Factor Investing and Factor-Neutral Investing

Rong Wang *

November 10, 2025

Latest draft available here.

Abstract

Factor investing based on machine-learning (ML) models promises high out-of-sample Sharpe ratios but fails to deliver in practice. I show that the out-of-sample performance of factor mean–variance portfolios constructed from leading ML models is economically unviable, turning strongly negative once realistic trading costs are accounted for. In contrast, factor-neutral investing—which builds optimal zero-beta portfolios to exploit model mispricing rather than harvest factor premia—proves far more robust and implementable. The zero-beta constraints act as a natural shrinkage device, mitigating turnover and improving resilience to trading frictions. In a comprehensive out-of-sample backtest spanning 1990–2024, factor-neutral strategies achieve net-of-cost annualized Sharpe ratios between 0.6 and 1.2 and monthly alphas of 0.4%–1.5%, unexplained by known factors. These results imply that, in practice, the most valuable signal in ML-based factor models is not their estimated factors, but their errors.

1. Introduction

The systematic pursuit of cross-sectional return predictability has made factor investing a cornerstone of modern empirical asset pricing. Theoretical and empirical studies have provided strong justifications for this approach, arguing that systematic exposure to well-defined factors can explain and predict variation in expected returns across assets. In practice, factor investing seeks to construct portfolios that tilt toward characteristics associated with higher expected returns—such as value, size, momentum, or profitability—and thereby harvest the corresponding factor risk premia. This framework has shaped both academic research and institutional portfolio management, offering a disciplined and transparent means of linking asset pricing theory with implementable investment strategies.

Recent advances in machine learning (ML) have introduced new optimism to this endeavor. By leveraging high-dimensional data and flexible functional forms, ML-based factor

*Fuqua School of Business, Duke University; rong.wang1996@duke.edu.

models promise to uncover complex nonlinearities and interaction effects that traditional linear specifications cannot capture (e.g., Kelly, Pruitt, and Su, 2019; Gu, Kelly, and Xiu, 2021). These models have demonstrated impressive in-sample and even out-of-sample explanatory power, often reporting remarkably high Sharpe ratios for their associated factor portfolios. Consequently, ML methods are increasingly viewed as the frontier of quantitative factor discovery and portfolio construction.

However, the practical performance of ML-based factor investing remains disappointing once realistic implementation costs are considered. Despite producing statistically strong signals, the corresponding factor investing portfolios—particularly those derived from sophisticated models such as Instrumented Principal Component Analysis (IPCA) and Autoencoder (AE) frameworks—exhibit extreme turnover, leverage, and trading intensity. The resulting frictions erode profitability to the point of economic infeasibility, reversing much of the apparent performance advantage observed in gross returns. This discrepancy between theoretical promise and empirical realizability highlights a growing disconnect between the goals of model-based factor identification and the constraints of implementable investment strategies.

This paper addresses this disconnect by developing and empirically validating a factor-neutral investing framework, which reinterprets the output of ML-based factor models from an investment perspective. Rather than harvesting risk premia associated with estimated factors, the proposed approach constructs optimal factor-neutral (zero-beta) portfolios designed to exploit model mispricing—that is, systematic deviations between predicted and realized returns. The factor-neutral constraint serves as a natural form of statistical shrinkage, regularizing portfolio weights, mitigating turnover, and enhancing robustness to estimation errors and trading frictions. Conceptually, factor-neutral investing transforms factor models from tools of risk pricing into instruments of "arbitrage": the most valuable information lies not in the estimated factors themselves, but in the structure of their residual errors.

Using a comprehensive out-of-sample backtest spanning 1990–2024, I show that this reorientation has substantial economic implications. Traditional ML-based factor portfolios that deliver gross Sharpe ratios exceeding 3.0 collapse to strongly negative net performance once transaction costs are incorporated. In contrast, factor-neutral portfolios achieve robust and persistent profitability, with net-of-cost annualized Sharpe ratios between 0.6 and 1.2 and monthly alphas ranging from 0.4% to 1.5%, unexplained by established risk factors. These results demonstrate that factor-neutrality constraints not only improve implementability but also isolate economically meaningful pricing errors across models.

Additional analyses underscore the distinctive properties of the two schools of investment. ML–based factor investing strategies exhibit extreme portfolio concentrations and leverage at the individual-stock level, reflecting their sensitivity to high-frequency changes in estimated factor loadings. In contrast, factor-neutral investing produces far more modest (short) positions and leverage across a wide range of factor models, consistent with the weights shrinkage effect of the factor-neutrality constraint. Moreover, while both approaches display decreasing returns to scale, the diseconomies of scale are substantially less pronounced for factor-neutral investing: as investment size increases, the marginal deterioration in performance is markedly smaller than in FI, suggesting greater scalability and capacity. Examining the persistence of predictive signals further highlights this contrast—alpha signals derived from model mispricing exhibit longer stability, whereas factor signals decay rapidly, rendering them difficult to monetize after accounting for trading frictions. Finally, the long-term dynamics of factor-neutral investing profitability reveal a close association with macroeconomic conditions, particularly the stock–bond covariance, which governs the pricing of systematic risk and hedging demand. As the stock–bond covariance has shifted toward positive territory in the post-COVID period, the relative appeal of factor-neutral strategies has increased, making them a particularly timely and resilient approach in the current macro-financial environment.

The investigation of factor investing versus factor-neutral investing has crucial implications on the methodology of factor model comparison. The widely cited framework of Barillas and Shanken (2017); Barillas et al. (2020) emphasizes the Sharpe ratios of factor portfolios. This paper shows that comparing the maximum attainable Sharpe ratio of factor-neutral portfolios is fundamentally equivalent to comparing the Sharpe ratios of factor portfolios. A lower maximum Sharpe ratio of factor-neutral portfolios implies a smaller distance between the factor portfolio and the efficiency frontier, which in turn indicates a higher Sharpe ratio of the factors. While preserving this economic intuition, factor-neutral portfolios are more feasible due to the modest short positions and leverage. And the use of them is broadly applicable: when the risk-free rate is unobservable, computing Sharpe ratios for factors becomes ambiguous, whereas the Sharpe ratios of zero-investment portfolios do not rely on the risk-free rate.

This paper contributes to three strands of literature. First, it systematically compares two categories of investment strategies: factor investing versus factor-neutral investing. I show that the latter extracts alpha signals from complex factor structures, offering robust and feasible investment opportunities. Second, it connects advances in ML–based asset pricing with the practical realities of trading and execution costs. Third, it extends the recent literature on model comparison by establishing factor-neutral portfolios as a feasible benchmark.

The recommended factor-neutral investment strategy is closely related to market-neutral and statistical arbitrage approaches. These strategies typically follow a two-step process: first, an "alpha generation" step identifies potentially mispriced assets by modeling their returns, either through purely statistical methods or factor models; second, a "hedging" step constructs a portfolio by going long on underpriced assets and short on overpriced ones while neutralizing exposure to systematic risk factors. Equity market-neutral strategies are widely employed by hedge funds (Fung and Hsieh, 2001), and other common statistical arbitrage approaches include pairs trading (Gatev et al., 2006; Avellaneda and Lee, 2010) and the betting-against-beta strategy (Frazzini and Pedersen, 2014). Kelly et al. (2019) and Kim et al. (2021) construct "arbitrage portfolios" to exploit the mispricing from factor models. More recently, Guijarro-Ordonez et al. (2021) apply deep learning (a convolutional transformer) to extract complex time-series trading signals from the residuals of a multi-factor model, which are then mapped to an optimal trading policy using a flexible neural network designed to maximize the Sharpe ratio. My approach aligns with this principle by using the factor model's pricing errors as the alpha signal and mean-variance optimization to construct a factor-neutral portfolio. Rather than introducing new techniques for alpha generation or hedging, the key contribution of this paper is to rigorously apply this established framework to compare the out-of-sample investment performance of different underlying factor models, with particular attention to the role of transaction costs.

