $t-1$	retvol	Ang, Hodrick, Xing & Zhang	2006, JF
23	Return on assets: Income before extraordinary items divided by one quarter lagged total assets	roaq	Balakrishnan, Bartov & Faurel	2010, JAE
24	Annual sales growth: Annual percent change in sales	sgr	Lakonishok, Shleifer & Vishny	1994, JF
25	Volatility of share turnover: Monthly standard deviation of daily share turnover	std_turn	Chordia, Subrahmanyan & Anshuman	2001, JFE
26	Unexpected quarterly earnings: Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings is I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file	sue	Rendelman, Jones & Latane	1982, JFE
27	Share turnover: Average monthly trading volume for most recent 3 months scaled by number of shares outstanding in current month	turn	Datar, Naik & Radcliffe	1998, JFM
28	Zero trading days: Turnover weighted number of zero trading days for most recent month	zerotrade	Liu	2006, JFE

20th cross-sectional percentile as in Brandt, Santa-Clara, and Valkanov (2009) and DeMiguel et al. (2020). We then rank stocks at the beginning of every month based on a particular firm characteristic and build a long value-weighted portfolio of stocks whose characteristic is above the 70th percentile and a short value-weighted portfolio of stocks below the 30th percentile.

4.2 Estimating the price-impact cost parameters

We explain in this section how we estimate the price-impact parameters of individual stocks in Equation (12), which are required for the computation of price-impact costs incurred by the factors. Following Novy-Marx and Velikov (2016), we use the Trade and Quote (TAQ) data from December 2003 to December 2020 to estimate the price-impact parameter (Kyle's lambda) of the nth stock in month t by regressing daily stock returns on daily order flows:

$$r_{n,\tau} = \alpha_n + d_{n,t}\text{OrderFlow}_{n,\tau} + \varepsilon_{n,\tau}, \tag{31}$$

where $r_{n,\tau}$ is the return of stock n on day τ and $\text{OrderFlow}_{n,\tau}$ is the order flow of stock n on day τ.[9] For the earlier part of our sample from January 1980 to December 2003, we estimate the price-impact parameter of the nth stock in month t, $d_{n,t}$, following DeMiguel et al. (2020, Appendix IA.2) who rely on the empirical results of Novy-Marx and Velikov (2016) based on Trade and Quote (TAQ) data.[10]

As in Korajczyk and Sadka (2004) and Novy-Marx and Velikov (2016), we express all quantities, including the optimal factor portfolio θ, in terms of market capitalization at the end of our sample (December 2020). To make price-impact costs comparable over the entire estimation window from 1980 to 2020, we scale the price-impact parameter, $d_{n,t}$,

[9] Order flow is defined as the dollar value of the difference between the buyer and seller initiated trades. The daily order flow data is obtained from the Millisecond Trade and Quote (TAQ) dataset, and the trades are signed using the algorithm of Lee and Ready (1991). The price-impact parameters are estimated monthly using daily observations from the previous year.

[10] Specifically, Novy-Marx and Velikov (2016) show that the R-squared of a cross-sectional regression of log transaction-cost parameters on log market capitalization is 70% and the slope is statistically indistinguishable from minus one. This suggests that a good approximation to the cross-sectional variation of price-impact cost parameters is to assume they are inversely proportional to market capitalization. Therefore, for months between December 1993 and December 2003, we use figure 4 in Novy-Marx and Velikov (2016) to recover estimates of the cross-sectional average price elasticity of stock supply, defined as the product between estimated price impacts per dollar traded and market capitalization, and estimate the price-impact parameter of stock n at month t as the ratio between average price elasticity of supply at month t and the market capitalization of stock n at month t. In addition, we estimate the price-impact parameter of stock n at month t from January 1980 to December 1993 as the ratio between 6.5 and the market capitalization of stock n at month t, where 6.5 is the time-series average of the average cross-sectional price elasticity.

by multiplying it with the ratio of the *aggregate* market capitalization month t to that in December 2020.

4.3 Factor summary statistics

Table 3 reports summary statistics for the 28 factors listed in Table 2. The first column gives the acronym of the factor. The second and third columns give the average monthly *gross* return of the factor and its t-statistic. The fourth and fifth columns give the average monthly *net-of-price-impact-costs* return of the factor and its t-statistic, when one invests one billion dollars on each leg of the factor. The sixth column gives the factor's monthly price-impact cost (PIC), the seventh column the factor's monthly turnover (TO), and the eighth column the factor's capacity. The ninth column reports the average of the monthly trade-weighted market capitalization, and the last column reports the average of the trade-weighted market capitalization at the end of June. Average returns, turnovers, and price-impact costs are reported in percentage. Investment positions, capacity, and trade-weighted market capitalization are reported in terms of market capitalization at the end of our sample, which spans the period from January 1980 to December 2020.

Consistent with the findings of Detzel et al. (2023), we find that, among the factors constructed from double and triple sorts, factors that are rebalanced monthly (UMD, ROE, IA, ME) have monthly turnovers ranging between 19.19% to 51.93% that are much higher than those of factors that are rebalanced annually (SMB, HML, RMW, CMA), which range between 7.90% and 15.14%. As a result, the annually rebalanced factors have, on average, lower price-impact costs and higher capacity than the monthly rebalanced factors. In particular, the average monthly price-impact cost and capacity of the four annually rebalanced factors are 0.016% and 13.83 billion dollars, respectively, while those of the four monthly rebalanced factors are 0.063% and 6.45 billion dollars, respectively. However, we also find that the relative performance of factors in terms of turnover is different from that in terms of price-impact costs. For instance, while UMD is the factor with the highest turnover, ROE is the factor with the highest price-impact costs. In particular, for the case where one invests one billion dollars on each leg of the factors, the monthly price-impact cost of UMD is around eight basis points, but that of the ROE factor is larger than ten basis points.

Table 3: Factor summary statistics

This table reports several summary statistics of the factors. The first column gives the acronym of the factor. The second and third columns give the average monthly *gross* return of the factor and its t-statistic. The fourth and fifth columns give the average monthly *net-of-price-impact-costs* return of the factor and its t-statistic, when one invests one billion dollars on each leg of the factor. The sixth column gives the factor's monthly price-impact cost (PIC), the seventh column the factor's monthly turnover (TO), and the eighth column the factor's capacity. The ninth column reports the average of the monthly trade-weighted market capitalization, and the last column reports the average of the trade-weighted market capitalization at the end of June. Average returns, turnovers, and price-impact costs are reported in percentage. Investment positions, capacity, and trade-weighted market capitalization are reported in terms of market capitalization at the end of our sample, which spans the period from January 1980 to December 2020.

Factor	Gross returns (%)		Net returns (%)		Costs (%), turnover (%), and capacity ($B)			Trade-weighted market cap ($B)	
	Average	t-statistic	Average	t-statistic	PIC	TO	Capacity	Monthly	June
Panel A: Market and factors constructed from double and triple sorts									
MKT	0.705	3.462	0.705	3.462	0.000	2.18	–	163.43	152.70
SMB	0.086	0.639	0.080	0.597	0.006	7.90	15.13	68.91	48.84
HML	0.163	1.196	0.147	1.076	0.016	10.61	10.13	70.15	54.54
RMW	0.350	3.293	0.333	3.136	0.016	10.62	21.20	64.59	55.52
CMA	0.240	2.666	0.213	2.354	0.027	15.14	8.85	80.69	84.35
UMD	0.557	2.744	0.476	2.343	0.081	51.93	6.86	90.52	73.47
ROE	0.521	4.394	0.420	3.536	0.101	35.42	5.16	63.41	55.13
IA	0.286	3.309	0.235	2.703	0.051	24.60	5.64	68.15	67.19
ME	0.147	1.108	0.129	0.974	0.018	19.19	8.12	70.11	58.26
Panel B: Factors constructed from single sorts									
agr	0.163	1.366	0.153	1.282	0.010	15.29	16.50	147.43	159.89
cashpr	0.013	0.092	0.010	0.072	0.003	8.04	4.60	158.52	169.58
chatoia	0.165	2.101	0.153	1.953	0.012	16.48	14.27	166.40	176.58
chcsho	0.297	2.852	0.288	2.760	0.009	14.00	31.56	160.23	186.96
convind	0.098	1.034	0.094	1.001	0.003	6.30	31.59	143.81	132.44
egr	0.164	1.496	0.154	1.409	0.010	15.13	17.24	147.72	160.82
ep	0.213	1.186	0.201	1.117	0.012	14.42	17.21	120.75	132.96
gma	0.220	1.674	0.218	1.661	0.002	6.76	121.82	164.20	142.50
idiovol	0.203	0.730	0.181	0.652	0.022	11.41	9.29	74.71	65.29
indmom	0.210	1.348	0.189	1.212	0.021	40.89	9.90	179.78	175.79
ps	0.160	1.695	0.141	1.493	0.019	18.17	8.35	148.30	173.55
rd_mve	0.409	2.390	0.394	2.300	0.015	10.82	26.68	167.53	176.52
retvol	0.388	1.489	0.225	0.861	0.163	83.62	2.38	98.89	87.82
roaq	0.272	1.841	0.245	1.655	0.027	25.76	9.97	110.46	94.00
sgr	0.100	0.743	0.090	0.669	0.010	15.30	10.07	159.01	173.36
std_turn	0.088	0.458	0.024	0.124	0.064	78.92	1.37	118.70	100.90
sue	0.238	2.307	0.170	1.642	0.068	45.74	3.48	107.80	87.05
turn	0.020	0.098	-0.000	-0.001	0.020	28.77	0.99	161.56	147.26
zerotrade	0.221	1.120	0.144	0.729	0.077	62.08	2.86	173.09	162.93

To understand the difference in the relative performance of factors in terms of turnover and price-impact cost, the last two columns of Table 3 report the average trade-weighted market capitalization (in billions of dollars) of the different factors listed in Table 2. In particular, for each factor we compute the monthly trade-weighted market capitalization of the stocks traded by the factor and report the time-series average. Table 3 shows that, as expected, the factor that trades in the largest, and thus, most liquid stocks is the market (MKT). Specifically, the average firm traded by the MKT factor has a market capitalization of 163.43 billion dollars. In contrast, the average market capitalization of the stocks traded by the return on equity (ROE) and the investment (IA) factors of Hou et al. (2015) is only 63.41 and 68.15 billion dollars, respectively. The low market capitalization of the average stock traded by the ROE factor explains why the price-impact cost of ROE is much higher than the price-impact cost of UMD, even though UMD has a substantially higher turnover. Finally, Panel B shows that the trade-weighted market capitalization of the factors constructed from single sorts is substantially higher than that of the factors obtained from double and triple sorts. This is because the factors obtained from single sorts assign a much lower weight to small stocks compared to factors obtained from double or triple sorts, which use market capitalization as one of the sorting variables. As a result, although the monthly turnover of the factors from single sorts is comparable to that of the factors from double and triple sorts, their price-impact costs are generally lower.

In summary, the results in this section show that the price-impact costs incurred by the different factors depend not only on the turnover required to rebalance them, which was highlighted by Detzel et al. (2023) as an important driver in the context of *proportional* transaction costs, but also on the size and liquidity of the stocks traded.

4.4 Model comparison using our proposed statistical tests

In this section, we compare the performance of the six factor models listed in Table 1 in the presence of price-impact costs using the statistical tests developed in Section 3. Like Gârleanu and Pedersen (2013), we consider a base case with an absolute risk-aversion parameter of $\gamma = 10^{-9}$, which corresponds to an investor with a relative risk-aversion parameter of five and an endowment of five billion dollars. For comparison, we also consider cases where the investor has the same relative risk-aversion parameter, but her endowment is ten times larger or smaller than in the base case; that is, when $\gamma = 10^{-10}$ or $\gamma = 10^{-8}$. For a constant relative risk-aversion level, a lower absolute risk-aversion parameter implies a larger investor, and therefore price-impact costs play a more important role in the investor's mean-variance utility.

Note that the CAPM is nested by all other models and that both FFC4 and FF5 are nested by the FF6 model. Thus, we use Proposition 6 to compare the CAPM with all the other models and to compare FFC4 and FF5 with FF6. Also, all models have one common market factor. Therefore, following our discussion in Section 3.2, we compare non-nested models with overlapping factors in two stages. First, we use Proposition 6 to test whether a model with all factors in the two models yields the same utility net of price-impact costs as a model with only the common factors. If the test does not reject the null, then the two models are statistically indistinguishable.[11] If the test rejects the null, we then implement the second-stage test that uses Proposition 5 to compare the two models.[12]

To understand how price-impact costs affect the relative performance of the six factor models, we first compare their performance in the *absence* of price-impact costs. Panel A in Table 4 reports the sample mean-variance utility of each model in the absence of price-impact costs and Panel B reports the p-values for all pairwise comparisons. To facilitate the comparison of utilities across different values of the absolute risk-aversion parameter, we report all mean-variance utilities scaled by multiplying them with 2γ. Our main observation is that in the absence of price-impact costs, HXZ4 is the best model. To see this, note first that the mean-variance utility delivered by the factors in the HXZ4 model is higher than those delivered by the factors in the other four low-dimensional models (CAPM, FFC4, FF5, and FF6). Moreover, the difference between the utility provided by the factors in the HXZ4 model and those derived by the CAPM and FFC4 models are statistically significant at the 1% level. Also, although the differences between the utility derived from the factors in the HXZ4 model and those derived from the factors in the FF5 and FF6 models are

[11]For every non-nested model comparison, we find in unreported results that the first-stage test rejects the null hypothesis at the 1% level, and thus we have to perform the second-stage test.

[12]In detail, the p-values are computed as follows. Assume without loss of generality that the sample mean-variance utilities net of price-impact costs for models A and B satisfy $\hat{U}_A^\gamma > \hat{U}_B^\gamma$. Then, we compute the p-value as the integral over the values greater than $\hat{U}_A^\gamma - \hat{U}_B^\gamma$ of the probability density function in (23) if the two models are non-nested and of the probability density function in (24) if they are nested. Like Barillas et al. (2020), we use the bias-adjusted values of $\hat{U}_A^\gamma$ and $\hat{U}_B^\gamma$ when comparing non-nested factor models using Proposition 5. This is because the asymptotic distribution in (23) fails to capture the finite-sample bias in estimates of mean-variance utility. Section IA.2 of the Internet Appendix details the procedure we use to adjust the bias. However, when using Proposition 6 to compare nested factor models, we use the raw values of $\hat{U}_A^\gamma$ and $\hat{U}_B^\gamma$ because the asymptotic distribution in (24) adequately captures the finite-sample bias of the sample mean-variance utility. This is also demonstrated by the bootstrap experiments in Section IA.3 of the Internet Appendix.

Table 4: Significance of difference in mean-variance utility without price-impact costs

This table reports the significance of the difference between the mean-variance utilities of the row and column models in the absence of trading costs. Panel A reports the scaled sample mean-variance utility of each of the six factor models in the absence of trading costs. Panel B reports the p-value for the difference in mean-variance utility for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model, for the case when the expected price-impact matrix $\Lambda = 0$.

Panel A: Mean-variance utilities without trading costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma \text{MVU}^{\gamma}$	0.0216	0.1334	0.0546	0.1012	0.1138	0.1569

Panel B: p-values

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.000	0.003	0.000	0.000	0.000
HXZ4			0.002	0.099	0.176	0.285
FFC4				0.049	0.000	0.005
FF5					0.035	0.077
FF6						0.137

not statistically significant, HXZ4 is the preferred model because it contains fewer factors than FF5 and FF6, and thus, it is more parsimonious. Finally, the high-dimensional model DMNU20 achieves a sample mean-variance utility that is higher than that delivered by the factors in the HXZ4 model, but the difference in utilities is not statistically significant and thus HXZ4 is again the preferred model because of its parsimony. Overall, we conclude that the HXZ4 model best spans the investment opportunity set in the absence of costs.