This paper contributes to the literature on comparing asset pricing models. Critically, as pointed out in Fama and French (2015), maximum Sharpe ratio factor portfolios used in both Gibbons et al. (1989) (GRS) testing and Barillas and Shanken (2017) (BS) comparison often require massive leverage and short positions, undermining the validity of these approaches in practical applications. In contrast, the shrinkage property of factor-neutrality constraints largely increase the robustness of the new framework to practical concerns. Related refinements in model testing comparison consider out-of-sample evaluation (Fama and French, 2018; Kan et al., 2024) and transaction costs (Detzel et al., 2023); these considerations are naturally incorporated in the factor-neutral portfolio framework. In summary, this generality makes factor-neutral portfolios a powerful and practical common ground (as advocated by Karolyi and Van Nieuwerburgh, 2020) for evaluating the full spectrum of modern asset pricing models, including the increasingly complex models that dominate contemporary research. In summary, the primary strength of the factor-neutral portfolio approach lies in its broad applicability and robustness: it accommodates a much broader range of model settings and it is more robust to real-world implementation costs.

The remainder of the paper is organized as follows. Section 2 describes the investment methodology, data, candidate factor models, and transaction cost models. Section 3 presents the investment performance and distinctive properties of factor investing versus factor-neutral investing. Section 4 explores the implications of factor-neutral portfolios in factor model comparison, and Section 5 concludes.

2. Methodology and Data

2.1. Factor Investing

Factor investing refers to a systematic investment approach that seeks to capture persistent sources of risk premia that explain differences in expected returns across securities. Rather than relying on discretionary stock selection, factor investors allocate to portfolios that are intentionally tilted toward specific characteristics—known as factors—such as value, size, momentum, profitability, or quality. These factors are motivated by asset-pricing theory, which links higher expected returns to compensation for bearing systematic risks, or by behavioral explanations that attribute premia to persistent investor biases and market inefficiencies. A notable subset of factor investing is smart beta investing, which applies factor principles within a passive, rules-based framework. Smart beta strategies deviate from traditional market-capitalization weighting by reweighting index constituents according to fundamental, statistical, or risk-based metrics (for example, book-to-market ratios, volatility, or earnings quality). The goal is to enhance risk-adjusted performance while maintaining transparency and cost efficiency. In essence, smart beta represents the index-based implementation of factor investing—systematically targeting factor exposures within a rules-driven structure that bridges the gap between purely passive indexing and active management.

Within a given factor model, this paper considers four complementary approaches to constructing dynamic factor-investing strategies that differ in how they exploit conditioning information and risk management principles. First, the factor mean-variance (FMV) portfolio represents the traditional mean-variance optimal combination of factors, obtained by estimating the conditional means and covariances of factor returns and solving for the portfolio that maximizes expected Sharpe ratio. It serves as the benchmark static factor-investing strategy. Second, following Haddad et al. (2020), the factor-timing portfolio (HKS) incorporates time-variation in expected factor premia by conditioning on predictive state variables—specifically valuation spreads such as the book-to-market ratio—to forecast factor returns. In their framework, factor timing corresponds to constructing the optimal stochastic discount factor when factor risk prices vary over time, yielding substantial improvements in Sharpe ratios relative to static factor exposure. Third, the volatility-timing approach of Moreira and Muir (2017) scales aggregate factor exposure inversely with recent realized variance, effectively managing portfolio volatility rather than expected returns. Their volatility-managed portfolios (MM) take less risk when volatility is high and more when volatility is low, thereby increasing risk-adjusted performance across multiple equity factors. Fourth, DeMiguel et al. (2024) extend this idea in a multifactor context, proposing a conditional mean-variance multifactor portfolio (DMU) that manages each factor individually according to its own volatility and then optimally combines the managed factors. Collectively, these four strategies provides a comprehensive empirical laboratory for studying the investment performance of multifactor investing.

Let $\mathbf{F}_t = (f_{1t}, f_{2t}, \ldots, f_{Kt})'$ denote the K-dimensional vector of factor excess returns at time t, with conditional mean $\boldsymbol{\mu}_{F,t} = \mathbb{E}_t[\mathbf{F}_{t+1}]$ and conditional covariance matrix $\boldsymbol{\Sigma}_{F,t} = \text{Var}_t(f_{t+1})$. Weights are estimated in the expanding windows and applied in the out-of-sample periods to construct portfolios.

2.1.1. Factor Mean-Variance Portfolio (FMV)

The factor mean-variance (FMV) portfolio represents the unconditional efficient allocation across factors under mean-variance utility. The optimal factor tangency portfolio weight is $\boldsymbol{\omega}_t^{\text{FMV}} = \frac{1}{\gamma}\boldsymbol{\Sigma}_{F,t}^{-1}\hat{\boldsymbol{\mu}}_{F,t}$ where γ is the coefficient of relative risk aversion and $\hat{\boldsymbol{\mu}}_{F,t}$ is estimated via historical mean in the expanding estimation window.

2.1.2. Factor Timing Portfolio (HKS)

Haddad et al. (2020) introduce factor timing by allowing expected factor premia to vary with conditioning information z_t, such as valuation ratios (e.g., book-to-market spreads). The conditional factor model is

$$\mathbb{E}_t[\mathbf{F}_{t+1}] = \boldsymbol{B}' z_t, \tag{1}$$

where z_t contains predictive state variables. The conditional mean-variance optimal portfolio is therefore

$$\boldsymbol{\omega}_t^{\text{HKS}} = \frac{1}{\gamma}\boldsymbol{\Sigma}_{F,t}^{-1}\mathbb{E}_t[\mathbf{F}_{t+1}] = \frac{1}{\gamma}\boldsymbol{\Sigma}_{F,t}^{-1}\boldsymbol{B}' z_t. \tag{2}$$

The HKS factor-timing strategy dynamically adjusts exposures as a function of forecasted factor premia. Empirically, I follow Haddad et al. (2020) and use the book-to-market ratio of the principal components of factor returns as z_t.

2.1.3. Volatility Timing Portfolio (MM)

Moreira and Muir (2017) propose volatility-managed portfolios, which scale exposure inversely with realized variance rather than expected returns. Their insight is that volatility is highly persistent, but expected returns are not, so Sharpe ratios can be improved by reducing exposure when volatility is high.

The MM portfolio scales the entire FMV portfolio:

$$r_{p,t+1}^{\mathrm{MM}} = c\,\frac{1}{\hat{\sigma}_{p,t}^2}(\boldsymbol{\omega}^{\mathrm{FMV}})'\mathbf{F}_{t+1}, \tag{3}$$

where $\hat{\sigma}_{p,t}^2$ is the realized variance of the FMV portfolio. This approach keeps relative factor weights fixed but varies total exposure to smooth portfolio risk.