Table 5 reports the performance of the six models in the *presence* of price impact for our base-case absolute risk-aversion parameter $\gamma = 10^{-9}$. Our main finding is that price-impact costs change the relative performance of the different models. Specifically, while HXZ4 was the best model in the absence of trading costs, it is significantly outperformed by FF6 and DMNU20 in the presence of price-impact costs. The explanation for the poor performance of the HXZ4 model in the presence of price impact is not only that its investment and profitability factors require higher turnover than those corresponding to the FF5 and FF6 models as shown in the seventh column of Table 3, but also that they require trading stocks with smaller market capitalization, and thus, less liquid as shown in the ninth column of Table 3. Consequently, investing in the factors in the HXZ4 model incurs high price-impact costs. FF6 emerges as the best low-dimensional model in the presence of price-impact costs

Table 5: Significance of difference in mean-variance utility with price-impact costs

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the baseline case with absolute risk-aversion parameter $\gamma = 10^{-9}$. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma \text{MVU}_{\Lambda}^{\gamma}$	0.0216	0.0414	0.0387	0.0577	0.0695	0.0982

Panel B: p-values

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.000	0.005	0.000	0.000	0.000
HXZ4			0.386	0.078	0.019	0.005
FFC4				0.077	0.000	0.006
FF5					0.003	0.033
FF6						0.086

because it significantly outperforms the CAPM, HXZ4, FFC4, and FF5 models.[13] Finally, although the high-dimensional model DMNU20 achieves higher sample mean-variance utility than the FF6 model, the difference of utilities between the FF6 model and the DMNU20 model is not statistically significant at the 5% level, and thus FF6 is the preferred model because of its parsimony. Overall, while the HXZ4 model was the best at spanning the investment opportunity set in the absence of costs, the FF6 model is best at spanning the achievable investment opportunity set in the presence of price-impact costs.

The finding that DMNU20 does *not* significantly outperform FF6 for the base case with $\gamma = 10^{-9}$ is surprising because DeMiguel et al. (2020) find that in the presence of trading costs, high-dimensional models are likely to perform well because of the benefits of trading diversification across factors. Moreover, DMNU20 includes value-weighted factors obtained from single sorts that, as shown in Table 3, trade liquid stocks with high market

[13]This result is counterintuitive because the FF6 model is obtained by adding the momentum factor to FF5 and trading the momentum factor incurs high price-impact costs as illustrated in the sixth column of Table 3. However, even though momentum is expensive when traded in isolation, it is a lot cheaper to trade in *combination* with the other five factors in the FF6 model because of trading diversification (DeMiguel et al., 2020). Indeed, Section IA.5 of the Internet Appendix reports summary statistics of the optimal portfolio weights for the different factor models, and shows that trading diversification greatly reduces the price-impact cost incurred by relatively higher-dimensional models such as FF6 and DMNU20.

Table 6: Significance of difference in mean-variance utility with costs for $\gamma = 10^{-10}$

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the case with absolute risk-aversion parameter $\gamma = 10^{-10}$. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma \text{MVU}_\Lambda^\gamma$	0.0215	0.0238	0.0242	0.0293	0.0317	0.0560

Panel B: p-values

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.017	0.056	0.004	0.001	0.000
HXZ4			0.411	0.077	0.024	0.003
FFC4				0.062	0.000	0.004
FF5					0.001	0.007
FF6						0.010

capitalization. To shed light on this result, we consider a case with a lower absolute risk aversion $\gamma = 10^{-10}$, which corresponds to an investor with the same relative risk aversion as in our base case, but with an endowment ten times higher than that in the base case. For this level of absolute risk aversion, price-impact costs should play a more important role and we expect that DMNU20 dominates other factor models. Table 6 confirms this intuition: the high-dimensional model DMNU20 significantly outperforms every low-dimensional model at the 1% confidence level. In addition, among low-dimensional models, FF6 is again the best model as it significantly outperforms the CAPM, HXZ4, FFC4, and FF5 models.

Finally, Table 7 reports the results for the case with a higher absolute risk-aversion parameter, $\gamma = 10^{-8}$, which corresponds to an investor with the same relative risk aversion as in the base case, but with an endowment ten times *lower* than that in the base case. For this case, price-impact costs are less important, and thus, we expect the relative performance of the different models to be similar to that in the *absence* of costs. Table 7 confirms this intuition: HXZ4 outperforms the CAPM, FFC4, and FF5 models, with the utility difference being statistically significant at the 1% level for CAPM and FFC4. Also, although FF6 and DMNU20 deliver higher mean-variance utilities net of price-impact costs than HXZ4, the

Table 7: Significance of difference in mean-variance utility with costs for $\gamma = 10^{-8}$

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the case with absolute risk-aversion parameter $\gamma = 10^{-8}$. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma \text{MVU}_{\Lambda}^{\gamma}$	0.0216	0.0984	0.0519	0.0904	0.1043	0.1297

Panel B: p-values

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.000	0.003	0.000	0.000	0.000
HXZ4			0.007	0.325	0.365	0.144
FFC4				0.058	0.000	0.007
FF5					0.021	0.093
FF6						0.198

difference between the utilities of these two models and HXZ4 is not statistically significant. Thus, HXZ4 emerges as the best model just as in the case without trading costs.

In summary, accounting for price-impact costs results in a more nuanced comparison of the various factor models we consider— HXZ4, FF6, and DMNU20 are the best models at spanning the achievable investment opportunities of investors with high, medium, and low absolute risk aversion, respectively.

4.5 Model comparison using out-of-sample bootstrap tests

In the previous section, we compared factor models using our proposed statistical tests, which address the main asset-pricing question: is the mean-variance utility in the presence of price-impact costs of a model significantly higher than that of another? As a robustness check, we now address a different question that is relevant for investment management: are the utility gains of a superior factor model achievable out of sample? To do this, we use the out-of-sample bootstrap test used by Fama and French (2018) and Detzel et al. (2023).

This bootstrap test guarantees that disjoint sets of observations are used for the in-sample and out-of-sample calculations. For each bootstrap sample, we carry out a four-step procedure. First, for every pair of consecutive months, we randomly assign one month to the set of in-sample (IS) observations and the other to the set of out-of-sample (OOS) observations. Second, within the IS set, we bootstrap with replacement a set with the same number of observations as the original sample, and allocate the corresponding partner months to the OOS set. Third, we use the factor returns and the factor-rebalancing trades of the months in the bootstrap IS set to calculate the optimal portfolio weights of each model using Equation (13).[14] Fourth, we apply the optimal portfolio weights from the third step to the bootstrap OOS set to obtain the OOS mean-variance utility net of price-impact costs for each factor model. We repeat these four steps 100,000 times, and thus, obtain 100,000 observations of the OOS mean-variance utility net of price-impact costs for each model. Finally, we compare models in terms of average mean-variance utility and the frequency with which one model outperforms another model across the bootstrap samples. This procedure not only guarantees that the IS and OOS sets for each bootstrap sample are disjoint, but also prevents the IS and OOS sets from having substantially different time-series properties because they are obtained from pairs of consecutive months.

Table 8 reports the out-of-sample bootstrap results for the base case with absolute risk aversion $\gamma = 10^{-9}$. Panel A reports the average mean-variance utility net of price-impact costs of each model and Panel B reports the frequency with which the row model outperforms the column model across the bootstrap samples.[15] As expected, the average *out-of-sample* mean-variance utilities of the different models in Panel A of Table 8 are much lower than the *in-sample* utilities in Panel A of Table 5 because of estimation error. However, the out-of-sample relative performance of the various models is generally consistent with that in sample.[16]

Note that the frequencies in Panel B of Table 8 are larger than the p-values based on our statistical tests in Panel B of Table 5. This is not surprising because even if a model

[14]We estimate the vector of factor-mean returns, μ, and the price-impact cost matrix, Λ, using their sample counterparts. For the covariance matrix of factor returns, Σ, we use the shrinkage estimator of Ledoit and Wolf (2004) to alleviate the impact of estimation error on the out-of-sample performance of the different models.

[15]Section IA.4 of the Internet Appendix reports the results for the cases with $\gamma = 10^{-10}$ and $\gamma = 10^{-8}$.

[16]There is one pairwise comparison of factor models for which the out-of-sample performance results differ from those in sample. In particular, the out-of-sample performance of the FF5 model is better than that of DMNU20, whereas the in-sample performance of DMNU20 was significantly better. Again, this is not surprising as the performance of the high-dimensional DMNU20 model is likely to be more impacted by estimation error out of sample than that of the FF5 model.

Table 8: Bootstrap out-of-sample utility net of price-impact costs

Panel A reports the average out-of-sample (OOS) scaled mean-variance utility net of price-impact costs across 100,000 bootstrap samples of each factor model under the baseline case with absolute risk-aversion parameter $\gamma = 10^{-9}$. Panel B reports the frequency with which the row model outperforms the column model out-of-sample across the 100,000 bootstrap samples.

Panel A: Average mean-variance utilities net of price-impact costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma\text{MVU}_{\Lambda}^{\gamma}$	0.0122	0.0247	0.0156	0.0318	0.0398	0.0265

Panel B: Frequency row model outperforms column model

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.167	0.360	0.213	0.186	0.355
HXZ4			0.664	0.310	0.235	0.439
FFC4				0.224	0.099	0.376
FF5					0.226	0.504
FF6						0.578

has a significantly higher mean-variance utility than another, it may deliver a lower out-of-sample mean-variance utility in a particular bootstrap sample because of estimation error. Nonetheless, the results in Panel B of Table 8 are generally consistent with those in Panel B of Table 5. In particular, we observe that, out of sample, HXZ4 outperforms FF5, FF6, and DMNU20 only on 31%, 23.5%, and 43.9% of the bootstrap samples, respectively. This is consistent with the finding in Panel B of Table 5 that the FF6 and DMNU20 models deliver higher mean-variance utility net of price-impact costs than the HXZ4 model. In addition, FF6 outperforms the CAPM, HXZ4, FFC4, and FF5 models on around 81%, 76%, 90%, and 77% of the bootstrap samples, respectively, which is consistent with the finding in Panel B of Table 5 that the FF6 model outperforms all other low-dimensional models. Finally, the FF6 model outperforms the DMNU20 model on 57.8% of the bootstrap samples, which again is coherent with our finding in Panel B of Table 8 that FF6 and DMNU20 are statistically indistinguishable.

In summary, the out-of-sample bootstrap tests confirm the main finding from our statistical tests in Table 5 that, in the base case with absolute risk-aversion parameter $\gamma = 10^{-9}$, the FF6 model emerges as the best model. Moreover, the out-of-sample test shows that the gains from using the FF6 factor model can actually be realized out of sample. Section IA.4 of the Internet Appendix shows that the findings from the out-of-sample bootstrap tests are also consistent with the findings from our statistical tests for the cases with lower and higher absolute risk-aversion parameters.

5 Conclusion

We show that the squared Sharpe ratio criterion is no longer sufficient to compare asset-pricing models in the presence of price impact because the efficient frontier spanned by a factor model is strictly concave. Instead, we propose comparing the ability of factor models to span the achievable investment opportunity set in terms of mean-variance utility net of price-impact costs and develop a formal statistical methodology to compare nested and non-nested factor models. Importantly, we observe that the relative performance of factor models depends on the absolute risk-aversion parameter, and thus comparing factor models in the presence of price impact is a more nuanced exercise than in the absence of trading costs.

Empirically, we find that while in the absence of trading costs the four-factor model of Hou et al. (2015) outperforms other low-dimensional models, in the presence of price-impact costs the six-factor model of Fama and French (2018) performs better. We also find that the high-dimensional model of DeMiguel et al. (2020) significantly outperforms the low-dimensional models *only* for the case with low absolute risk aversion, where price impact is important enough for the trading diversification benefits of combining a large number of factors to dominate other effects such as the impact of estimation error.

An implication of our work is that different benchmark factor models should be used to evaluate the performance of investment strategies designed for different investors, depending on their absolute risk aversion. Our proposed statistical test can be used not only to compare factor models, but also to evaluate the significance of the increase in mean-variance utility net of price impact-costs that an investor can achieve by having access to a particular investment strategy in addition to the factors in the benchmark model.

References

Avramov, Doron, and John C. Chao, 2006, An exact Bayes test of asset pricing models with application to international markets, *Journal of Business* 79, 293–324.

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.

Barroso, Pedro, and Andrew Detzel, 2021, Do limits to arbitrage explain the benefits of volatility-managed portfolios?, *Journal of Financial Economics* 140, 744–767.

Barroso, Pedro, and Pedro Santa-Clara, 2015, Beyond the carry trade: Optimal currency portfolios, *Journal of Financial and Quantitative Analysis* 50, 1037–1056.

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Campbell, John Y., 2017, *Financial decisions and markets: A course in asset pricing* (Princeton University Press, Princeton, NJ).

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chen, Andrew, and Mihail Velikov, 2022, Zeroing in on the expected returns of anomalies, forthcoming in *Journal of Financial and Quantitative Analysis*.

Chib, Siddhartha, Xiaming Zeng, and Lingxiao Zhao, 2020, On comparing asset pricing models, *Journal of Finance* 75, 551–577.

Cochrane, John H., 2009, *Asset Pricing: Revised Edition* (Princeton University Press).

DeMiguel, Victor, Alberto Martin-Utrera, Francisco J. Nogales, and Raman Uppal, 2020, A transaction-cost perspective on the multitude of firm characteristics, *Review of Financial Studies* 33, 2180–2222.

DeMiguel, Victor, Alberto Martin-Utrera, and Raman Uppal, 2022, A multifactor perspective on volatility-managed portfolios, available at SSRN 3982504.

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

Edelen, Roger M., Richard B. Evans, and Gregory B. Kadlec, 2007, Scale effects in mutual fund performance: The role of trading costs, *Available at SSRN* 951367.

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, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

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

Ferson, Wayne E., Andrew F. Siegel, and Junbo L. Wang, 2019, Asymptotic variances for tests of portfolio efficiency and factor model comparisons with conditioning information, *Available at SSRN* 3330663.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2015, Trading costs of asset pricing anomalies, Fama-Miller working paper.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2018, Trading costs, *Available at SSRN* 3229719.

Gârleanu, Nicolae, and Lasse H. Pedersen, 2013, Dynamic trading with predictable returns and transaction costs, *Journal of Finance* 68, 2309–2340.

Gârleanu, Nicolae, and Lasse H. Pedersen, 2022, Active and passive investing: Understanding Samuelson's dictum, forthcoming in *Review of Asset Pricing Studies*.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, Zhongzhi Lawrence He, and Sahn-Wook Huh, 2018, Distance-based metrics: A Bayesian solution to the power and extreme-error problems in asset-pricing tests, Swiss Finance Institute Research Paper.

Hayashi, Fumio, 2000, *Econometrics* (Princeton University Press).

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

Jensen, Theis I., Bryan T. Kelly, Semyon Malamud, and Lasse H. Pedersen, 2022, Machine learning and the implementable efficient frontier, *Available at SSRN* 4187217.

Jobson, J. D., and B. Korkie, 1980, Estimation for markowitz efficient portfolios, *Journal of the American Statistical Association* 75, 544–554.

Kan, Raymond, and Cesare Robotti, 2008, Specification tests of asset pricing models using excess returns, *Journal of Empirical Finance* 15, 816–838.

Kan, Raymond, and Cesare Robotti, 2009, Model comparison using the Hansen-Jagannathan distance, *Review of Financial Studies* 22, 3449–3490.

Kan, Raymond, Cesare Robotti, and Jay Shanken, 2013, Pricing model performance and the two-pass cross-sectional regression methodology, *Journal of Finance* 68, 2617–2649.

Kan, Raymond, Xiaolu Wang, and Xinghua Zheng, 2019, In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models, *Available at SSRN* 3454628.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039–1082.

Ledoit, Olivier, and Michael Wolf, 2004, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.

Lee, Charles MC, and Mark J Ready, 1991, Inferring trade direction from intraday data, *The Journal of Finance* 46, 733–746.

Lintner, John, 1965, Security, risk, and maximal gains from diversification, *Journal of Finance* 20, 587–615.

Maruyama, Yosihito, and Takashi Seo, 2003, Estimation of moment parameter in elliptical distributions, *Journal of the Japan Statistical Society* 33, 215–229.

Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *Journal of Finance* 72, 1611–1644.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Ross, Stephen, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

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

A Proofs of all results

This appendix contains the proofs of all novel propositions in the manuscript, with the exception of Proposition 4, which is proven and discussed in Appendix B. For expositional purposes, we also state in Proposition 1 a well-known result that is proven, for instance, in Campbell (2017, Section 2.2.6).

A.1 Proof of Proposition 2

Note that the proportional-trading-cost function given in Definition 1 is not convex in general and this complicates the proof, which consists of two parts. Part (i) shows that there exists a nonzero maximizer to the mean-variance problem. Part (ii) shows that the efficient frontier is a straight line.

Part (i): existence of a nonzero maximizer to mean-variance problem

We first show that for any absolute risk-aversion parameter γ, the objective function of problem (2) has a nonzero maximizer and its maximum is strictly positive.