2.1.4. *Multifactor Volatility Timing Portfolio (DMU)*

DeMiguel et al. (2024) generalize volatility management to a conditional multifactor setting where each factor's exposure is adjusted individually according to its own volatility and then optimally recombined using mean–variance optimization.

For each factor k:

$$r_{k,t+1}^{\sigma} = c_k\,\frac{1}{\hat{\sigma}_{k,t}^2} f_{k,t+1}, \tag{4}$$

and the vector of original and managed factors is $\mathbf{F}_{t+1}^{\sigma} = (f_{1t}, f_{2t}, \ldots, f_{Kt}, r_{1,t+1}^{\sigma}, \ldots, r_{K,t+1}^{\sigma})'$. The conditional mean-variance optimal weights are

$$\boldsymbol{\omega}_t^{\mathrm{DMU}} = \frac{1}{\gamma}\boldsymbol{\Sigma}_{\sigma,t}^{-1}\boldsymbol{\mu}_{\sigma,t}, \tag{5}$$

where $\boldsymbol{\mu}_{\sigma,t} = \mathbb{E}_t[\mathbf{F}_{t+1}^{\sigma}]$ and $\boldsymbol{\Sigma}_{\sigma,t} = \mathrm{Var}_t(\mathbf{F}_{t+1}^{\sigma})$ reflect the combined factors set.

2.2. *Factor-Neutral Investing*

Factor-neutral investing represents an investment framework that deliberately eliminates exposure to systematic risk factors—such as market beta, value, size, or momentum—in order to isolate idiosyncratic or alpha-driven sources of return. Unlike factor investing, which seeks to harvest risk premia by tilting toward rewarded characteristics, factor-neutral strategies are constructed so that portfolio performance is independent of these common drivers. In practice, factor neutrality is achieved through regression-based residualization or explicit hedging: expected returns or signals are orthogonalized with respect to known factors, and any remaining predictive component is attributed to firm-specific inefficiencies or temporary mispricings. The resulting portfolios are typically long–short and beta-neutral, ensuring that gains or losses cannot be explained by broad market movements. This framework aligns closely with the goal of extracting pure alpha and serves as an empirical test of market efficiency—if residual returns persist after removing all systematic exposures, markets are not fully efficient with respect to those factors. Moreover, factor-neutral investing plays a stabilizing role in portfolio construction by reducing drawdowns during periods of market stress and by providing diversification benefits when traditional factor exposures become crowded or highly correlated.

Let $\mathbf{R}_t$ denote the N-dimensional vector of asset returns at time t, with conditional mean $\boldsymbol{\mu}_t = \mathbb{E}_t[\mathbf{R}_{t+1}]$ and conditional covariance matrix $\boldsymbol{\Sigma}_t = \text{Var}_t(\mathbf{R}_{t+1})$. Weights are estimated in the expanding windows and applied in the out-of-sample periods to construct portfolios.

I construct the optimal zero-investment, zero-beta portfolio with no exposure to any systematic risk factors of the model and evaluate its investment performance. This is achieved by solving the following constrained mean–variance optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{\omega}} \boldsymbol{\omega}'\boldsymbol{\mu} - \frac{\gamma}{2}\boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega} \\ &s.t. \quad \boldsymbol{\omega}'\boldsymbol{\iota} = 0, \quad \boldsymbol{\omega}'\boldsymbol{\beta} = \mathbf{0}_K \end{aligned} \tag{6}$$

where γ denotes the risk aversion coefficient, $\boldsymbol{\beta}$ the estimated betas, and $\boldsymbol{\iota}$ a vector of ones. The analytical solution for the optimal zero-investment, zero-beta portfolio weights is:

$$\boldsymbol{\omega}_z^* = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\left[\mathbf{I} - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}'\boldsymbol{\Sigma}^{-1}\right]\boldsymbol{\mu} \tag{7}$$

where $\boldsymbol{\Pi} = [\boldsymbol{\iota},\ \boldsymbol{\beta}]$. Recall that the unconstrained optimal portfolio weights are given by $\boldsymbol{\omega}^* = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Comparing the two, equation (7) can be interpreted as the optimal portfolio based on the projected mean return $\mathbf{P}_\Pi\boldsymbol{\mu}$, where $\mathbf{P}_\Pi \equiv \left[\mathbf{I} - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}'\boldsymbol{\Sigma}^{-1}\right]$ is a generalized projection matrix with weights $\boldsymbol{\Sigma}^{-1}$. This matrix projects the mean return vector onto the subspace orthogonal to $\boldsymbol{\Pi}$, so that $\tilde{\boldsymbol{\alpha}} \equiv \mathbf{P}_\Pi\boldsymbol{\mu}$ can be interpreted as pricing errors (alphas) and $\boldsymbol{\omega}_z^*$ represents the optimal portfolio of alphas that is beta neutral.

Factor neutrality is not a new concept. Maintaining low or minimal correlation between hedge fund returns and market returns has long been considered a desirable feature of hedge funds (Fung and Hsieh, 2001). In particular, the equity market-neutral strategy (also known as the beta-neutral strategy) is popular among hedge funds and mutual funds. These strategies are designed to eliminate systematic market risk and deliver returns that are independent of broad market movements.[1] According to BarclayHedge, the total as-

[1] Quote from the prospectus of AQR's Equity Market Neutral Fund: "The Fund seeks to provide investors with returns from the potential gains from its long and short equity positions. The Fund is designed to be market- or beta-neutral, which means that the Fund seeks to achieve returns that are not closely correlated with the returns of the equity markets in which the Fund invests. Accordingly, the Adviser, on average, intends to target a portfolio beta of zero to equity markets in which the Fund invests over a normal business cycle. Achieving zero portfolio beta would result in returns with no correlation to the returns of equity markets in which the Fund invests over a normal business cycle."

sets under management for Equity Market-Neutral hedge funds was \$76.6 billion as of the first quarter of 2025 (Figure B.1). In practice, portfolio construction is considerably more complex. Typically, an Equity Market-Neutral fund (EMN) uses proprietary quantitative models to go long undervalued stocks and short overvalued stocks (alpha generation) and then manages the aggregate portfolio to maintain a market beta of approximately zero (the zero-beta constraint). Frazzini and Pedersen (2014)'s betting-against-beta (BAB) strategy follows the same logic. According to the prospectus of AQR's Equity Market Neutral Fund, "The Adviser employs a model which aggregates many measures, or signals, that are used to determine a stock's relative attractiveness, utilizing a wide variety of traditional and non-traditional, public and proprietary data sources...Applying these signal categories, the Adviser takes long or short positions in sectors, industries and companies that it believes are attractive or unattractive (to achieve market neutrality)."

The notion of market neutrality has been extended to multi-factor neutrality through statistical arbitrage. Building on the idea of pairs trading, Avellaneda and Lee (2010) propose a factor model-based statistical arbitrage strategy. They decompose a stock's return into a "systematic" component (driven by multiple risk factors) and an "idiosyncratic" component (the residual). The residual is modeled as an Ornstein–Uhlenbeck process, reflecting the key assumption that it is mean-reverting. A trading signal is then triggered by the deviation of the current residual from its estimated long-term mean (analogous to an alpha). After this alpha-generation step, they go long the underpriced stock and simultaneously short a basket of risk factors in proportion to the stock's estimated betas. By construction, the combined position has net-zero exposure to the systematic risk factors. More recently, Guijarro-Ordonez et al. (2021) employ deep learning (a convolutional transformer) to extract complex time-series trading signals from the residual portfolios of a multi-factor model. These signals are then mapped to an optimal trading policy using another flexible neural network designed to explicitly maximize the Sharpe ratio.

All beta-neutral strategies, including the zero-investment zero-beta strategies discussed in this paper, follow the same general principle of portfolio construction. The process begins with alpha extraction—identifying underpriced and overpriced stocks—followed by establishing long and short positions to enforce the zero-beta constraint. Different approaches vary in both the method of alpha generation and the design of the long/short positions. Specifically, the zero-beta strategy in this paper relies on factor models for alpha extraction and mean–variance optimization for portfolio construction.

2.3. Modeling Transaction Costs

Investment strategies are evaluated accounting for transaction costs. Suppose $\boldsymbol{\pi}_t$ denotes an $N \times 1$ vector of portfolio allocation (dollar amounts) across individual stocks. The $N \times 1$ turnover vector of individual stocks required to rebalance the investment portfolio is:

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\pi}_{t+1} - \boldsymbol{\pi}_t \circ (\boldsymbol{\iota} + \mathbf{r}_t) \tag{8}$$

where $\boldsymbol{\iota}$ an $N \times 1$ vector of ones, and $\mathbf{r}_t$ the $N \times 1$ vector of individual returns. $\circ$ is the component-wise product. $\boldsymbol{\pi}_t \circ (\boldsymbol{\iota} + \mathbf{r}_t)$ represents the effective holdings prior to rebalancing.

An important insight from DeMiguel et al. (2024) is that netting trades across multiple portfolios—a form of trading diversification—can yield substantial transaction-cost savings. Following this idea, I first net the rebalancing trades across the 273 characteristic-sorted portfolios before applying transaction costs at the individual-stock level. This procedure captures the cost reduction from offsetting trades among portfolios while accurately accounting for the actual costs incurred when adjusting positions in the underlying stocks.

I consider two types of transaction costs. First, proportional trading costs increase proportionally to turnover trades:

$$f(\boldsymbol{\tau}_t) = \left\|\boldsymbol{\Phi}_t \circ \boldsymbol{\tau}_t\right\|_1 \tag{9}$$

where $\|\cdot\|_1 = \sum_{i=1}^{N} |\cdot|$ denotes the 1-norm, and $\boldsymbol{\Phi}_t$ is a $N \times 1$ vector of individual stock-level transaction-cost parameters, measured by the average low-frequency effective bid–ask spreads (Chen and Velikov, 2023). The individual transaction-cost parameter, $\boldsymbol{\Phi}_t$, is measured using the average low-frequency (LF) effective bid–ask spreads described in Chen and Velikov (2023). They provide both high-frequency (HF) measures, derived from intraday trade and quote data, and low-frequency (LF) measures, based only on daily price and volume data. Since HF measures are available only from 1983 onward, I use the average of four LF measures (Hasbrouck, 2009; Corwin and Schultz, 2012; Kyle and Obizhaeva, 2016; and Abdi and Ranaldo, 2017), which are available across my full sample. Chen and Velikov (2023) finds that LF measures tend to be biased upward compared to HF measures in the modern era of electronic trading (post-2005). Moreover, Frazzini et al. (2018) argues that actual transaction costs may be substantially lower than suggested by previous studies. Consequently, the transaction costs in this analysis may be overestimated, implying that the investment performance reported in Section 3 could be understated. Figure B.6 shows the time variation of the mean, median, 5th percentile, and 95th percentile of individual transaction costs from January 1960 to December 2024.

Because the proportional cost function is non-linear due to the absolute value operator, I apply transaction costs after constructing the optimal portfolio weights from the standard mean–variance optimization problem. This approach is conservative, as the resulting investment performance serves as a lower bound for the true performance that would obtain if transaction costs were incorporated directly into the portfolio optimization stage.

Second, I consider price impact costs that are quadratic functions of turnover trades:

$$f(\boldsymbol{\tau}_t) = \frac{1}{2}\boldsymbol{\tau}_t'\boldsymbol{\Lambda_t}\boldsymbol{\tau}_t \tag{10}$$

where $\frac{1}{2}\boldsymbol{\Lambda_t}\boldsymbol{\tau}_t$ represents the price impact, and $\boldsymbol{\Lambda}_t$ is a $N \times 1$ vector of individual stock-level Kyle's lambda, calibrated such that the market impact, $\frac{1}{2}\boldsymbol{\Lambda_t}\boldsymbol{\tau}_t$, is 0.1% when trading 1% of the daily dollar volume of a stock (Jensen et al., 2024).[2] The expected daily volume is defined as the average daily dollar volume over the preceding six months.

Because the price impact cost function is quadratic in portfolio allocations, I incorporate these costs in the portfolio optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{\pi}} \boldsymbol{\pi}'\boldsymbol{\mu} - \frac{\gamma}{2}\boldsymbol{\pi}'\boldsymbol{\Sigma}\boldsymbol{\pi} - \frac{W}{2}\boldsymbol{\pi}'\boldsymbol{\Lambda}\boldsymbol{\pi} \\ &s.t. \quad \boldsymbol{\omega}'\boldsymbol{\iota} = 0, \quad \boldsymbol{\omega}'\boldsymbol{\beta} = \mathbf{0}_K \end{aligned} \tag{11}$$

where γ is the risk aversion coefficient. W denotes investor wealth, which directly enters into the optimization problem because of the quadratic form of trading costs. The analytical solution for the optimal zero-investment, zero-beta portfolio weights is:

$$\boldsymbol{\omega}_z^* = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\left[\mathbf{I} - \boldsymbol{\Omega}\left(\boldsymbol{\Omega}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Omega}\right)^{-1}\boldsymbol{\Omega}'\boldsymbol{\Sigma}^{-1}\right]\boldsymbol{\mu} \tag{12}$$

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma} + W\boldsymbol{\Lambda}/\gamma \tag{13}$$

To evaluate investment performance with price impact costs, I consider three investors who have \$5, \$50, and \$100 billion dollars at the end of 2024. I assume that investors' wealth grows at the same rate as the market, i.e. $W_t = W_{t-1}(1 + R_{m,t})$, where $R_{m,t}$ denotes the realized market return.

[2] I am using the same example as in Jensen et al. (2024): trading \$5 million over a day in a stock with a daily volume of \$500 million moves the price by $\frac{1}{2}\frac{0.2}{\$500m}$, leading to a transaction cost of 1 2 $\frac{1}{2}\frac{0.2}{\$500m} \times (\$5m)^2 =$ \$5000.

2.4. Data

I obtain monthly individual stock returns and characteristics from the Global Factor Data website organized by Jensen, Kelly, and Pedersen (2023) (JKP).[3] My sample spans January 1960 to December 2024, covering 780 months (65 years). In total, there are 3,658,843 stock-month observations for 28,828 unique stocks, averaging 4,691 stocks per month. For each characteristic, I fill missing values with the cross-sectional median by 2-digit SIC industry each month. After this step, I retain 136 characteristics with complete coverage across the full sample (see Appendix A.1 for the full list). All characteristics are lagged one month. JKP update characteristics using the most recent accounting data four months after the fiscal period ends, ensuring that lagged characteristics are in the public information set and avoiding look-ahead bias.