Denote the mean-variance utility in problem (2) as

$$g_\gamma(\theta) = \theta^\top \mu - f(\theta) - \frac{\gamma}{2}\theta^\top \Sigma\theta.$$

By Assumption 2.3, we have that the set $S = \{\theta | \theta^\top \mu - f(\theta) \geq 0\}$ is nonempty. Moreover, by Assumption 2.2, $f(\theta)$ is continuous in S, and hence, S is compact. Furthermore, $g_\gamma(\theta)$ is also continuous in S, and thus, by the extreme-value theorem we have that there exists $\theta^* \in S$ such that $g_\gamma(\theta^*) \geq g_\gamma(\theta)$ for all $\theta \in S$. Also, by Assumption 2.3, we know that there are values of θ in S such that $g_\gamma(\theta) > 0$. Therefore, the maximum value, $g_\gamma(\theta^*)$, must be strictly positive. Consequently, $\theta^* \neq 0$ because $g_\gamma(0) = 0$.

Part (ii): the efficient frontier is a straight line

We first show by contradiction that if θ_1 is a maximizer for the case with absolute risk aversion γ, then for any $c > 0$ we have that $c\theta_1$ is a maximizer for the case with absolute risk aversion γ/c. Suppose $c\theta_1$ is not a maximizer for absolute risk aversion is γ/c, then there exists θ_2 such that

$$\theta_2^\top \mu - f(\theta_2) - \frac{\gamma}{2c}\theta_2^\top \Sigma \theta_2 > c\theta_1^\top \mu - f(c\theta_1) - \frac{\gamma}{2c} c\theta_1^\top \Sigma c\theta_1, \tag{32}$$

which is equivalent to

$$\frac{\theta_2^\top}{c}\mu - f\left(\frac{\theta_2}{c}\right) - \frac{\gamma}{2}\frac{\theta_2^\top}{c}\Sigma\frac{\theta_2}{c} > \theta_1^\top \mu - f(\theta_1) - \frac{\gamma}{2}\theta_1^\top \Sigma \theta_1, \tag{33}$$

which contradicts θ_1 being a maximizer for the case with absolute risk aversion γ. Note that this argument also shows that if θ_1 is a maximizer for the case with absolute risk aversion γ, then $c\theta_1$ with $c > 0$ is *not* a maximizer for the case with absolute risk aversion γ.

Next, we show by contradiction that given two maximizers θ_1 and θ_2 for the case with absolute risk aversion γ, we must have

$$\theta_1^\top \Sigma \theta_1 = \theta_2^\top \Sigma \theta_2, \tag{34}$$

and thus $\theta_1^\top \mu - f(\theta_1) = \theta_2^\top \mu - f(\theta_2)$. To see this, suppose without loss of generality that $\theta_2^\top \Sigma \theta_2 > \theta_1^\top \Sigma \theta_1$. Because both θ_1 and θ_2 are maximizers, by Part (i), we have $\theta_2^\top \mu - f(\theta_2) > \theta_1^\top \mu - f(\theta_1) > 0$. Thus, there exists $c > 1$, such that

$$c\theta_1^\top \mu - cf(\theta_1) = \theta_2^\top \mu - f(\theta_2). \tag{35}$$

Moreover, because we have shown that for $c > 0$, we have that $c\theta_1$ is not a maximizer for the case with absolute risk aversion γ, then we must have that

$$(c\theta_1^\top)\Sigma(c\theta_1) > \theta_2^\top \Sigma \theta_2. \tag{36}$$

Thus,

$$c\theta_1^\top \mu - cf(\theta_1) - \frac{\gamma}{2c}(c\theta_1^\top)\Sigma(c\theta_1) < \theta_2^\top \mu - f(\theta_2) - \frac{\gamma}{2c}\theta_2^\top \Sigma \theta_2, \tag{37}$$

which contradicts $c\theta_1$ being optimal for the case with absolute risk aversion is γ/c. Therefore, $\theta_1^\top \Sigma \theta_1 = \theta_2^\top \Sigma \theta_2$ and $\theta_2^\top \mu - f(\theta_2) = \theta_1^\top \mu - f(\theta_1)$, and thus, any two maximizers θ_1 and θ_2 for the case with absolute risk aversion γ must have the same Sharpe ratio.

We now show that the efficient frontier is a straight line by showing every efficient portfolio has the same Sharpe ratio, SR_{PTC}. The Sharpe ratio of $c\theta^*$, a maximizer for the case with absolute risk aversion γ/c is

$$\frac{c\theta^{*\top}\mu - f(c\theta^*)}{c\sqrt{\theta^{*\top}\Sigma\theta^*}} = \frac{\theta^{*\top}\mu - f(\theta^*)}{\sqrt{\theta^{*\top}\Sigma\theta^*}}, \tag{38}$$

which is also the Sharpe ratio of θ^*. Therefore, every efficient portfolio has the same Sharpe ratio of returns net of proportional trading costs, and thus the efficient frontier is a straight line starting at the origin of the standard deviation-mean diagram. Moreover, by Assumption 2.2 we have that $f(\theta) > 0$ for any $\theta \neq 0$, and thus,

$$SR_{PTC} = \frac{\theta^{*\top}\mu - f(\theta^*)}{\sqrt{\theta^{*\top}\Sigma\theta^*}} < \frac{\theta^{*\top}\mu}{\sqrt{\theta^{*\top}\Sigma\theta^*}} \leq SR.$$

A.2 Proof of Proposition 3

The proof consists of two parts. Part (i) provides an alternative condition to define a price-impact-cost function. Part (ii) shows that the efficient frontier is strictly concave.

Part (i): an alternative condition to define a price-impact-cost function

Definition 2 states that a price-impact-cost function must satisfy condition (8). We now show that this condition is equivalent to

$$f(c'\theta) < c'f(\theta) \quad \text{for } \theta \neq 0 \text{ and } 0 < c' < 1. \tag{39}$$

We first prove that (8) implies (39). Let $\theta' = c\theta$ with $c > 1$. Then (8) becomes

$$\frac{1}{c}f(\theta') > f\big(\frac{1}{c}\theta'\big). \tag{40}$$

If we define $c' = 1/c \in (0,1)$, then the previous inequality becomes

$$c'f(\theta') > f(c'\theta'), \tag{41}$$

which is (39). Using a similar argument, it is straightforward to show that (39) implies (8).

Part (ii): the efficient frontier is concave

Part (i) of the proof of Proposition 2 shows that for any γ, there exists a nonzero maximizer to problem (2). Let θ^* and θ^*_c be the maximizers to problem (2) for the cases with absolute risk aversion γ and $c\gamma$, respectively, where $0 < c < 1$. We first show that the variance of portfolio θ^*_c is greater than or equal to that of portfolio θ^*. We then show that the Sharpe ratio of θ^*_c is strictly lower than that of θ^* when the variance of θ^*_c is strictly greater than that of θ^*, and thus the efficient frontier is strictly concave.

Step 1: the variance of θ_c^* is greater than or equal to that of θ^*.

We show by contradiction that $(\theta_c^*)^\top \Sigma \theta_c^* \geq \theta^{*\top} \Sigma \theta^*$. Suppose $(\theta_c^*)^\top \Sigma \theta_c^* < \theta^{*\top} \Sigma \theta^*$. The optimality of θ^* and θ_c^* for the cases with absolute risk aversion γ and $c\gamma$, respectively, implies that

$$\theta^{*\top}\mu - f(\theta^*) - \frac{c\gamma}{2}\theta^{*\top}\Sigma\theta^* \leq (\theta_c^*)^\top \mu - f(\theta_c^*) - \frac{c\gamma}{2}(\theta_c^*)^\top \Sigma \theta_c^*, \tag{42}$$

$$(\theta_c^*)^\top \mu - f(\theta_c^*) - \frac{\gamma}{2}(\theta_c^*)^\top \Sigma \theta_c^* \leq \theta^{*\top}\mu - f(\theta^*) - \frac{\gamma}{2}\theta^{*\top}\Sigma\theta^*. \tag{43}$$

Combining these two inequalities yields

$$\frac{\gamma}{2}(\theta^{*\top}\Sigma\theta^* - (\theta_c^*)^\top \Sigma \theta_c^*) \leq \theta^{*\top}\mu - f(\theta^*) - (\theta_c^*)^\top \mu + f(\theta_c^*) \leq \frac{c\gamma}{2}(\theta^{*\top}\Sigma\theta^* - (\theta_c^*)^\top \Sigma \theta_c^*). \tag{44}$$

Because we have assumed that $(\theta_c^*)^\top \Sigma \theta_c^* < \theta^{*\top} \Sigma \theta^*$ and $0 < c < 1$, the leftmost term is strictly greater than the rightmost term in (44), and thus we have a contradiction. Therefore, we must have that $(\theta_c^*)^\top \Sigma \theta_c^* \geq \theta^{*\top} \Sigma \theta^*$.

Step 2: the Sharpe ratio of the portfolio θ_c^* is not greater than that of θ^*.

We show that

$$\frac{(\theta_c^*)^\top \mu - f(\theta_c^*)}{\sqrt{(\theta_c^*)^\top \Sigma \theta_c^*}} \leq \frac{\theta^{*\top}\mu - f(\theta^*)}{\sqrt{\theta^{*\top}\Sigma\theta^*}}, \tag{45}$$

and the equality holds only when $(\theta_c^*)^\top \Sigma \theta_c^* = \theta^{*\top} \Sigma \theta^*$.

When $(\theta_c^*)^\top \Sigma \theta_c^* = \theta^{*\top} \Sigma \theta^*$, (44) implies that $\theta^{*\top}\mu - f(\theta^*) = (\theta_c^*)^\top \mu - f(\theta_c^*)$, and thus (45) holds with equality.

When $(\theta_c^*)^\top \Sigma \theta_c^* > \theta^{*\top} \Sigma \theta^*$, let $(\theta_c^*)^\top \Sigma \theta_c^* = c^2 \theta^{*\top} \Sigma \theta^*$ where $c > 1$. To prove (45) with strict inequality, we prove by contradiction that

$$(\theta_c^*)^\top \mu - f(\theta_c^*) < c(\theta^{*\top}\mu - f(\theta^*)). \tag{46}$$

Suppose (46) does not hold and thus $\theta^{*\top}\mu - f(\theta^*) \leq ((\theta_c^*)^\top \mu - f(\theta_c^*))/c$, then

$$\begin{aligned}\theta^{*\top}\mu - f(\theta^*) - \frac{\gamma}{2}\theta^{*\top}\Sigma\theta^* &\leq \frac{1}{c}(\theta_c^*)^\top \mu - \frac{1}{c}f(\theta_c^*) - \frac{\gamma}{2}\frac{(\theta_c^*)^\top}{c}\Sigma\frac{\theta_c^*}{c} \\ &< \frac{1}{c}(\theta_c^*)^\top \mu - f(\frac{1}{c}\theta_c^*) - \frac{\gamma}{2}\frac{(\theta_c^*)^\top}{c}\Sigma\frac{\theta_c^*}{c},\end{aligned} \tag{47}$$

where the second inequality comes from Part (i). This contradicts θ^* being a maximizer for the case with absolute risk aversion is γ. Thus, when $(\theta_c^*)^\top \Sigma \theta_c^* > \theta^{*\top} \Sigma \theta^*$, (46) holds.

Dividing both sides of (46) by $\sqrt{(\theta_c^*)^\top \Sigma \theta_c^*} = c\sqrt{\theta^{*\top}\Sigma\theta^*}$, (45) holds with strict inequality. Therefore, the efficient frontier is strictly concave. Moreover, since $f(\theta) > 0$ for any $\theta \neq 0$ both sides of (45) are less than the Sharpe ratio in the absence of trading costs, SR.

A.3 Proof of Proposition 5

The proof consists of two parts. Part (i) derives the asymptotic distribution of the sample mean-variance utility net of price-impact costs of a factor model. Part (ii) derives the asymptotic distribution of the difference between the sample mean-variance utilities net of price-impact costs of two factor models. For ease of notation, we drop the superscript γ from MVU^γ throughout this proof.

Part (i): asymptotic distribution of sample mean-variance utility of one model

The proof of Part (i) contains two steps. We first show that the sample mean-variance utility of a model is asymptotically normally distributed and then derive the variance of the asymptotic normal distribution.

Step 1: $\sqrt{T}(\widehat{MVU} - MVU)$ is asymptotically normally distributed. We extend the notation in the proof of Proposition 2 of Barillas et al. (2020) to the case with price-impact costs. In particular, let

$$\varphi = [\mu, \text{vec}(\Sigma), \text{vec}(\Lambda/\gamma)] \in \mathbb{R}^{K+2K^2}, \tag{48}$$

$$\hat{\varphi} = [\hat{\mu}, \text{vec}(\hat{\Sigma}), \text{vec}(\hat{\Lambda}/\gamma)] \in \mathbb{R}^{K+2K^2}, \tag{49}$$

$$r_t(\varphi) = [F_t - \mu, \text{vec}(\Sigma_t - \Sigma), \text{vec}((\Lambda_t - \Lambda)/\gamma)] \in \mathbb{R}^{K+2K^2}. \tag{50}$$

Under standard regularity conditions[17], the central limit theorem implies that,

$$\sqrt{T}(\hat{\varphi} - \varphi) \overset{A}{\sim} N(0, S_0), \quad \text{where } S_0 = \sum_{j=-\infty}^{\infty} E[r_t(\varphi) r_{t+j}^\top(\varphi)].$$

Using the delta method, we have that

$$\sqrt{T}(\widehat{\text{MVU}} - \text{MVU}) \overset{A}{\sim} N(0, \frac{\partial \text{MVU}}{\partial \varphi^\top} S_0 \frac{\partial \text{MVU}}{\partial \varphi}). \tag{51}$$

[17] For example, we could assume that the returns and the rebalancing trades are stationary and ergodic, and the corresponding Gordin's condition is satisfied, as in Proposition 6.10 of Hayashi (2000)

Step 2: variance of asymptotic normal distribution, $h_t(\varphi)$.

Let

$$h_t(\varphi) = 2\gamma \frac{\partial \mathrm{MVU}}{\partial \varphi^\top} r_t(\varphi), \tag{52}$$

then (51) can be rewritten as

$$\sqrt{T}(\widehat{\mathrm{MVU}} - \mathrm{MVU}) \overset{A}{\sim} N(0, W), \quad \text{where } W = \sum_{j=-\infty}^{\infty} E\left[\frac{h_t(\varphi)h_{t+j}(\varphi)}{4\gamma^2}\right]. \tag{53}$$

Assumption 3.1 implies that $h_t(\varphi)$ is serially uncorrelated, and thus, we have that

$$W = E\left[\frac{h_t^2(\varphi)}{4\gamma^2}\right]. \tag{54}$$

Also, note that

$$\frac{\partial \mathrm{MVU}}{\partial \mu} = \frac{1}{\gamma}(\Sigma + \Lambda/\gamma)^{-1}\mu = \theta^*,$$

$$\frac{\partial \mathrm{MVU}}{\partial \Sigma} = \gamma \frac{\partial \mathrm{MVU}}{\partial \Lambda} = -\frac{1}{2\gamma}(\Sigma + \Lambda/\gamma)^{-1}\mu\mu^\top(\Sigma + \Lambda/\gamma)^{-1} = -\frac{\gamma}{2}\theta^*\theta^{*\top},$$

and thus,

$$\frac{\partial \mathrm{MVU}}{\partial \mathrm{vec}(\Sigma)} = \gamma \frac{\partial \mathrm{MVU}}{\partial \mathrm{vec}(\Lambda)} = -\frac{\gamma}{2}\theta^* \otimes \theta^*,$$

where $\otimes$ denotes the Kronecker product. Plugging these partial derivatives in the definition of $h_t(\varphi)$ in (52), we have that

$$\begin{aligned} h_t(\varphi) &= 2\gamma\left[\frac{\partial \mathrm{MVU}}{\partial \mu^\top}(F_t - \mu) + \frac{\partial \mathrm{MVU}}{\partial \mathrm{vec}(\Sigma)^\top}\mathrm{vec}(\Sigma_t - \Sigma) + \frac{\partial \mathrm{MVU}}{\partial \mathrm{vec}(\Lambda)^\top}\mathrm{vec}(\Lambda_t - \Lambda)\right] \\ &= 2\gamma\theta^{*\top}(F_t - \mu) - \gamma^2\theta^{*\top}\Sigma_t\theta^* - \gamma\theta^{*\top}\Lambda_t\theta^* + \gamma^2\theta^{*\top}\Sigma\theta^* + \gamma\theta^{*\top}\Lambda\theta^* \\ &= \mu^\top(\Sigma + \Lambda/\gamma)^{-1}(2F_t - \mu) - \mu^\top(\Sigma + \Lambda/\gamma)^{-1}(\Sigma_t + \Lambda_t/\gamma)(\Sigma + \Lambda/\gamma)^{-1}\mu, \end{aligned} \tag{55}$$

which completes the first part of the proof.