I use characteristic-sorted portfolios in model evaluation and zero-beta rate estimation. Following Jensen et al. (2023), I construct portfolios and factors for each characteristic and retain the two corner portfolios (top and bottom terciles), since much of the relevant information resides in the extremes (Lettau and Pelger, 2020). I also include the middle tercile portfolio sorted by size so that the market return is spanned by the testing portfolios. This yields a total of $136 \times 2 + 1 = 273$ univariate-sorted portfolios. Each factor is constructed as the return spread between portfolios in the top and bottom terciles of a given characteristic. The factor's sign is adjusted, if necessary, to ensure that its average return over the sample period is positive. The timing of my portfolio formation differs a bit from standard practice: while Fama–French form portfolios annually in June and JKP form them monthly, I construct portfolios each December, aligned with the rolling out-of-sample periods in my following analysis. During this procedure, I store portfolio weights on individual stocks. My empirical results are robust to the portfolio formation method.

In conditional factor models (described in Section **??**), lagged characteristics also serve as determinants of model parameters. Following Gu et al. (2020), Gu et al. (2021), and others, I cross-sectionally rank-normalize all characteristics into the $(-1, 1)$ interval each month.[4]

Although factor-neutral strategies are constructed without a risk-free rate, factor-investing typically require a funding cost of capital to calculate Sharpe ratios. Rather than the Treasury bill rate, which is distorted by convenience yields, I use short-term unsecured borrowing rates, which more closely reflect actual institutional funding costs. Specifically, I use: (i) the 30-Day AA Financial Commercial Paper rate (FRED: CPF1M) from January 1997 to

[3] I thank the authors for making the data easily accessible (a WRDS account with access to CRSP and Compustat is required).

[4] Stock characteristics often display high skewness and kurtosis. This rank transformation reduces sensitivity to outliers.

December 2024, (ii) the 1-Month Commercial Paper Rate (FRED: CP1M) from April 1971 to December 1996, and (iii) the 1-Month Finance Paper Placed Directly Average Offering Rate (FRED: H0RIFSPPFM01NM) from January 1960 to March 1971.

2.5. Candidate Factor Models

I study a wide range of factor models in this paper. For each model, I consider two asset universes: individual stocks and 272 characteristic-sorted portfolios (see Section 2.4). I evaluate specifications with 1, 3, 5, and 6 factors. A general representation of a factor model is given by

$$r_{i,t+1} = \alpha_{i,t}(z_{i,t}) + \boldsymbol{\beta}(z_{i,t})'\mathbf{f}_{t+1} + \varepsilon_{i,t+1} \tag{14}$$

where $\mathbf{f}_{t+1}$ is a K-dimensional vector of factors, $\alpha_{i,t}(z_{i,t})$ and $\boldsymbol{\beta}(z_{i,t})$ denote the intercept and risk loadings, potentially functions of the 136 stock characteristics.

First, I consider *unconditional linear models with pre-specified factors*, the most widely used class of models. These models assume that a small set of observable, economically motivated factors explain stock returns, with constant intercepts and loadings: $\alpha_{i,t}(z_{i,t}) = \alpha_i$ and $\boldsymbol{\beta}(z_{i,t}) = \boldsymbol{\beta}$. Prominent examples include Fama and French (1993), Carhart (1997), Hou et al. (2015), Stambaugh and Yuan (2017), and Fama and French (2018). In the 1-factor case, I include only the market factor. In the 3-factor case, I include market, size, and value factors. In the 5-factor case, I include market, size, value, profitability, and investment factors. In the 6-factor case, I add the momentum factor to the 5-factor specification.[5] For individual stocks, I require at least 60 months of observations to include them in the time-series regressions used for estimating betas. I refer to these models collectively as “FF”.

Second, I turn to *unconditional linear models with PCA factors*, rooted in the Arbitrage Pricing Theory (APT) (Ross, 1976; Huberman, 1982; Chamberlain and Rothschild, 1983; Ingersoll Jr, 1984; Connor and Korajczyk, 1986, among others). Unlike the CAPM or the Intertemporal CAPM (ICAPM), which derive from equilibrium models with explicit preferences and market assumptions, APT is a reduced-form framework.[6] It assumes a factor structure in which returns decompose into systematic and idiosyncratic components. With sufficiently many assets, idiosyncratic risk diversifies away, and the absence of arbitrage opportunities yields an approximate linear beta-pricing relation. As in the “FF” case, loadings

[5] The market factor is the weighted average of all stocks. The size, value, profitability, investment, and momentum factor are constructed using the characteristics “market_equity”, “be_me”, “ope_me”, “at_gr1”, and “ret_12_1”.

[6] Extensions that embed APT in an equilibrium setting include Connor (1984) and Connor and Korajczyk (1988).

are static: $\alpha_{i,t}(z_{i,t}) = \alpha_i$ and $\boldsymbol{\beta}(z_{i,t}) = \boldsymbol{\beta}$. The APT naturally motivates the use of principal component analysis (PCA) to extract statistical factors. Cooper et al. (2021) demonstrate that such statistically constructed factors outperform most of the traditional "FF"-style multi-factor models, in both economic and statistical terms. Following this insight, I extract 1, 3, 5, and 6 factors using the standard PCA. For individual stocks, I require at least 60 months of observations to include them in the PCA used for factor extraction and beta estimation.[7] I refer to these specifications collectively as "PCA".

Third, I examine *conditional linear models with latent factors*, which allow risk loadings to vary with stock characteristics. Specifically, I implement the instrumented principal component analysis (IPCA) models of Kelly et al. (2019) with 1, 3, 5, and 6 factors. In this framework, the intercepts and loadings are modeled as linear functions of observable characteristics: $\alpha_{i,t}(z_{i,t}) = \Gamma'_{\alpha} z_{i,t}$ and $\boldsymbol{\beta}(z_{i,t}) = \Gamma'_{\beta} z_{i,t}$. Unlike FF models, which rely on pre-specified factors, or PCA models, which assume static loadings, IPCA jointly estimates latent factors and their time-varying exposures using an alternating least squares (ALS) algorithm. I collectively refer to these models as "IPCA".[8]

Finally, I consider *conditional non-linear models with latent factors*, which leverage machine learning methods to capture richer relationships between characteristics and risk exposures. While IPCA models imposes linearity, neural networks can approximate complex nonlinear mappings. I use the conditional autoencoder model of Gu et al. (2021). Autoencoders are neural networks designed for unsupervised dimension reduction, which can be viewed as nonlinear analogues of PCA. They aim to learn a compressed, low-dimensional representation of input data by training the network to reconstruct their own inputs as accurately as possible. A standard latent factor model can be interpreted as a simple autoencoder, while conditional autoencoders extend this by incorporating observable characteristics. The architecture consists of two networks: a multi-layer beta network capturing non-linear mappings from characteristics to loadings, and a single-layer factor network generating latent factors as linear combinations of portfolios. The two are then combined as in equation (14). My implementation follows Gu et al. (2021) but adds an intercept term in the beta network, allowing $\alpha_{i,t}$ to vary flexibly with characteristics, and uses the 272 characteristic-sorted portfolios as

[7] Unlike IPCA and AE, which do not impose filters on individual stocks, both FF and PCA require a minimum number of observations for regressions and for PCA. As a result, Tables B.2 and B.1 do not compare models rigorously on exactly the same universe of individual stocks. This limitation does not affect the new model comparison method, since the comparison is based on zero-beta portfolios regardless of their composition.

[8] Another strand of conditional linear factor models emphasizes time-varying risk premia in addition to time-varying loadings, pioneered by Ferson and Harvey (1991), who attribute much of cross-sectional return predictability to variations in risk premia than by variations in betas. Gagliardini et al. (2016) further develop econometric methods for large panels of individual stocks, modeling both risk premia and risk loadings as parametric functions of macro instruments and stock characteristics.

the input layer to the factor network. Estimation relies on stochastic gradient descent (SGD), with learning rate tuning, LASSO (l_1) penalization, and early stopping for regularization.[9] I refer to these models collectively as "AE".[10]

For the conditional autoencoder models, I split the full sample into training, validation, and testing sets. The initial training period is 1960–1977 (18 years), the validation period is 1978–1989 (12 years), and the testing period is 1990–1991 (1 year). Following the literature (e.g., Gu et al., 2020), I refit the models annually. At each refit, the training sample expands by one year, while the validation sample is rolled forward with a fixed length, always including the most recent 12 years. This setup yields an out-of-sample period from 1990 to 2024, totaling 35 years. Since non–deep learning models typically do not require hyperparameter tuning, I combine the training and validation samples for estimation and use the same 1-year testing window for out-of-sample evaluation.

3. Factor Investing versus Factor-Neutral Investing

3.1. Investment Performance of Factor Investing

Table 1 reports the investment performance metrics of factor mean-variance portfolios (FMV) across different factor models and numbers of factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%)[11], and hit rates (%)[12]. I scale the mean-variance portfolio weights each month to target a 15% annualized volatility using historical estimates of mean returns and the covariance matrix. Performance is shown both before transaction costs ("gross") and after proportional transaction costs ("net"). Investment performances with price impact costs display similar implications and are shown in Appendix .

Before accounting for transaction costs, the performance of factor-based portfolios is striking, with machine learning models clearly dominating the conventional FF and PCA benchmarks. For example, portfolios with five or six IPCA and AE factors achieve Sharpe ratios above 3.0 and annualized gross returns exceeding 80%, highlighting the large potential

[9] Other machine learning approaches include Feng et al. (2024), who use feed-forward networks to map characteristics into deep characteristics that generate latent deep factors, and Chen et al. (2024), who incorporate no-arbitrage directly into the loss function via a generative adversarial network (GAN) framework. Their architecture pairs an SDF network that constructs the pricing kernel with a conditional network that selects assets and moments yielding the largest mispricings, iterating until arbitrage opportunities are eliminated.

[10] I use two hidden layers in the beta network, with 32 and 16 neurons, respectively. The empirical results are robust to the choice of network depth. For robustness, I also consider architectures with a single hidden layer and with three hidden layers in Appendix **??**.

[11] Maximum drawdown (MDD): largest peak-to-trough decline in investment values over a specific period.

[12] Hit rate: percentage of months in which the portfolio generates a positive return.

gains from exploiting systematic factor exposures. These results are consistent with Kelly et al. (2019) and Gu et al. (2021). Moreover, maximum drawdowns for these portfolios are as low as 20% and hit rates approach 100%, indicating highly favorable risk–return trade-offs in gross terms.

However, these extraordinary results do not survive once transaction costs are incorporated. Net Sharpe ratios and net returns fall substantially for all models, and in the case of IPCA and AE, net performance turns highly negative, reflecting the extreme trading turnover required to implement such strategies. This high turnover not only erodes mean returns but also inflates standard deviations and maximum drawdowns, making these strategies economically unviable. The contrast highlights the importance of accounting for practical frictions: while machine learning methods can generate striking gross performance, their real-world implementability is severely limited. By comparison, factor mean-variance portfolios based on FF and PCA deliver modest but positive net Sharpe ratios, outperforming machine learning models once costs are considered. Nonetheless, their net Sharpe ratios remain lower than those of factor-neutral investing in Table 2, although maximum drawdowns and hit rates are broadly similar.

To enhance multi-factor portfolio performance, I conduct factor timing (Haddad et al., 2020) and volatility management (Moreira and Muir, 2017; DeMiguel et al., 2024). Prior work has shown that scaling down factor exposure during high-volatility periods improves Sharpe ratios. Moreira and Muir (2017) (MM) manage the FMV portfolio as a whole based on portfolio volatility, while DeMiguel et al. (2024) (DMU) manage each factor individually based on its factor volatility and then optimally combine them with their managed counterparts. Both approaches significantly improve multi-factor portfolio performance. I replicate their results in Tables B.3 and B.4, confirming that MM and DMU portfolios generate higher Sharpe ratios for FF and PCA factors, both before and after transaction costs.[13] By contrast, while IPCA and AE improve mean-variance efficiency in theory, they fail as implementable trading strategies due to extreme transaction costs.

In summary, although the literature has highlighted the remarkable factor mean-variance efficiency via machine learning methods, with reported Sharpe ratios easily exceeding 3.0, these results are not economically meaningful once transaction costs are accounted for. In practice, factor-neutral investing through the factor-neutral strategy delivers stronger and more robust investment performance, both before and after costs, than factor-based mean-variance portfolios.

[13]For FF models, the improvement of MM and DMU is clearer when expanding the factor set beyond the six standard factors to include a BAB factor and two mispricing factors from Stambaugh and Yuan (2017). In unreported results, the net-of-cost (before-cost) Sharpe ratios are 0.74 (0.96) for FMV, 0.82 (1.05) for MM, and 0.83 (1.03) for DMU portfolios.

Table 1: Investment Performance of Fator Mean-Variance Portfolios

Models	# Factors	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
		Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
FF	1	0.53	0.51	8.0	7.8	15.1	15.1	54.3	55.1	62.9	62.9
	3	0.46	0.38	8.1	6.8	17.9	17.9	57.2	59.2	56.9	56.7
	6	0.64	0.50	17.2	13.6	26.9	27.1	79.6	82.5	61.7	60.5
	9	0.63	0.51	15.7	12.7	24.9	25.1	73.3	76.4	64.0	62.6
PCA	1	0.46	0.44	7.1	6.8	15.3	15.4	51.1	51.6	60.0	60.0
	3	0.58	0.52	10.2	9.3	17.7	17.8	71.9	73.0	66.0	66.0
	6	0.59	0.49	12.2	10.3	20.8	20.9	63.5	65.8	60.7	59.8
	9	0.93	0.79	15.3	13.1	16.4	16.7	47.6	51.4	65.5	65.0
IPCA	1	0.42	-4.76	6.5	--	15.5	++	44.6	++	57.6	0.0
	3	1.82	-3.48	35.2	--	19.3	++	58.4	++	74.3	0.0
	6	3.99	-4.66	85.7	--	21.5	++	19.6	++	91.0	0.0
	9	3.88	-4.69	87.8	--	22.6	++	21.0	++	91.2	0.0
AE	1	0.52	-5.11	8.4	--	16.1	++	42.6	++	59.2	0.0
	3	1.73	-4.11	34.8	--	20.1	++	59.1	++	73.1	0.0
	6	3.85	-4.89	85.7	--	22.2	++	20.7	++	89.7	0.0
	9	3.68	-4.97	84.3	--	22.9	++	21.2	++	90.2	0.0

Notes: The table reports the performance of factor mean-variance portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%), and hit rates (%). I scale the mean-variance portfolio weights each month to target a 15% annualized volatility using historical estimates of mean returns and the covariance matrix. Performance is shown both before transaction costs ("gross") and after transaction costs ("net"). "++" indicates a number above 100 (%), and "--" indicates a number below -100 (%).