Part (ii): asymptotic distribution of difference between utilities of two models

Following the same steps as in Part (i), we have that

$$\sqrt{T}\big([\widehat{\mathrm{MVU}}_A - \widehat{\mathrm{MVU}}_B] - [\mathrm{MVU}_A - \mathrm{MVU}_B]\big) \overset{A}{\sim} N\left(0, \frac{\partial(\mathrm{MVU}_A - \mathrm{MVU}_B)}{\partial \varphi^\top} S_0 \frac{\partial(\mathrm{MVU}_A - \mathrm{MVU}_B)}{\partial \varphi}\right). \tag{56}$$

By Assumption 3.1, we have that

$$\sqrt{T}\left([\widehat{\mathrm{MVU}}_A - \widehat{\mathrm{MVU}}_B] - [\mathrm{MVU}_A - \mathrm{MVU}_B]\right) \overset{A}{\sim} N\left(0, E\left[\frac{(h_{t,A} - h_{t,B})^2}{4\gamma^2}\right]\right), \tag{57}$$

where $h_{t,A}$ and $h_{t,B}$ are obtained by applying Equation (52) to models A and B, respectively. This completes the proof.

Remark: When model A nests model B and the extra factors of model A are redundant, or when models A and B share common factors and the extra factors of both models are redundant, the two models have the same optimal factor portfolio. In either case, the null hypothesis $\mathrm{MVU}_A = \mathrm{MVU}_B$ holds and Equation (55) suggests that $h_{t,A} = h_{t,B}$ for all t, and thus the variance in (57), $E[(h_{t,A} - h_{t,B})^2/(4\gamma^2)] = 0$. Consequently, the distribution in (57) is not applicable to perform a statistical test in these cases. Instead, in these cases we use the asympotic distribution in Proposition 6.

A.4 Proof of Proposition 6

Let the mean-variance portfolio in the presence of price-impact costs for model A be $\theta_A^* = [\theta_1^*, \theta_2^*]$. Note that the null hypothesis that models A and B have the same mean-variance utility holds if and only if $\theta_2^* = 0$. Using this condition, we prove this proposition in three parts. Part (i) derives the asymptotic distribution of the sample factor portfolio $\hat{\theta}_A^*$. Part (ii) provides an expression for the difference between the mean-variance utilities net of price-impact costs of models A and B as a function of θ_2^*. Part (iii) uses the asymptotic distribution of $\hat{\theta}_2^*$ to derive the asymptotic distribution of the difference between the sample mean-variance utilities net of price-impact costs of models A and B. Similar to the proof of Proposition 5, we drop the superscript γ from MVU^γ throughout this proof.

Part (i): asymptotic distribution for $\hat{\theta}_A^*$.

Following similar steps as those in Part (i) of the proof of Proposition 5, the asymptotic distribution of $\hat{\theta}_A^*$ is

$$\sqrt{T}(\hat{\theta}_A^* - \theta_A^*) \overset{A}{\sim} N(0, \frac{E[l_t l_t^\top]}{\gamma^2}), \tag{58}$$

where

$$l_t = (\Sigma_A + \Lambda_A/\gamma)^{-1} R_{A,t} - (\Sigma_A + \Lambda_A/\gamma)^{-1}(\Sigma_{A,t} + \Lambda_{A,t}/\gamma)(\Sigma_A + \Lambda_A/\gamma)^{-1}\mu_A \in \mathbb{R}^{K_1+K_2}. \tag{59}$$

Part (ii): expression for $\mathbf{MVU}_A - \mathbf{MVU}_B$ as a function of θ_2^*.

The difference $\mathrm{MVU}_A - \mathrm{MVU}_B$ can be written as

$$= \frac{1}{2\gamma}\left[\mu_1^\top, \mu_2^\top\right]\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & \Sigma_{22}+\Lambda_{22}/\gamma\end{bmatrix}^{-1}\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix} \tag{60}$$

$$- \frac{1}{2\gamma}\left[\mu_1^\top, \mu_2^\top\right]\begin{bmatrix}(\Sigma_{11}+\Lambda_{11}/\gamma)^{-1} & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}$$

$$= \frac{\gamma}{2}\theta_A^{*\top}\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & \Sigma_{22}+\Lambda_{22}/\gamma\end{bmatrix}\theta_A^*$$

$$- \frac{\gamma}{2}\theta_A^{*\top}\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & \Sigma_{22}+\Lambda_{22}/\gamma\end{bmatrix}\begin{bmatrix}(\Sigma_{11}+\Lambda_{11}/\gamma)^{-1} & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & \Sigma_{22}+\Lambda_{22}/\gamma\end{bmatrix}\theta_A^*$$

$$= \frac{\gamma}{2}\theta_A^{*\top}\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & \Sigma_{22}+\Lambda_{22}/\gamma\end{bmatrix}\theta_A^*$$

$$- \frac{\gamma}{2}\theta_A^{*\top}\begin{bmatrix}\Sigma_{11}+\Lambda_{11}/\gamma & \Sigma_{12}+\Lambda_{12}/\gamma \\ \Sigma_{21}+\Lambda_{21}/\gamma & (\Sigma_{21}+\Lambda_{21}/\gamma)(\Sigma_{11}+\Lambda_{11}/\gamma)^{-1}(\Sigma_{12}+\Lambda_{12}/\gamma)\end{bmatrix}\theta_A^*$$

$$= \frac{\gamma}{2}\theta_2^{*\top}\left[(\Sigma_{22}+\Lambda_{22}/\gamma) - (\Sigma_{21}+\Lambda_{21}/\gamma)(\Sigma_{11}+\Lambda_{11}/\gamma)^{-1}(\Sigma_{12}+\Lambda_{12}/\gamma)\right]\theta_2^*$$

$$= \frac{\gamma}{2}\theta_2^{*\top}W\theta_2^*, \tag{61}$$

where $W = (\Sigma_{22}+\Lambda_{22}/\gamma) - (\Sigma_{21}+\Lambda_{21}/\gamma)(\Sigma_{11}+\Lambda_{11}/\gamma)^{-1}(\Sigma_{12}+\Lambda_{12}/\gamma)$. Replacing the population parameters in Equation (61) with their sample counterparts we have that

$$\widehat{\mathrm{MVU}}_A - \widehat{\mathrm{MVU}}_B = \frac{\gamma}{2}\hat{\theta}_2^{*\top}\hat{W}\hat{\theta}_2^*, \quad \text{where} \quad \hat{W} \overset{a.s.}{\to} W. \tag{62}$$

Part (iii): asymptotic distribution for $T(\widehat{\mathbf{MVU}}_A - \widehat{\mathbf{MVU}}_B)$.

We now use (58) and (62) to derive the asymptotic distribution for $T(\widehat{\mathrm{MVU}}_A - \widehat{\mathrm{MVU}}_B)$. Let

$$z = \lim_{T\to\infty}\sqrt{T}\left(\frac{E[l_t l_t^\top]_{22}}{\gamma^2}\right)^{-\frac{1}{2}}\hat{\theta}_2^*.$$

Under the null hypothesis that $\theta_2^* = 0$, from the asymptotic distribution in (58) we have that that $z \sim N(0, I_{K_2})$, where I_{K_2} is a K_2-dimensional identity matrix. Thus, from Equation (62) we have that

$$\begin{aligned} T(\widehat{\mathrm{MVU}}_A - \widehat{\mathrm{MVU}}_B) &= \frac{\gamma}{2}T\hat{\theta}_2^{*\top}\hat{W}\hat{\theta}_2^* \\ &\overset{A}{\sim} \frac{1}{2\gamma}z^\top(E[l_t l_t^\top]_{22})^{\frac{1}{2}}W(E[l_t l_t^\top]_{22})^{\frac{1}{2}}z. \end{aligned} \tag{63}$$

Let $Q\Xi Q^\top$ be the eigenvalue decomposition of $(E[l_t l_t^\top]_{22})^{\frac{1}{2}} W (E[l_t l_t^\top]_{22})^{\frac{1}{2}}/2\gamma$, where Q is the orthogonal matrix whose columns contain the eigenvectors and Ξ is a diagonal matrix whose diagonal elements contain the eigenvalues ξ_i for $i = 1, \ldots, K_2$. Note the eigenvalues in the diagonal of Ξ are also the eigenvalues of $E[l_t l_t^\top]_{22} W/2\gamma$. Let $\bar{z} = Q^\top z \sim N(0, I_{K_2})$, then (63) can be rewritten as

$$T(\widehat{\text{MVU}}_A - \widehat{\text{MVU}}_B) \overset{A}{\sim} \bar{z}^\top \Xi \bar{z} = \sum_{i=1}^{K_2} \xi_i x_i,$$

where x_i for $i = 1, \ldots, K_2$ are independent chi-square random variables with one degree of freedom.

A.5 Proof of Proposition 7

The proof consists of two parts. Part (i) derives a closed-form expression for the asymptotic variance of the sample mean-variance utility of a factor model. Part (ii) derives a closed-form expression for the asymptotic variance of the difference between the sample mean-variance utilities of two factor models.

Part (i): closed-form asymptotic variance of the mean-variance utility of a model

We first provide a closed-form expression for the asymptotic variance of the sample mean-variance utility of a model, $E[h_t^2]/(4\gamma^2)$, and then simplify this expression.

Step 1: express $E[h_t^2]$ as a function of u_t, $v_{n,t}$, and $\bar{u} = E[u_t]$.
Plugging $\bar{u}, u_t$, and $v_{n,t}$ into (22), we have that

$$h_t = 2(u_t - \bar{u}) - \left[(u_t - \bar{u})^2 + \sum_{n=1}^{N} v_{n,t}^2\right] + \bar{u}.$$

Therefore,

$$\begin{aligned} E[h_t^2] =& E\Bigg[4(u_t - \bar{u})^2 - 4(u_t - \bar{u})^3 - 4(u_t - \bar{u}) \sum_{n=1}^{N} v_{n,t}^2 + 4(u_t - \bar{u})\bar{u} \\ &+ (u_t - \bar{u})^4 + 2(u_t - \bar{u})^2 \sum_{n=1}^{N} v_{n,t}^2 - 2(u_t - \bar{u})^2 \bar{u} \\ &+ \left(\sum_{n=1}^{N} v_{n,t}^2\right)^2 - 2\bar{u} \sum_{n=1}^{N} v_{n,t}^2 + \bar{u}^2\Bigg]. \end{aligned} \tag{64}$$

Lemma 2 of Maruyama and Seo (2003) shows that if (X_i, X_j, X_k, X_l) are jointly normally distributed with zero mean, then

$$E[X_i X_j X_k] = 0, \tag{65}$$

$$E[X_i X_j X_k X_l] = (\sigma_{ij}\sigma_{kl} + \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}), \tag{66}$$

where σ_{ab} is the covariance between X_a and X_b. Because $(u_t - \bar{u})$ and $v_{n,t}$ for $n = 1, \ldots, N$ are jointly normally distributed, using Equation (65), we can drop the third-order moments from Equation (64) to obtain

$$\begin{aligned} E[h_t^2] =& E\Bigg[4(u_t - \bar{u})^2 + (u_t - \bar{u})^4 + 2(u_t - \bar{u})^2 \sum_{n=1}^{N} v_{n,t}^2 - 2(u_t - \bar{u})^2 \bar{u} \\ &+ \left(\sum_{n=1}^{N} v_{n,t}^2\right)^2 - 2\bar{u}\sum_{n=1}^{N} v_{n,t}^2 + \bar{u}^2\Bigg]. \end{aligned} \tag{67}$$

Step 2: simplify (67). Using Equation (66), we can rewrite the terms on the right-hand side of Equation (67) as

$$\begin{aligned} E\left[(u_t - \bar{u})^2\right] &= \text{var}(u_t) = \mu^\top(\Sigma + \Lambda/\gamma)^{-1}\Sigma(\Sigma + \Lambda/\gamma)^{-1}\mu, \\ E\left[(u_t - \bar{u})^4\right] &= 3\left[\text{var}(u_t)\right]^2, \\ E\left[\sum_{n=1}^{N} v_{n,t}^2\right] &= \sum_{n=1}^{N} \text{var}(v_{n,t}) = \mu^\top(\Sigma + \Lambda/\gamma)^{-1}(\Lambda/\gamma)(\Sigma + \Lambda/\gamma)^{-1}\mu, \\ E\left[(u_t - \bar{u})^2 \sum_{n=1}^{N} v_{n,t}^2\right] &= E\left[(u_t - \bar{u})^2\right] \sum_{n=1}^{N} E\left[v_{n,t}^2\right] + 2\sum_{n=1}^{N}\left(E\left[(u_t - \bar{u})v_{n,t}\right]\right)^2 \\ &= \text{var}(u_t)\sum_{n=1}^{N}\text{var}(v_{n,t}) + 2\sum_{n=1}^{N}\left[\text{cov}(u_t, v_{n,t})\right]^2, \\ E\left[\Big(\sum_{n=1}^{N} v_{n,t}^2\Big)^2\right] &= \sum_{i=1}^{N}\sum_{j=1}^{N}\Big(\text{var}(v_{i,t})\text{var}(v_{j,t}) + 2\big[\text{cov}(v_{i,t}, v_{j,t})\big]^2\Big), \\ \bar{u} &= \mu^\top(\Sigma + \Lambda/\gamma)^{-1}\mu = \text{var}(u_t) + \sum_{n=1}^{N}\text{var}(v_{n,t}). \end{aligned}$$

Plugging these equations into (67), we have that

$$E[h_t^2] = 4\text{var}(u_t) + 3\left[\text{var}(u_t)\right]^2 + 2\left(\text{var}(u_t)\sum_{n=1}^{N}\text{var}(v_{n,t}) + 2\sum_{n=1}^{N}\left[\text{cov}(u_t, v_{n,t})\right]^2\right)$$

$$- 2\text{var}(u_t)\left(\text{var}(u_t) + \sum_{n=1}^{N}\text{var}(v_{n,t})\right) + \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\text{var}(v_{i,t})\text{var}(v_{j,t}) + 2\left[\text{cov}(v_{i,t}, v_{j,t})\right]^2\right)$$

$$- 2\sum_{n=1}^{N}\text{var}(v_{n,t})\left(\text{var}(u_t) + \sum_{n=1}^{N}\text{var}(v_{n,t})\right) + \left(\text{var}(u_t) + \sum_{n=1}^{N}\text{var}(v_{n,t})\right)^2$$

$$= 4\text{var}(u_t) + 2\left[\text{var}(u_t)\right]^2 - \left(\sum_{n=1}^{N}\text{var}(v_{n,t})\right)^2 + 4\sum_{n=1}^{N}\left[\text{cov}(u_t, v_{n,t})\right]^2$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\text{var}(v_{i,t})\text{var}(v_{j,t}) + 2\left[\text{cov}(v_{i,t}, v_{j,t})\right]^2\right)$$

$$= 4\text{var}(u_t) + 2\left[\text{var}(u_t)\right]^2 + 4\sum_{n=1}^{N}\left[\text{cov}(u_t, v_{n,t})\right]^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\text{cov}(v_{i,t}, v_{j,t})\right]^2.$$

Part (ii): asymptotic variance for difference between utilities of two models

The asymptotic variance of the difference between the sample mean-variance utilities of two models is

$$\frac{E[(h_{t,A} - h_{t,B})^2]}{4\gamma^2} = \frac{1}{4\gamma^2}\left(E[h_{t,A}^2] + E[h_{t,B}^2] - 2E[h_{t,A}h_{t,B}]\right). \tag{68}$$

The closed-form expressions of $E[h_{t,A}^2]$ and $E[h_{t,B}^2]$ are given in Part (i), and thus we focus on finding the closed-form expression of $E[h_{t,A}h_{t,B}]$. Similar to Part (i), we first express $E[h_{t,A}h_{t,B}]$ as a function of $\bar{u}$, u_t, and $v_{n,t}$, and then simplify this expression.