3.2. Investment Performance of Factor-Neutral Investing

This section examines the investment performance of factor-neutral portfolios as actual investment strategies. The factor-neutral portfolios are constructed on a purely out-of-sample basis by solving the constrained optimization problem (6) using the beta estimates through t, and tracking the post-formation $t+1$ return.

Table 2: Investment Performance of factor-neutral Portfolios

Models	# Factors	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
		Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
FF	1	1.49	1.17	25.4	20.1	17.0	17.1	37.0	44.8	67.6	65.0
	3	1.49	1.17	25.7	20.3	17.2	17.3	33.6	43.0	67.9	65.2
	6	1.39	1.08	24.2	18.8	17.4	17.4	32.8	41.5	66.9	63.3
	9	1.38	1.07	24.4	18.8	17.7	17.7	32.7	40.2	66.7	62.6
PCA	1	1.49	1.17	25.3	20.0	17.0	17.1	37.1	44.7	67.6	64.8
	3	1.49	1.17	25.5	20.2	17.1	17.2	36.9	44.6	66.9	64.3
	6	1.39	1.07	24.3	18.8	17.4	17.6	36.5	45.5	66.4	62.1
	9	1.32	0.99	23.5	17.8	17.8	17.9	37.9	46.8	64.0	60.5
IPCA	1	1.48	1.15	24.7	19.5	16.7	16.9	41.1	49.3	68.1	65.0
	3	1.42	1.10	23.7	18.4	16.6	16.8	43.2	50.9	67.6	63.8
	6	1.25	0.92	19.8	14.6	15.8	15.9	41.4	49.3	66.7	63.3
	9	1.23	0.88	18.7	13.5	15.1	15.3	39.2	47.2	66.0	62.4
AE	1	1.48	1.16	24.9	19.6	16.8	16.9	40.8	49.4	67.9	65.0
	3	1.34	1.02	22.3	17.1	16.7	16.7	38.3	46.6	65.5	63.8
	6	1.36	0.98	19.9	14.6	14.7	14.9	38.9	45.3	66.4	62.4
	9	1.12	0.77	17.8	12.4	15.9	16.1	36.4	43.7	65.7	63.1

Notes: The table reports the performance of optimal factor-neutral portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%), and hit rates (%). Factor-neutral portfolio weights are rescaled each month to target a 15% annualized volatility, based on historical estimates of betas and the covariance matrix. Performance is shown both before transaction costs ("gross") and after transaction costs ("net").

Table 2 reports backtesting metrics for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors, including annualized Sharpe ratios, annualized mean returns (%), annualized standard deviations (%), maximum drawdowns (%), and hit rates (%). Although the models are re-estimated annually, portfolio positions are rebalanced monthly based on the most recent estimates of betas. The conditional optimal factor-neutral portfolio weights are computed each month using equation (7), and then rescaled to target a 15% annualized volatility using historical estimates of betas and the covariance matrix. Investment performance is reported both before transaction costs ("gross") and after proportional transaction costs ("net"). These portfolios are constructed based on characteristics-sorted portfolio betas.[14]

During the out-of-sample period from January 1990 to December 2024, the factor-neutral portfolios deliver consistently strong investment performance, even after accounting for transaction costs. The Sharpe ratios across all specifications are remarkably high relative to the market portfolio benchmark, whose annualized Sharpe ratio is 0.53 before costs and 0.52 after costs.[15] By comparison, the factor-neutral strategies achieve annualized net Sharpe ratios between 0.72 and 1.17, with many gross Sharpe ratios exceeding 1.3. A first observation is that the FF and PCA models stand out for their stability and persistence of high Sharpe ratios. For both models, the one- and three-factor portfolios achieve a gross Sharpe ratio of 1.49 (1.17 net), more than doubling the performance of the market. As dimensionality increases, performance modestly declines: the six-factor versions still deliver net Sharpe ratios around 1.07–1.08, which remain very strong in absolute terms. In contrast, IPCA and AE display a sharper deterioration in Sharpe ratios as the number of factors rises. For IPCA, the one-factor specification attains a net Sharpe ratio of 1.15, competitive with FF and PCA, but performance falls to 0.88 with six factors. AE follows a similar trajectory, starting at 1.16 net with one factor and declining to 0.77 with six factors. These declines are consistent with the findings in Section 4.3: as model specification improves—particularly through machine-learning methods—the profitability of factor-neutral portfolios diminishes. The strong performance of one-factor strategies across all models is particularly striking. Net Sharpe ratios of 1.17 for FF and PCA, 1.15 for IPCA, and 1.16 for AE indicate that even simple specifications—though far from fully capturing true risk structure—translate into highly profitable investment strategies. This suggests that from an investment perspective, a simple market-neutral strategy may suffice, while extending to multi-factor, beta-neutral portfolios offers limited incremental benefits in real-world trading.

Turning to returns and volatility, the mean annualized net returns range between 12% and 20% depending on the model and factor count. This compares favorably to typical equity premium estimates. Standard deviations are tightly centered around 15%–18% by construction, reflecting the monthly volatility-scaling rule. Importantly, volatility is well-controlled across models and specifications, which means that differences in Sharpe ratios are primarily driven by variations in mean returns rather than excess volatility. The risk management properties of these portfolios are more nuanced. Maximum drawdowns (MDD) are typically around 40–50% net of costs, comparable in magnitude to that of the market portfolio (51.9%). This suggests that, although the strategies deliver superior risk-adjusted returns on

[14] I do not use individual stock betas to construct factor-neutral portfolios, since inverting the individual-stock covariance matrix is notoriously unstable and beta estimates are considerably noisier than those obtained from characteristic-sorted portfolios.

[15] Transaction costs associated with trading the market portfolio are minimal.

average, they remain exposed to occasional severe losses. In other words, high Sharpe ratios coexist with drawdowns of a scale similar to traditional equity markets. Importantly, however, these drawdowns are smaller than those of the standard BAB and momentum strategy, whose maximum drawdown reached 74.7% and 75.3%, respectively, over the same period. Finally, the hit rate (the fraction of months with positive returns) provides additional insight into return consistency. Across models, hit rates range from roughly 60% to 70%, closely matching the market's 64.5%. This suggests that factor-neutral strategies succeed not by winning more often than the market, but by delivering larger average payoffs per winning trade, consistent with their elevated Sharpe ratios.

Overall, the evidence indicates that factor-neutral portfolios represent attractive investment opportunities, particularly in parsimonious one-factor implementations. Despite lacking systematic factor risk exposures by construction, they achieve high and persistent risk-adjusted returns, maintain volatility near targeted levels, and generate hit rates on par with the market. Particularly noteworthy is that one-factor models based on FF and PCA specifications deliver net Sharpe ratios near 1.2—more than double the market's Sharpe ratio—underscoring their remarkable efficiency as practical investment strategies.

3.2.1. Portfolio Positions

I investigate the portfolio weights on individual stocks for factor investing and factor-neutral investing. Panel (A) of Table 3 reports the maximum short positions assigned to individual stocks for all models with 1, 3, 6, and 9 factors. The results show that the largest short positions for factor investing are modest for conventional FF and PCA models, but we observe extreme weights for ML-based models such as IPCA and AE. Positions for factor-neutral investing do not exceed 3.4% of the portfolio' value, indicating that these portfolios are well diversified and free from extreme concentration risk. Figures B.10 - B.11 address this concern by plotting the maximum, minimum, 1st percentile, and 99th percentile of factor-neutral portfolio weights on individual stocks from January 1990 to December 2024. The results show that the largest absolute weight assigned to any individual stock does not exceed 5%, indicating that the portfolios are well diversified and free from extreme concentration.