Step 1: express $E[h_{t,A}h_{t,B}]$ as a function of $\bar{u}$, u_t, and $v_{n,t}$.
Because $\left(u_t^A - \bar{u}^A\right)$, $\left(u_t^B - \bar{u}^B\right)$, $v_{n,t}^A$, and $v_{n,t}^B$ for $n = 1, \ldots, N$ are jointly normally distributed. Using Equation (65), we have that

$$E[h_{t,A}h_{t,B}] = E\Bigg[4\left(u_t^A - \bar{u}^A\right)\left(u_t^B - \bar{u}^B\right) + \left(u_t^A - \bar{u}^A\right)^2\left(u_t^B - \bar{u}^B\right)^2$$

$$+ \left(u_t^A - \bar{u}^A\right)^2\sum_{n=1}^{N}(v_{n,t}^B)^2 + \left(u_t^B - \bar{u}^B\right)^2\sum_{n=1}^{N}\left(v_{n,t}^A\right)^2$$

$$- \left(u_t^A - \bar{u}^A\right)^2\bar{u}^B - \left(u_t^B - \bar{u}^B\right)^2\bar{u}^A + \left(\sum_{n=1}^{N}(v_{n,t}^A)^2\right)\left(\sum_{n=1}^{N}\left(v_{n,t}^B\right)^2\right)$$

$$- \bar{u}^A\sum_{n=1}^{N}\left(v_{n,t}^B\right)^2 - \bar{u}^B\sum_{n=1}^{N}\left(v_{n,t}^A\right)^2 + \bar{u}^A\bar{u}^B\Bigg], \tag{69}$$

Step 2: simplify (69). Using Equation (66), we can rewrite the terms on the right-hand side of Equation (69) as

$$E\left[\left(u_t^A-\bar{u}^A\right)\left(u_t^B-\bar{u}^B\right)\right]=\operatorname{cov}\left(u_t^A,u_t^B\right),$$

$$E\left[\left(u_t^A-\bar{u}^A\right)^2\left(u_t^B-\bar{u}^B\right)^2\right]=\operatorname{var}\left(u_t^A\right)\operatorname{var}\left(u_t^B\right)+2\left[\operatorname{cov}\left(u_t^A,u_t^B\right)\right]^2,$$

$$E\left[\sum_{n=1}^N\left(v_{n,t}^A\right)^2\right]=\sum_{n=1}^N\operatorname{var}\left(v_{n,t}^A\right),$$

$$E\left[\sum_{n=1}^N\left(v_{n,t}^B\right)^2\right]=\sum_{n=1}^N\operatorname{var}\left(v_{n,t}^B\right),$$

$$E\Big[(u_t^A-\bar{u}^A)^2\sum_{n=1}^N(v_{n,t}^B)^2\Big]=\operatorname{var}(u_t^A)\sum_{n=1}^N\operatorname{var}(v_{n,t}^B)+2\sum_{n=1}^N\left[\operatorname{cov}(u_t^A,v_{n,t}^B)\right]^2,$$

$$E\Big[(u_t^B-\bar{u}^B)^2\sum_{n=1}^N(v_{n,t}^A)^2\Big]=\operatorname{var}(u_t^B)\sum_{n=1}^N\operatorname{var}(v_{n,t}^A)+2\sum_{n=1}^N\left[\operatorname{cov}(u_t^B,v_{n,t}^A)\right]^2,$$

$$E\Big[\Big(\sum_{n=1}^N(v_{n,t}^A)^2\Big)\Big(\sum_{n=1}^N(v_{n,t}^B)^2\Big)\Big]=\sum_{i=1}^N\sum_{j=1}^N\Big(\operatorname{var}(v_{i,t}^A)\operatorname{var}(v_{j,t}^B)+2\big[\operatorname{cov}(v_{i,t}^A,v_{j,t}^B)\big]^2\Big),$$

$$\bar{u}^A=\operatorname{var}(u_t^A)+\sum_{n=1}^N\operatorname{var}(v_{n,t}^A),$$

$$\bar{u}^B=\operatorname{var}(u_t^B)+\sum_{n=1}^N\operatorname{var}(v_{n,t}^B).$$

Plugging these equations into Equation (69), we have that

$$\begin{aligned}E[h_{t,A}h_{t,B}]=&\,4\operatorname{cov}(u_t^A,u_t^B)+\operatorname{var}(u_t^A)\operatorname{var}(u_t^B)+2\big[\operatorname{cov}(u_t^A,u_t^B)\big]^2\\&+\operatorname{var}(u_t^A)\sum_{n=1}^N\operatorname{var}(v_{n,t}^B)+2\sum_{n=1}^N\left[\operatorname{cov}(u_t^A,v_{n,t}^B)\right]^2\\&+\operatorname{var}(u_t^B)\sum_{n=1}^N\operatorname{var}(v_{n,t}^A)+2\sum_{n=1}^N\left[\operatorname{cov}(u_t^B,v_{n,t}^A)\right]^2\\&-\operatorname{var}(u_t^A)\Big(\operatorname{var}(u_t^B)+\sum_{n=1}^N\operatorname{var}(v_{n,t}^B)\Big)-\operatorname{var}(u_t^B)\Big(\operatorname{var}(u_t^A)+\sum_{n=1}^N\operatorname{var}(v_{n,t}^A)\Big)\\&+\sum_{i=1}^N\sum_{j=1}^N\Big(\operatorname{var}(v_{i,t}^A)\operatorname{var}(v_{j,t}^B)+2\big[\operatorname{cov}(v_{i,t}^A,v_{j,t}^B)\big]^2\Big)\end{aligned}$$

$$
\begin{aligned}
&- \left(\sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{B})\right)\left(\mathrm{var}(u_t^{A}) + \sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{A})\right) \\
&- \left(\sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{A})\right)\left(\mathrm{var}(u_t^{B}) + \sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{B})\right) \\
&+ \left(\mathrm{var}(u_t^{A}) + \sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{A})\right)\left(\mathrm{var}(u_t^{B}) + \sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{B})\right) \\
&= 4\mathrm{cov}(u_t^{A}, u_t^{B}) + 2\big[\mathrm{cov}(u_t^{A}, u_t^{B})\big]^2 - \left(\sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{A})\right)\left(\sum_{n=1}^{N} \mathrm{var}(v_{n,t}^{B})\right) \\
&+ 2\sum_{n=1}^{N} \left(\big[\mathrm{cov}(u_t^{A}, v_{n,t}^{B})\big]^2 + \big[\mathrm{cov}(u_t^{B}, v_{n,t}^{A})\big]^2\right) \\
&+ \sum_{i=1}^{N}\sum_{j=1}^{N} \left(\mathrm{var}(v_{i,t}^{A})\mathrm{var}(v_{j,t}^{B}) + 2\big[\mathrm{cov}(v_{i,t}^{A}, v_{j,t}^{B})\big]^2\right) \\
&= 4\mathrm{cov}(u_t^{A}, u_t^{B}) + 2\big[\mathrm{cov}(u_t^{A}, u_t^{B})\big]^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{N} \big[\mathrm{cov}(v_{i,t}^{A}, v_{j,t}^{B})\big]^2 \\
&+ 2\sum_{n=1}^{N} \left(\big[\mathrm{cov}(u_t^{A}, v_{n,t}^{B})\big]^2 + \big[\mathrm{cov}(u_t^{B}, v_{n,t}^{A})\big]^2\right),
\end{aligned}
$$

which completes the proof.

B Proof and discussion of Proposition 4

In this appendix, we provide a proof and interpretation for Proposition 4. Section B.1 gives the proof, Section B.2 discusses the relation between Proposition 4 and the GRS test of Gibbons et al. (1989), and Section B.3 provides interpretation for the net alpha introduced in Proposition 4.

B.1 Proof of Proposition 4

Let the vector $S_t = (F_t^\top, R_t^\top)^\top$ stack the returns of the factors and test assets. Thus, the average of S_t is $\mu_S = (\mu_F^\top, \mu_R^\top)^\top$ and its covariance matrix is

$$\Sigma_{S,S} = \begin{bmatrix} \Sigma_{F,F} & \Sigma_{F,R} \\ \Sigma_{R,F} & \Sigma_{R,R} \end{bmatrix}.$$

Similarly, the expected price-impact matrix for S_t is

$$\Lambda_{S,S} = \begin{bmatrix} \Lambda_{F,F} & \Lambda_{F,R} \\ \Lambda_{R,F} & \Lambda_{R,R} \end{bmatrix},$$

where $\Lambda_{F,F} = E[(\tilde{X}_t^F)^\top D_t \tilde{X}_t^F]$ is the expected price-impact matrix for the factors, $\Lambda_{R,R} = E[(\tilde{X}_t^R)^\top D_t \tilde{X}_t^R]$ is the expected price-impact matrix for the test assets, and $\Lambda_{R,F} = \Lambda_{F,R}^\top = E[(\tilde{X}_t^R)^\top D_t \tilde{X}_t^F]$ is the expected price-impact matrix for the test assets when the investor is also holding the factors.

Consider an investor with absolute risk aversion γ who faces the quadratic price-impact costs defined in (12). Then, Equation (12) implies that an investor holding a portfolio $\theta_S = [\theta_F, \theta_R]$ of the factors and test assets incurs the following expected price-impact cost:

$$f(\theta_S) = \frac{1}{2}\theta_F^\top \Lambda_{F,F}\theta_F + \frac{1}{2}\theta_R^\top \Lambda_{R,R}\theta_R + \theta_R^\top \Lambda_{R,F}\theta_F^\top. \tag{70}$$

The first term in the right-hand side of (70) is the price-impact cost associated with rebalancing the portfolio of the factors in isolation, θ_F, the second term is the price-impact cost associated with rebalancing the portfolio of the test assets in isolation, θ_R, and the third term is the price-impact cost associated with the interaction between the trades required to rebalance the portfolios of the test assets and the factors.

Equation (14) implies that the mean-variance utility net of price-impact costs of the investor when she has access to both the test assets and factors is

$$\mathrm{MVU}^{\gamma}\left([F,R]\right)=\frac{\mu_S^{\top}\left(\Sigma_{S,S}+\Lambda_{S,S}/\gamma\right)^{-1}\mu_S}{2\gamma}, \tag{71}$$

and that when she only has access to the factors is

$$\mathrm{MVU}^{\gamma}\left(F\right)=\frac{\mu_F^{\top}\left(\Sigma_{F,F}+\Lambda_{F,F}/\gamma\right)^{-1}\mu_F}{2\gamma}. \tag{72}$$

To prove the proposition, we first note that for an invertible matrix

$$U=\begin{bmatrix} A & B \\ B^{\top} & D \end{bmatrix},$$

where A is an invertible square matrix, we have

$$U^{-1}=\begin{bmatrix} A^{-1}+A^{-1}B\left(D-B^{\top}A^{-1}B\right)^{\top}B^{\top}A^{-1} & -A^{-1}B\left(D-B^{\top}A^{-1}B\right)^{-1} \\ -\left(D-B^{\top}A^{-1}B\right)^{-1}B^{\top}A^{-1} & \left(D-B^{\top}A^{-1}B\right)^{-1} \end{bmatrix}.$$

Let U be $\Sigma_{S,S}+\Lambda_{S,S}/\gamma$, and thus A, B, and D correspond to $\Sigma_{F,F}+\Lambda_{F,F}/\gamma$, $\Sigma_{F,R}+\Lambda_{F,R}/\gamma$, and $\Sigma_{R,R}+\Lambda_{R,R}/\gamma$, respectively. In this case, we have

$$\begin{aligned}\mu_S^{\top}U^{-1}\mu_S &= \mu_F^{\top}A^{-1}\mu_F+\mu_F^{\top}A^{-1}B\left(D-B^{\top}A^{-1}B\right)^{\top}B^{\top}A^{-1}\mu_F \\ &\quad -\mu_F^{\top}A^{-1}B\left(D-B^{\top}A^{-1}B\right)^{-1}\mu_R-\mu_R^{\top}\left(D-B^{\top}A^{-1}B\right)^{-1}B^{\top}A^{-1}\mu_F \\ &\quad +\mu_R^{\top}\left(D-B^{\top}A^{-1}B\right)^{-1}\mu_R \\ &= \left(\mu_R^{\top}-\mu_F^{\top}A^{-1}B\right)\left(D-B^{\top}A^{-1}B\right)^{-1}\left(\mu_R-B^{\top}A^{-1}\mu_F\right)+\mu_F^{\top}A^{-1}\mu_F.\end{aligned}$$

Thus,

$$\mu_S^{\top}U^{-1}\mu_S-\mu_F^{\top}A^{-1}\mu_F=\left(\mu_R^{\top}-\mu_F^{\top}A^{-1}B\right)\left(D-B^{\top}A^{-1}B\right)^{-1}\left(\mu_R-B^{\top}A^{-1}\mu_F\right). \tag{73}$$

Note that

$$B=\Sigma_{F,R}+\Lambda_{F,R}/\gamma=\left(\Sigma_{F,F}+\Lambda_{F,F}/\gamma\right)\beta^{\top}+\left(\Lambda_{F,R}/\gamma-\Lambda_{F,F}\beta^{\top}/\gamma\right),$$

where β is the slope obtained from an OLS regression of the test asset returns on the factor returns. Thus, we have

$$\begin{aligned}\mu_R-B^{\top}A^{-1}\mu_F &= \mu_R-\left[\beta\left(\Sigma_{F,F}+\Lambda_{F,F}/\gamma\right)+\left(\Lambda_{R,F}-\beta\Lambda_{F,F}/\gamma\right)\right]\left(\Sigma_{F,F}+\Lambda_{F,F}/\gamma\right)^{-1}\mu_F \\ &= \mu_R-\beta\mu_F-\left(\gamma\Lambda_{R,F}/\gamma-\gamma\beta\Lambda_{F,F}/\gamma\right)\frac{1}{\gamma}\left(\Sigma_{F,F}+\Lambda_{F,F}/\gamma\right)^{-1}\mu_F\end{aligned}$$

$$= \alpha - (\Lambda_{R,F} - \beta\Lambda_{F,F}/\gamma)\, \theta_F^* \equiv \alpha^{\text{net}}, \tag{74}$$

where α is the intercept obtained from regressing the test asset returns on the factor returns and the last equality follows from Equation (13). Thus, Equations (73) and (74) imply that

$$\text{MVU}^\gamma([F,R]) - \text{MVU}^\gamma(F) = \left(\alpha^{\text{net}}\right)^\top H_\gamma^{-1} \alpha^{\text{net}}, \tag{75}$$

where

$$H_\gamma = 2\gamma(\Sigma_{R,R} + \Lambda_{R,R}/\gamma) - 2\gamma(\Sigma_{R,F} + \Lambda_{R,F}/\gamma)(\Sigma_{F,F} + \Lambda_{F,F}/\gamma)^{-1}(\Sigma_{F,R} + \Lambda_{F,R}/\gamma), \tag{76}$$

which is positive definite because H_γ is the Schur complement of $2\gamma(\Sigma_{S,S} + \Lambda_{S,S}/\gamma)$, which is positive definite by assumption.

B.2 Relation to the GRS test

We now show that for the case without trading costs, Proposition 4 implies the result in equation (23) of Gibbons et al. (1989) that the increase in the squared Sharpe ratio of the investor when she has access to the test assets in addition to the factors in the model is a quadratic form of the gross alpha. To see this, note that for the case without trading costs ($\Lambda_{S,S} = 0$), we have that $\alpha^{\text{net}} = \alpha$, $\text{MVU}^\gamma([F,R]) = SR^2([F,R])/(2\gamma)$, $\text{MVU}^\gamma(F) = SR^2(F)/(2\gamma)$, and $H_\gamma = 2\gamma\Sigma_{R,R} - 2\gamma\Sigma_{R,F}\Sigma_{F,F}^{-1}\Sigma_{F,R}$. Thus, Equation (75) becomes

$$SR^2([F,R]) - SR^2(F) = \alpha^\top \left(\Sigma_{R,R} - \Sigma_{R,F}\Sigma_{F,F}^{-1}\Sigma_{F,R}\right)^{-1} \alpha. \tag{77}$$

B.3 Interpretation of the adjusted alpha

Consider an investor with absolute risk aversion γ. Then, the net alpha (α^{net}) defined in Equation (20) of Proposition 4 is the incremental return net of price-impact costs that the investor can achieve by making a marginal investment in the test assets when she is already holding the mean-variance portfolio of the factors in the model.

To see this, consider first the case without trading costs. Assume the investor holds the mean-variance portfolio of the factors in the model $\theta^* = \Sigma_{F,F}^{-1}\mu_F/\gamma$ and M dollars of the ith test asset with return $R_{i,t}$. Then, the average return of the investor's portfolio is $(\theta^*)\top\mu_F + M\mu_{R_i}$, where μ_{R_i} is the average return of the ith asset. Moreover, the beta of the investor's portfolio with respect to the factors in the model is $\theta^* + M\beta_i$, where β_i is the beta of the ith asset with respect to the factors. Thus, the average return of the investor's portfolio explained by the factors in the model is $(\theta^* + M\beta_i)^\top \mu_F$, and the average return of the investor's portfolio that is not explained by the factors in the model, per dollar invested in the ith test asset is:

$$\frac{1}{M}\left[(\theta^*)^\top \mu_F + M\mu_{R_i} - \left(\theta^* + M\beta_i\right)^\top \mu_F\right] = \alpha_i,$$

which is the alpha of the ith asset with respect to the factors in the model. Importantly, in the absence of trading costs the alpha of asset i does not depend on the mean-variance portfolio of the factors θ^*.

In the presence of price-impact costs, however, the *net* alpha of the ith test asset depends on the investor's mean-variance factor portfolio θ^*, and thus, on the investor's absolute risk aversion γ. To see this, note that the price-impact cost associated with holding the portfolio of the investor is

$$\frac{1}{2}\begin{bmatrix}(\theta^*)^\top & M\end{bmatrix}\Lambda_{S,S}\begin{bmatrix}\theta^* \\ M\end{bmatrix} = \frac{1}{2}(\theta^*)^\top \Lambda_{F,F}\theta^* + M(\theta^*)^\top \Lambda_{F,R_i} + \frac{M^2}{2}\Lambda_{R_i,R_i}.$$

Moreover, the beta of the investor's portfolio with respect to the factors is $\theta^* + M\beta_i$, and thus, the price-impact cost of the projection of the investor's portfolio on the factors is

$$\frac{1}{2}\left(\theta^* + M\beta_i\right)^\top \Lambda_{F,F}\left(\theta^* + M\beta_i\right) = \frac{1}{2}(\theta^*)^\top \Lambda_{F,F}\theta^* + M(\theta^*)^\top \Lambda_{F,F}\beta_i + \frac{M^2}{2}\beta_i^\top \Lambda_{F,F}\beta_i.$$

Then the average return net of price-impact costs of the investor's portfolio that is not explained by the factors in the model per dollar invested in the ith test asset is:

$$\frac{1}{M}\underbrace{\left((\theta^*)^\top \mu_F + M\mu_{R_i} - \left(\frac{1}{2}(\theta^*)^\top \Lambda_{F,F}\theta^* + M(\theta^*)^\top \Lambda_{F,R_i} + \frac{M^2}{2}\Lambda_{R_i,R_i}\right)\right)}_{\text{Average net return of the investor's portfolio}}$$

$$-\frac{1}{M}\underbrace{\left(\left(\theta^* + M\beta_i\right)^\top \mu_F - \left(\frac{1}{2}(\theta^*)^\top \Lambda_{F,F}\theta^* + M(\theta^*)^\top \Lambda_{F,F}\beta_i + \frac{M^2}{2}\beta_i^\top \Lambda_{F,F}\beta_i\right)\right)}_{\text{Average net return of the projection of the investor's portfolio on the factors}}$$

$$= \alpha_i - \left((\theta^*)^\top \Lambda_{F,R_i} - (\theta^*)^\top \Lambda_{F,F}\beta_i\right) - \frac{M}{2}\left(\Lambda_{R_i,R_i} - \beta_i^\top \Lambda_{F,F}\beta_i\right). \tag{78}$$

Furthermore, for the case where M is arbitrarily small we get

$$\lim_{M\to 0}\alpha_i - \left((\theta^*)^\top \Lambda_{F,R_i} - (\theta^*)^\top \Lambda_{F,F}\beta_i\right) - \frac{M}{2}\left(\Lambda_{R_i,R_i} - \beta_i^\top \Lambda_{F,F}\beta_i\right) = \alpha^{net}.$$

That is, α^{net} measures the incremental return net of price-impact costs that the investor can achieve by making a marginal investment in the test assets when she is already holding the mean-variance portfolio of the factors in the model.

Internet Appendix to

Comparing Factor Models with Price-Impact Costs

This Internet Appendix contains several robustness checks and additional information. Section IA.1 compares the p-values from our Proposition 6 with those from the GRS test for comparing nested models in the absence of trading costs. Section IA.2 discusses how we correct the upward bias in sample mean-variance utilities. Section IA.3 uses bootstrap to check the finite-sample accuracy of our proposed asymptotic distributions. Section IA.4 gives the results for the out-of-sample bootstrap tests for different values of absolute risk aversion. Section IA.5 gives summary statistics for the optimal portfolio weights of the different factor models. Finally, Section IA.6 shows that the relative performance of the six factor models we consider is robust to considering factors constructed using the *banding* transaction-cost mitigation strategy used in section 5 of Detzel et al. (2023).

IA.1 Comparing Proposition 6 and the GRS test

Although our Proposition 6 is designed to compare factor models in the presence of price-impact costs, one can also use it to compare factor models in the absence of trading costs by setting $\Lambda_t = \Lambda = 0$. As a robustness check, we now compare the p-values for model comparisons in the absence of trading costs obtained using Proposition 6 and the GRS test, which is the test recommended by Barillas et al. (2020) to compare nested models in the absence of costs. Specifically, suppose model A nests model B. We first use Proposition 6 to compare the two models in terms of the maximum mean-variance utility and obtain the p-value of this test. We then let the extra factors of model A be the left-hand side test assets and let the factors of model B be the right-hand side factors, and run a time-series regression of the test assets on the factors. Then, we calculate the GRS test statistic based on the time-series alpha and obtain the p-value of this test.

Table IA.1 reports the p-values from the seven tests for the seven sets of nested models in our dataset. The first column lists the acronym of the nested model comparison. The second column reports the p-value of the test based on Proposition 6. The third column reports the p-value of the finite-sample GRS test in which the test statistic has an F distribution, and the fourth column reports the p-value of the asymptotic GRS test in which the test statistic has a χ^2 distribution. From this table, we find that the p-value of the GRS test (both the finite-sample version and the asymptotic version) is very similar to that of the test based on Proposition 6, although when comparing FF5 and FF6, the p-value of the test based on Proposition 6 is slightly larger, and thus, less significant than its GRS counterpart. Therefore, we conclude that the test based on Proposition 6 is very similar to the GRS test in the absence of trading costs, and it can be viewed as a generalization of the GRS test because it is also applicable to compare factor models in the presence of price-impact costs.

IA.2 Correcting the upward bias in sample utilities

As mentioned in Footnote 12 of the main body of the manuscript, the sample mean-variance utility net of price-impact costs of a factor model suffers from a small-sample upward bias as documented by Jobson and Korkie (1980) and Barillas et al. (2020). In this appendix, we show how we correct this upward bias.

We use bootstrap to estimate the upward bias of the sample mean-variance utility net of price-impact costs of each model. First, we bootstrap with replacement a sample with T^b months, and read the factor returns and factor rebalancing trades of the bootstrapped months.[18] Second, we calculate the mean-variance utility net of price-impact costs of each factor model on the bootstrap sample. We then repeat the two steps for 100,000 times. For each model, we calculate its average mean-variance utility net of price-impact costs on the 100,000 bootstrap samples. The difference between the average mean-variance utility net of price-impact costs on the bootstrap samples and the utility in the original sample is our bootstrap estimator of the upward bias of each model, and we denote it as Δ_Λ. The bias-corrected mean-variance utility net of price-impact costs of a model is obtained by subtracting Δ_Λ from its mean-variance utility net of price-impact costs estimated using the original sample. In the main body of the manuscript, all reported mean-variance utility net of price-impact costs are bias corrected, and we implement bias correction when comparing factor models using Proposition 5.

IA.3 Finite-sample accuracy of asymptotic distributions

Propositions 5 and 6 provide two asymptotic distributions for the difference between the sample mean-variance utilities net of price-impact costs of two models. In this appendix, we use bootstrap simulations to check how accurately these asymptotic distributions fit their

[18] In Sections 4.4 and 4.5, T^b is chosen to be 491, which equals to the size of our original sample.

finite-sample counterparts. We set the absolute risk-aversion parameter $\gamma = 10^{-9}$ as in our base case. To simplify notation, we drop the superscript γ from the mean-variance utility net of price-impact costs MVU^{γ}.

IA.3.1 Asymptotic distribution from Proposition 5

In this section, we use bootstrap simulations to check how accurately the asymptotic distribution in Proposition 5 fits its finite-sample counterpart.

We assume that the true data generating process (DGP) is characterized by the sample estimators $\hat{\mu}, \hat{\Sigma}$, and $\hat{\Lambda}$. We use the superscript g to denote the true DGP, and use the superscript b to denote values obtained from bootstrap samples. We bootstrap with replacement 100,000 samples of T^b observations from our original sample. In other words, each bootstrap sample is generated from the true DGP. On each bootstrap sample, we estimate the mean-variance utility net of price-impact costs for every factor model, and adjust its finite-sample bias following Barillas et al. (2020) using the procedures in Appendix IA.2 to obtain U^b.[19] We then compute the following quantity on each bootstrap sample and for each pair of models A and B:

$$\sqrt{T^b}\big([\text{MVU}_A^b - \text{MVU}_B^b] - [\text{MVU}_A^g - \text{MVU}_B^g]\big), \tag{IA1}$$

where MVU_A^g and MVU_B^g denote the mean-variance utilities net of price-impact costs of models A and B under the true DGP g, which are known by construction. To simplify the notation, in this section we drop the superscript γ from the mean-variance utility net of price-impact costs. The 100,000 values of (IA1) characterize the finite-sample distribution of (IA1), and Proposition 5 characterizes the asymptotic distribution of (IA1) when $T^b \to \infty$.

Figure IA.1 compares the finite-sample distribution when $T^b = 491$ (blue histogram) and the asymptotic distribution based on Proposition 5 (orange curve) of (IA1) for all pairs of factor models that are non-nested. We observe that for most pairwise model comparisons the finite-sample distribution is very close to the asymptotic distribution. In some cases, such as the comparison of FF5 and DMNU20, the asymptotic distribution does not fit the finite-sample distribution closely. The reason of this is that the number of observations

[19] In particular, for each model we subtract the quantity Δ_Λ defined in Appendix IA.2 from the sample mean-variance utility net of price-impact costs on each bootstrap sample, and thus the average bias-adjusted mean-variance utility net of price-impact costs over the 100,000 bootstrap samples equals to U^g.

in each bootstrap sample, $T^b = 491$, is not large enough to guarantee the convergence of the finite-sample distribution to the asymptotic distribution. To validate this argument, Figure IA.2 depicts the finite-sample distribution when $T^b = 2{,}000$ and the asymptotic distribution of (IA1). This figure shows that the asymptotic distribution provides a good fit when the number of observations in each bootstrap sample is large enough.

IA.3.2 Asymptotic distribution from Proposition 6

In this section, we use bootstrap simulations to check how accurately the asymptotic distribution in Proposition 6 fits its finite-sample counterpart.

One difficulty in this experiment is that the asymptotic distribution in Proposition 6 holds only under the null hypothesis that $\text{MVU}_A = \text{MVU}_B$, but this null hypothesis does not hold in our sample for any pair of nested models. To simplify notation, we drop the γ superscript from the mean-variance utility net of price-impact costs in this section. To address this, we assume that the true DGP is characterized by our original sample estimators $\hat{\Sigma}$ and $\hat{\Lambda}$, and we adjust $\hat{\mu}$ to make the null hypothesis hold under the true DGP.

We now describe how to adjust $\hat{\mu}$ using the notation in the proof of Proposition 6. Let the mean-variance portfolio estimated on the original sample of the larger model A be

$$\hat{\theta}_A^* = \frac{1}{\gamma}\left(\hat{\Sigma}_A + \hat{\Lambda}_A/\gamma\right)^{-1}\begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_2\end{bmatrix} = \begin{bmatrix}\hat{\theta}_1^*\\ \hat{\theta}_2^*\end{bmatrix},$$

where $\hat{\mu}_1$ and $\hat{\mu}_2$ are the sample average returns of the factors F_1 and F_2, respectively, and $\hat{\theta}_1^*$ and $\hat{\theta}_2^*$ are the sample estimates of the mean-variance portfolio weights of model A on factors F_1 and F_2, respectively. We find a vector $c \in \mathbb{R}^{K_2}$, such that

$$\hat{\theta}_A^{*\prime} = \frac{1}{\gamma}\left(\hat{\Sigma}_A + \hat{\Lambda}_A/\gamma\right)^{-1}\begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_2 - c\end{bmatrix} = \begin{bmatrix}\hat{\theta}_1^{*\prime}\\ 0\end{bmatrix}.$$

In other words, we adjust the mean returns of the extra factors F_2 by c so that the mean-variance portfolio of model A assigns zero weight to F_2. The vector c must satisfy

$$\frac{1}{\gamma}\left[\left(\hat{\Sigma}_A + \hat{\Lambda}_A/\gamma\right)^{-1}\right]_{22} c = \hat{\theta}_2^*,$$

and thus it is uniquely identified because matrix $\left[(\hat{\Sigma}_A + \hat{\Lambda}_A/\gamma)^{-1}\right]_{22}$ is invertible. We assume that the true DGP has the adjusted mean return vector $[\hat{\mu}_1, \hat{\mu}_2 - c]$. Note that with the adjusted mean return, the mean-variance portfolio of model A assigns zero weight to F_2. Therefore, under the true DGP, models A and B have the same mean-variance utility net of price-impact costs.

We use the superscript g to denote the true DGP, and use the superscript b to denote values obtained from the bootstrap samples. To make our original sample follow the true DGP, we adjust the sample returns of F_2 by setting $R'_{2,t} = R_{2,t} - c$ for all t. We then bootstrap with replacement from this adjusted sample to generate 100,000 bootstrap samples with T^b observations. Thus, each bootstrap sample comes from the true DGP, which satisfies the null hypothesis of Proposition 6. On each bootstrap sample, we calculate the mean-variance utility net of price-impact costs U^b for every factor model. We do not adjust the finite-sample bias of U^b for the reasons discussed in Footnote 12. We then compute the following quantity on each bootstrap sample for every pair of nested models A and B:

$$T^b(\mathrm{MVU}_A^b - \mathrm{MVU}_B^b), \tag{IA2}$$

The 100,000 values of (IA2) characterize the finite-sample distribution of (IA2), and Proposition 6 characterizes the asymptotic distribution of (IA2) when $T^b \to \infty$.

Figure IA.3 compares the finite-sample distribution when $T^b = 491$ (blue histogram) and the asymptotic distribution based on Proposition 6 (orange curve) of (IA2) for all pairs of nested models. The figure shows that the asymptotic distribution fits its finite-sample counterpart very accurately. Moreover, Figure IA.3 is based on sample mean-variance utilities net of price-impact costs that are *not* adjusted for finite-sample bias, and thus the figure also demonstrates that the asymptotic distribution in Proposition 6 adequately captures the finite-sample bias of the mean-variance utility.

IA.4 OOS bootstrap tests for different risk aversion

In the main body of the manuscript, we discuss the results of the out-of-sample bootstrap test for the base case with absolute risk aversion $\gamma = 10^{-9}$. In this section, we report the results for the cases with a lower and a higher absolute risk-aversion parameters. The exact procedure of the bootstrap test is the same as that in Section 4.5.

Tables IA.2 and IA.3 report the out-of-sample bootstrap results for the cases with absolute risk-aversion parameters $\gamma = 10^{-10}$ and $\gamma = 10^{-8}$, respectively. In each table, Panel A reports the average mean-variance utility net of price-impact costs of each model, and Panel B reports the frequency with which the row model outperforms the column model across the bootstrap samples. In both cases, we find that the average out-of-sample mean-variance utility net of price-impact costs of each model is lower than its in-sample counterpart because of estimation error. However, the out-of-sample relative performance of the models is generally consistent with that in sample.

Specifically, for the case with absolute risk aversion $\gamma = 10^{-10}$, DMNU20 outperforms all other models in over 70% of the bootstrap samples, which is consistent with the in-sample results that DMNU20 significantly outperforms all other models. Although HXZ4 has a higher average out-of-sample utility than FFC4, FFC4 still outperforms HXZ4 in 37.5% of the bootstrap samples, which is consistent with the in-sample results that the two models are statistically indistinguishable. For the case with absolute risk aversion $\gamma = 10^{-8}$, consistent with the in-sample results, HXZ4 has the highest average out-of-sample mean-variance utility net of price-impact costs, and it outperforms FFC4 in 92.2% of the bootstrap samples. Furthermore, it outperforms the CAPM, FF5, FF6, and DMNU20 models in 91.3, 70.2, 58.3, and 79.0% of the bootstrap samples, respectively. This confirms the in-sample results that when $\gamma = 10^{-8}$, HXZ4 is the best-performing model.[20] In summary, the results of the out-of-sample bootstrap tests for the two cases with lower and higher absolute risk-aversion parameters than the base case confirm the main finding based on our statistical test.

IA.5 Summary statistics of optimal portfolios

Table IA.4 reports the summary statistics of the optimal portfolios for the base-case absolute risk aversion parameter $\gamma = 10^{-9}$. Panel A reports summary statistics of the portfolios computed ignoring price-impact costs and Panel B reports the performance of portfolios computed accounting for price-impact costs and trading diversification. For each panel, the second to fourth columns report the mean, *t*-statistic, and standard deviation of the

[20] The result of the comparison between FFC4 and DMNU20 in this table is different from its in-sample counterpart. In Table 7, DMNU20 significantly outperforms FFC4, while FFC4 has higher out-of-sample mean-variance utility net of price-impact costs in 53.0% of the bootstrap samples. This is not surprising because the higher absolute risk-aversion parameter makes the price-impact costs matrix Λ have lower impact on the optimal portfolio, and the covariance matrix of the returns Σ has relatively higher impact on the optimal portfolio. Accurately estimating the covariance matrix of a twenty-factor model is hard and thus the performance of DMNU20 is likely to more impacted by estimation error.

monthly gross returns. The fifth and six columns report the monthly average price-impact cost ignoring trading diversification (No TD) and considering trading diversification (With TD). The last two columns report the monthly turnover ignoring and considering trading diversification. Panel C reports the weight on each factor (in billion dollars) of the optimal portfolios that ignore price impact-costs (No cost) and that account for price-impact costs (PIC).

A few comments are in order. First, portfolios that ignore price-impact costs have higher turnovers and substantially higher price-impact costs than those that account for price-impact costs. For example, by accounting for price-impact costs in constructing the optimal portfolio, HXZ4's monthly turnover decreases from 17.62% to 12.46%, and its monthly price-impact costs decreases from 0.827% to 0.100%. Second, trading diversification substantially decreases the turnover and price-impact costs of the models, and the effect is more substantial for models with more factors, such as FF6 and DMNU20. For example, for portfolios that account for price-impact costs, the turnover of DMNU20 decreases by 55% from 15.56% to 7.05%, and its monthly price-impact costs decreases by 75% from 0.096% to 0.024%; while the decrease of turnover and price-impact costs for a low-dimensional FFC4 in this case are only 12.8% and 20.3%, respectively. This is consistent with DeMiguel et al. (2020), who find that in the presence of trading costs, high-dimensional models are likely to perform well because of the benefits of trading diversification across factors.

Panel C of Table IA.4 reports the weight of the optimal portfolio (in dollar) assigned to each factor. We find that price-impact costs restrain the participation of an investor in the financial market because the total capital invested in all factors becomes lower when the investor considers price-impact costs. More importantly, we find that the weights on high-cost factors such as ROE become substantially lower, and the relative weight on the market factor becomes substantially higher when considering price-impact costs because the cost of the market factor is negligible. This implies that all candidate factor models become similar to CAPM when pricing the returns of extremely large investments. Moreover, this also motivates using high-dimensional models to price the returns of extremely large investments because high-dimensional models better exploit trading diversification, and thus, have a higher chance to improve the investment opportunity set of CAPM than low-dimensional models in the presence of large price-impact costs.

IA.6 Transaction-cost mitigated factors

Consistent with the findings of Detzel et al. (2023), Table 3 in the main body of the manuscript shows that among the factors constructed from double and triple sorts, factors that are rebalanced monthly (UMD, ROE, IA, ME) have much higher turnover and price-impact cost than factors that are rebalanced annually (SMB, HML, RMW, CMA). In the main body of the manuscript, we consider UMD, ROE, IA, and ME factors constructed as in the original papers that proposed the factors. In this section, we show that the relative performance of the six factor models we consider is robust to considering UMD, ROE, IA, and ME factors constructed using the *banding* transaction-cost mitigation strategy used in section 5 of Detzel et al. (2023).[21]

Table IA.5 compares the summary statistics for the UMD, ROE, IA, and ME factors constructed as in the original papers to those of the factors constructed using the banding transaction-cost mitigated strategy. The table shows that, consistent with the findings of Detzel et al. (2023), the banding strategy substantially decreases the turnover of these four factors. For instance, the monthly turnover of the momentum (UMD) factor decreases from 51.93% to 29.87% when we use banding to construct the portfolio. Moreover, the table also shows that the banding strategy is effective at decreasing the price-impact cost incurred to hold these factors. For instance, the monthly price-impact cost associated with rebalancing the ROE factor decreases from 10.1 to 8.5 basis points when we use banding to construct the portfolio.

Nonetheless, our main finding is that the relative performance of the six factor models is robust to considering UMD, ROE, IA, and ME factors constructed using the banding transaction-cost mitigation strategy. In particular, Tables IA.6–IA.10 replicate Tables 4–8 in the main body of the manuscript for the case where the UMD, ROE, IA, and ME factors are constructed using the banding transaction-cost mitigation strategy. Although the tables show that the banding transaction-cost mitigation strategy helps to increase the mean-variance utility net of price-impact costs of the models that use monthly-rebalanced factors (HXZ4, FFC4, FF6), the relative performance of the various models does not change: the HXZ4, FF6, and DMNU20 models continue to be the best at spanning the investment opportunities of investors with high, medium, and low absolute risk aversion, respectively.

[21] We thank Detzel et al. (2023) for making the replication code for their paper publicly available.

For instance, Table IA.7 shows that for the base-case with absolute risk aversion $\gamma = 10^{-9}$, the FF6 factor model continues to dominate because it significantly outperforms all other low-dimensional models and although it has a smaller mean-variance utility net of price-impact costs than the high-dimensional DMNU20 model, the difference between their utilities is not statistically significant, a result that parallels that obtained from Table 5 in the main body of the manuscript. Table IA.8 shows that for the case with $\gamma = 10^{-10}$, in which price-impact costs are more important, the high-dimensional model DMNU20 continues to significantly outperform every other model for the case where the monthly-rebalanced factors are constructed using the banding strategy, a result that parallels that obtained from Table 6 in the main body of the manuscript. Table IA.9 shows for the case with $\gamma = 10^{-8}$, in which price-impact costs are less important, the HXZ4 model is preferred because it significantly outperforms CAPM and FF4, it ouperforms FF5, and although it is outperformed by FF6 and DMNU20, the difference in mean-variance utility net of price-impact costs with these models is not statistically significant, a result that parallels that from Table 7 in the main body of the manuscript. Finally, Table IA.10 reports the out-of-sample boostrap test for the models based on the monthly rebalanced factors that are constructed using the banding approach, and the relative performance of the six factor models is very similar to that in Table 8 in the main body of the manuscript.

Figure IA.1: Distribution of the difference in mean-variance utilities: finite-sample distribution when $T^b = 491$ and asymptotic distribution based on Proposition 5

This figure compares the finite-sample distribution (blue histogram) and the asymptotic distribution (orange curve) of the difference in mean-variance utilities net of price-impact costs in (IA1) for all pairs of models that are non-nested. The title of each sub-figure illustrates the two models for comparison. The finite-sample distribution is obtained by evaluating (IA1) on 100,000 bootstrap samples with $T^b = 491$ observations, and the asymptotic distribution is based on Proposition 5.

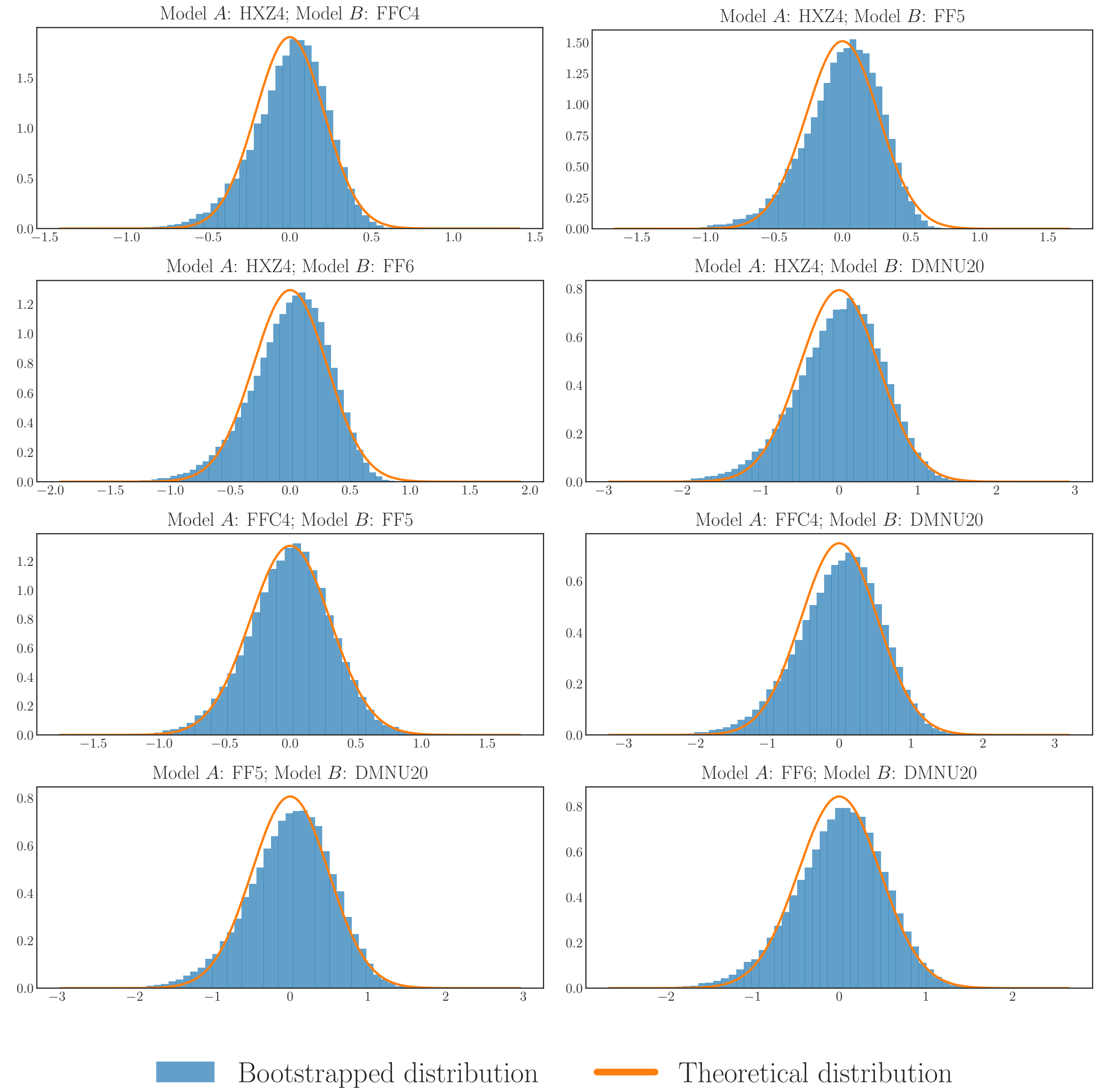


Figure IA.2: Distribution of the difference in mean-variance utilities: finite-sample distribution when $T^b = 2000$ and asymptotic distribution based on Proposition 5

This figure compares the finite-sample distribution (blue histogram) and the asymptotic distribution (orange curve) of the difference in mean-variance utilities net of price-impact costs in (IA1) for all pairs of models that are non-nested. The title of each sub-figure illustrates the two models for comparison. The finite-sample distribution is obtained by evaluating (IA1) on 100,000 bootstrap samples with $T^b = 2000$ observations, and the asymptotic distribution is based on Proposition 5.

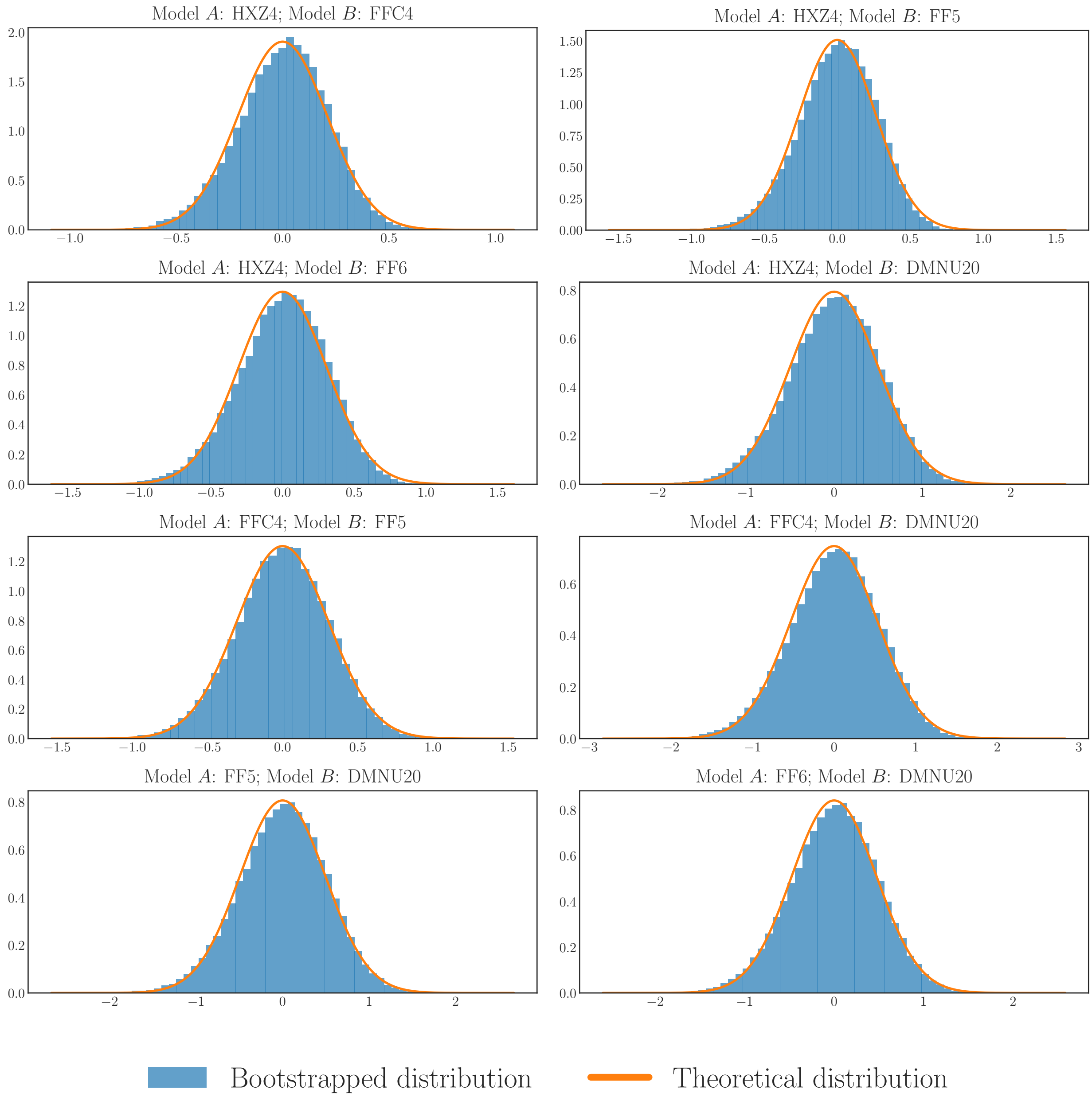


Figure IA.3: Distribution of the difference in mean-variance utilities: finite-sample distribution when $T^b = 491$ and asymptotic distributions based on Proposition 6

This figure compares the finite-sample distribution (blue histogram) and the asymptotic distribution (orange curve) of the difference in mean-variance utilities net of price-impact costs in (IA2) for all pairs of models that are nested. The title of each sub-figure illustrates the two models for comparison. The finite-sample distribution is obtained by evaluating (IA2) on 100,000 bootstrap samples with $T^b = 491$ observations, and the asymptotic distribution is based on Proposition 6

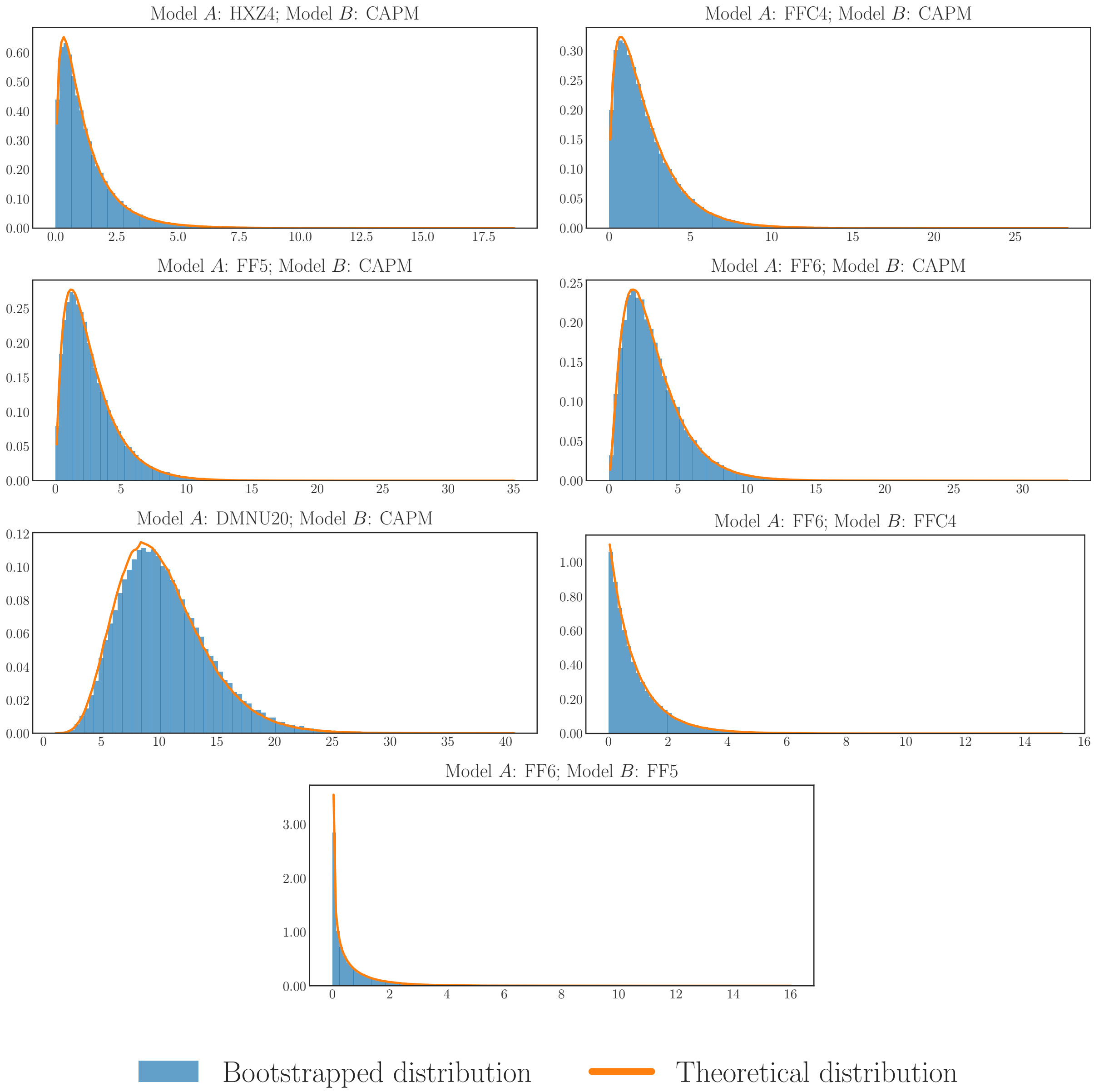


Table IA.1: Comparing p-values using Proposition 6 and the GRS test

This table reports the p-values of the test using Proposition 6 and of the GRS test to compare nested models in the absence of trading costs. The first column lists the acronyms of the nested models. The second column reports the p-value of the test based on Proposition 6. The third and the fourth columns report the p-value of the finite-sample GRS test and of the asymptotic GRS test, respectively.

	p-values		
	Proposition 6	Finite-sample GRS	Asymptotic GRS
CAPM v. HXZ4	0.000	0.000	0.000
CAPM v. FFC4	0.003	0.000	0.000
CAPM v. FF5	0.000	0.000	0.000
CAPM v. FF6	0.000	0.000	0.000
CAPM v. DMNU20	0.000	0.000	0.000
FFC4 v. FF6	0.000	0.000	0.000
FF5 v. FF6	0.036	0.007	0.007

Table IA.2: Bootstrap out-of-sample utility net of price-impact costs ($\gamma = 10^{-10}$)

Panel A reports the average out-of-sample (OOS) scaled mean-variance utility net of price-impact costs across 100,000 bootstrap samples of each factor model under the case with absolute risk-aversion parameter $\gamma = 10^{-10}$. Panel B reports the frequency with which the row model outperforms the column model out-of-sample across the 100,000 bootstrap samples.

Panel A: Average mean-variance utilities net of trading costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma\text{MVU}_{\Lambda}^{\gamma}$	0.0121	0.0124	0.0101	0.0142	0.0158	0.0264

Panel B: Frequency row model outperforms column model

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.356	0.526	0.327	0.287	0.245
HXZ4			0.625	0.324	0.269	0.244
FFC4				0.200	0.120	0.215
FF5					0.205	0.254
FF6						0.272

Table IA.3: Bootstrap out-of-sample utility net of price-impact costs ($\gamma = 10^{-8}$)

Panel A reports the average out-of-sample (OOS) scaled mean-variance utility net of price-impact costs across 100,000 bootstrap samples of each factor model under the case with absolute risk-aversion parameter $\gamma = 10^{-8}$. Panel B reports the frequency with which the row model outperforms the column model out-of-sample across the 100,000 bootstrap samples.

Panel A: Average mean-variance utilities net of trading costs

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
$2\gamma\text{MVU}_{\Lambda}^{\gamma}$	0.0122	0.0727	0.0208	0.0555	0.0606	0.0062

Panel B: Frequency row model outperforms column model

	CAPM	HXZ4	FFC4	FF5	FF6	DMNU20
CAPM		0.087	0.317	0.151	0.173	0.476
HXZ4			0.922	0.702	0.583	0.790
FFC4				0.198	0.099	0.530
FF5					0.305	0.711
FF6						0.745

Table IA.4: Summary statistics for optimal portfolios ($\gamma = 10^{-9}$)

This table reports the summary statistics of the optimal portfolios for the base-case absolute risk aversion parameter $\gamma = 10^{-9}$. Panel A reports summary statistics of the portfolios computed ignoring price-impact costs and Panel B reports the performance of portfolios computed accounting for price-impact costs and trading diversification. For each panel, the second to fourth columns report the mean, t-statistic, and standard deviation of the monthly gross returns. The fifth and six columns report the monthly average price-impact cost ignoring trading diversification (No TD) and considering trading diversification (With TD). The last two columns report the monthly turnover ignoring and considering trading diversification. Panel C reports the weight on each factor (in billion dollars) of the optimal portfolios that ignore price impact-costs (No cost) and that account for price-impact costs (PIC).

Panel A: Statistics of portfolios that ignore trading costs

Model	Gross return			PIC (%)		TO (%)	
	Mean (%)	t-stat	Std. (%)	No TD	With TD	No TD	With TD
HXZ4	0.433	8.46	1.134	1.068	0.827	23.22	17.62
FFC4	0.463	5.71	1.798	0.163	0.133	20.13	17.77
FF5	0.296	7.52	0.871	0.346	0.191	10.64	7.24
FF6	0.348	8.05	0.959	0.333	0.178	13.77	9.70
DMNU20	0.192	10.29	0.414	0.953	0.357	35.73	20.09

Panel B: Statistics of portfolios that account for price-impact costs

Model	Gross return			PIC (%)		TO (%)	
	Mean (%)	t-stat	Std. (%)	No TD	With TD	No TD	With TD
HXZ4	0.527	6.09	1.918	0.133	0.100	16.42	12.46
FFC4	0.500	5.43	2.042	0.069	0.055	16.36	14.26
FF5	0.387	6.75	1.268	0.104	0.054	8.99	5.75
FF6	0.406	7.49	1.202	0.130	0.061	13.11	8.92
DMNU20	0.266	8.65	0.681	0.096	0.024	15.56	7.05

Panel C: Factor weights ($B)

Model	Cost	MKT	SMB	HML	RMW	CMA	UMD	ME	IA	ROE
HXZ4	No cost	6.65						3.24	12.67	11.10
	PIC	4.19						-0.23	2.39	2.08
FFC4	No cost	4.85	0.24	4.94			4.31			
	PIC	4.24	-0.02	2.91			2.16			
FF5	No cost	6.33	3.56	-3.65	11.32	14.09				
	PIC	4.79	1.83	0.89	5.98	3.79				
FF6	No cost	6.80	3.28	-1.68	10.37	12.67	3.11			
	PIC	5.17	1.67	1.53	5.72	3.66	1.95			

Table IA.5: Factor summary statistics: Cost-mitigated factors

This table compares the summary statistics of the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors constructed following the procedure in the papers that originally proposed the factors to those of the factors constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023). The first column gives the acronym of the factor. The second and third columns give the average monthly *gross* return of the factor and its t-statistic. The fourth and fifth columns give the average monthly *net-of-price-impact-costs* return of the factor and its t-statistic, when one invests one billion dollars on each leg of the factor. The sixth column gives the factor's monthly price-impact cost (PIC), the seventh column the factor's monthly turnover (TO), and the eighth column the factor's capacity. The ninth column reports the average of the monthly trade-weighted market capitalization, and the last column reports the average of the trade-weighted market capitalization at the end of June. Average returns, turnovers, and price-impact costs are reported in percentage. Investment positions, capacity, and trade-weighted market capitalization are reported in terms of market capitalization at the end of our sample, which spans the period from January 1980 to December 2020.

Factor	Gross returns (%)		Net returns (%)		Costs (%), turnover (%), and capacity ($B)			Trade-weighted market cap ($B)	
	Average	t-statistic	Average	t-statistic	PIC	TO	Capacity	Monthly	June
Panel A: Standard factors									
UMD	0.557	2.744	0.476	2.343	0.081	51.93	6.86	90.52	73.47
ROE	0.521	4.394	0.420	3.536	0.101	35.42	5.16	63.41	55.13
IA	0.286	3.309	0.235	2.703	0.051	24.60	5.64	68.15	67.19
ME	0.147	1.108	0.129	0.974	0.018	19.19	8.12	70.11	58.26
Panel B: Transaction-cost mitigated factors									
UMD^{CM}	0.672	3.272	0.624	3.035	0.048	29.87	13.93	77.52	68.91
ROE^{CM}	0.556	4.395	0.470	3.712	0.085	27.15	6.50	49.85	50.34
IA^{CM}	0.305	3.313	0.253	2.729	0.052	20.65	5.88	56.29	57.76
ME^{CM}	0.125	0.971	0.107	0.837	0.017	16.21	7.18	59.13	55.30

Table IA.6: Significance of difference in mean-variance utility without price-impact costs: Cost-mitigated factors

This table reports the significance of the difference between the mean-variance utilities of the row and column models in the absence of trading costs for the case where the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors are constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023) and the rest of the factors are constructed following the procedure in the papers that originally proposed the factors. Panel A reports the scaled sample mean-variance utility of each of the six factor models in the absence of trading costs. Panel B reports the p-value for the difference in mean-variance utility for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities without trading costs

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
$2\gamma MVU^{\gamma}$	0.0216	0.1343	0.0654	0.1012	0.1206	0.1569

Panel B: p-values

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
CAPM		0.000	0.001	0.000	0.000	0.000
$HXZ4^{CM}$			0.005	0.108	0.285	0.296
$FFC4^{CM}$				0.112	0.000	0.011
FF5					0.015	0.077
$FF6^{CM}$						0.183

Table IA.7: Significance of difference in mean-variance utility with price-impact costs: Cost-mitigated factors

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the baseline case with absolute risk-aversion parameter $\gamma = 10^{-9}$ and where the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors are constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023) and the rest of the factors are constructed following the procedure in the papers that originally proposed the factors. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	HXZ4^{CM}	FFC4^{CM}	FF5	FF6^{CM}	DMNU20
$2\gamma\text{MVU}^{\gamma}_{\Lambda}$	0.0216	0.0436	0.0488	0.0577	0.0780	0.0982

Panel B: p-values

	CAPM	HXZ4^{CM}	FFC4^{CM}	FF5	FF6^{CM}	DMNU20
CAPM		0.000	0.000	0.000	0.000	0.000
HXZ4^{CM}			0.337	0.115	0.016	0.007
FFC4^{CM}				0.284	0.000	0.019
FF5					0.001	0.033
FF6^{CM}						0.175

Table IA.8: Significance of difference in mean-variance utility with costs for $\gamma = 10^{-10}$: Cost-mitigated factors

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the case with absolute risk-aversion parameter $\gamma = 10^{-10}$ and where the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors are constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023) and the rest of the factors are constructed following the procedure in the papers that originally proposed the factors. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
$2\gamma MVU_{\Lambda}^{\gamma}$	0.0215	0.0243	0.0272	0.0293	0.0346	0.0560

Panel B: p-values

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
CAPM		0.010	0.009	0.004	0.000	0.000
$HXZ4^{CM}$			0.159	0.096	0.011	0.003
$FFC4^{CM}$				0.299	0.000	0.007
FF5					0.000	0.007
$FF6^{CM}$						0.018

Table IA.9: Significance of difference in mean-variance utility with costs for $\gamma = 10^{-8}$: Cost-mitigated factors

This table reports the significance of the difference between the mean-variance utilities net of price-impact costs of the row and column models for the case with absolute risk-aversion parameter $\gamma = 10^{-8}$ and where the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors are constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023) and the rest of the factors are constructed following the procedure in the papers that originally proposed the factors. Panel A reports the scaled sample mean-variance utility net of price-impact costs of each of the six factor models. Panel B reports the p-value for the difference in mean-variance utility net of price impact costs for every pairwise model comparison. The p-value is computed using Proposition 5 when the row and column models overlap and Proposition 6 when the row model is nested in the column model.

Panel A: Mean-variance utilities net of price-impact costs

	CAPM	HXZ4^{CM}	FFC4^{CM}	FF5	FF6^{CM}	DMNU20
$2\gamma\text{MVU}_{\Lambda}^{\gamma}$	0.0216	0.1011	0.0630	0.0904	0.1118	0.1297

Panel B: p-values

	CAPM	HXZ4^{CM}	FFC4^{CM}	FF5	FF6^{CM}	DMNU20
CAPM		0.000	0.001	0.000	0.000	0.000
HXZ4^{CM}			0.032	0.292	0.299	0.171
FFC4^{CM}				0.147	0.000	0.019
FF5					0.009	0.093
FF6^{CM}						0.281

Table IA.10: Bootstrap out-of-sample utility with price-impact costs: Cost-mitigated factors

Panel A reports the average out-of-sample (OOS) scaled mean-variance utility net of price-impact costs across 100,000 bootstrap samples of each factor model under the case with absolute risk-aversion parameter $\gamma = 10^{-9}$ and where the momentum (UMD), profitability (ROE), investment (IA), and size (ME) factors are constructed using the banding transaction-cost mitigation strategy considered in Detzel et al. (2023) and the rest of the factors are constructed following the procedure in the papers that originally proposed the factors. Panel B reports the frequency with which the row model outperforms the column model out-of-sample across the 100,000 bootstrap samples.

Panel A: Average mean-variance utilities net of price-impact costs

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
$2\gamma MVU_{\Lambda}^{\gamma}$	0.0122	0.0264	0.0239	0.0318	0.0462	0.0265

Panel B: Frequency row model outperforms column model

	CAPM	$HXZ4^{CM}$	$FFC4^{CM}$	FF5	$FF6^{CM}$	DMNU20
CAPM		0.165	0.276	0.213	0.175	0.355
$HXZ4^{CM}$			0.482	0.345	0.218	0.453
$FFC4^{CM}$				0.392	0.111	0.441
FF5					0.205	0.504
$FF6^{CM}$						0.638