I also examine their leverage ratios. Following Fama and French (2015), the leverage ratio is defined as the total value of short positions divided by the total value of the portfolio. Panel (B) of Table 3 shows that for all models with 1, 3, 6, and 9 factors. Similarly, the leverage ratios explodes for factor investing with ML-based models, and remain at reasonable and implementable levels for factor-neutral investing, ranging from 1.84 to 4.10.

These result reflect the shrinkage benefits of imposing factor-neutrality in portfolio construction. Consider the expected return decomposition $\mu = \alpha + \beta' \lambda$ and the covariance matrix

Table 3: Maximum Short Positions and Leverage Ratios

Strategies		Panel A: Maximum Short Positions (%)				Panel B: Leverage Ratios			
		FF	PCA	IPCA	AE	FF	PCA	IPCA	AE
FI	1	0	0	37	37	0	0	1.35	1.38
	3	2.98	2.99	608	634	3.87	3.85	5,522	5,623
	6	3.01	3.01	1,435	1,451	3.98	3.90	12,924	13,343
	9	3.09	3.10	1,575	1,592	4.02	4.03	14,356	14,671
FNI	1	3.12	3.12	3.11	3.11	3.96	3.95	3.91	3.91
	3	3.12	3.12	3.09	3.11	4.00	3.97	3.86	3.86
	6	3.25	3.25	3.07	3.12	4.10	4.10	3.78	3.81
	9	3.34	3.34	3.03	3.09	4.14	4.24	3.74	3.78

Notes: This table reports the maximum short positions (Panel A) on individual stocks and the portfolio leverage ratio (Panel B) for FF, PCA, IPCA, and AE models with 6 factors. Both in-sample and out-of-sample portfolio constructions are considered. Transaction costs include proportional trading costs and price impact costs.

decomposition $\Sigma = \beta'\Sigma_f\beta + E$, where α captures pricing errors, $\beta'\lambda$ represents factor risk compensation, Σ_f is the systematic factor covariance matrix, and E denotes idiosyncratic risk. The factor-neutral portfolio is constructed by solving

$$\max_{\omega} \ \omega'\mu - \frac{\gamma}{2}\omega'\Sigma\omega \quad \text{s.t.} \quad \omega'\beta = 0_K,$$

which, substituting μ and Σ, becomes

$$\max_{\omega} \ \omega'\alpha - \frac{\gamma}{2}\sum_{i=1}^{N}\omega_i^2 E_{ii} - \frac{\gamma}{2}\sum_{i\neq j}\omega_i\omega_j E_{ij}.$$

The first term penalizes pricing errors, while the second and third terms correspond to a ridge penalty and a hedging magnifier, respectively. The key advantage of this formulation is that factor-neutral portfolios typically entail smaller position sizes. Intuitively, when the off-diagonal elements of E are small, the shrinkage effect induced by the ridge penalty dominates the magnifying effect from cross-asset hedging, leading to more stable and feasible portfolio weights.

3.2.2. Market Betas and Idiosyncratic Risks of Long vs Short Positions

Another well-known beta-neutral strategy is the BAB portfolio of Frazzini and Pedersen (2014), which takes long positions in low-beta stocks and short positions in high-beta stocks.

The strategy builds on the "beta anomaly" documented in the literature.[16] This naturally raises the question: do the proposed optimal factor-neutral portfolios resemble BAB in their positioning? In other words, do they systematically go long low-beta stocks and short high-beta stocks? Figure B.12 - B.13 plot the weighted average market betas of the long and short sides of the optimal factor-neutral portfolios. The evidence shows otherwise: long and short positions have similar weighted average betas, with no clear tilt toward low- or high-beta stocks. By contrast, Figure B.14 illustrates the BAB case, which clearly features low-beta longs and high-beta shorts.

Liu et al. (2018) argue that the beta anomaly is largely driven by the strong positive cross-sectional correlation between beta and idiosyncratic volatility (IVOL). Figure B.17 confirms that the BAB strategy mainly takes long positions in low-IVOL stocks and short positions in high-IVOL stocks. To explore whether the optimal factor-neutral portfolios display a similar pattern, I examine the weighted average IVOL of their long and short sides. Figures B.15–B.16 show that long and short positions have comparable average IVOL, with only a mild tilt toward low-IVOL in longs and high-IVOL in shorts.

3.2.3. Other Benchmark Strategies

I also consider prominent investment strategies proposed in the literature. Following Gu et al. (2020) and Gu et al. (2021), I construct long–short portfolios based on model-predicted returns. For each model, stocks are sorted into deciles according to their out-of-sample predicted returns. The strategy goes long the top-decile (highest expected return) stocks and short the bottom-decile (lowest expected return) stocks, with equal weights applied within each leg. Table B.5 reports the performance of these prediction-sorted portfolios. Since unconditional models are poor at out-of-sample return prediction (see Table B.2), FF and PCA yield non-positive Sharpe ratios. In contrast, machine learning methods substantially improve predictability and, consequently, portfolio performance. However, their out-of-sample net-of-cost Sharpe ratios remain below those of factor-neutral strategies.

I also examine conventional model-free investment strategies. Table B.6 reports results

[16] There are a number of theories explaining the beta anomaly. Several theories have been proposed to explain the beta anomaly. For example, Black (1972), Frazzini and Pedersen (2014), and Asness et al. (2020) argue that leverage-constrained investors, unable to borrow to lever up market exposure, instead tilt toward high-beta stocks. This excess demand inflates the prices of high-beta stocks, leading to their underperformance. Bali et al. (2017) attribute the anomaly to behavioral biases. High-beta stocks often have lottery-like features (e.g., high skewness, high maximum daily returns), which attracts investors who are willing to accept lower average returns for the small chance of a large payoff. Liu et al. (2018) argue that the anomaly reflects the correlation between beta and idiosyncratic volatility (IVOL): high-beta stocks tend to have high IVOL, which deters arbitrage and makes them more likely to be overpriced. Schneider et al. (2020) find that low-beta stocks carry higher coskewness risk—underperforming in extreme downturns—so their higher returns may reflect rational compensation for higher-order risks rather than mispricing.

for the market portfolio (MKT), the mean–variance portfolio (MVP), the global minimum-variance portfolio (GMV), the equal-weighted portfolio (EW), the risk-parity portfolio (RP), and the Kan et al. (2022) optimal portfolio (KWZ). Consistent with the literature, MVP underperforms relative to simple rules such as GMV, EW, and RP. Among these model-free strategies, the KWZ rule achieves a before-cost Sharpe ratio of 1.31 and a net-of-cost Sharpe ratio of 1.02, offering attractive investment opportunities.

4. Implications on Factor Model Comparison

4.1. Conceptual Framework

A pivotal shift in the methodology of factor model comparison is proposed by Barillas and Shanken (2017), who argue that a proper comparison should assess a model's ability to price all available assets, including the factors from competing models. Under this framework, the choice of test assets becomes mathematically irrelevant for model comparison. As a result, the exercise reduces to comparing the maximum squared Sharpe ratio $SR^2(f)$ that each model's factors can generate. The paper's title, *"Which Alpha?"*, highlights that the relevant alpha for model comparison is not the test-asset alpha but the "excluded-factor alpha"—the pricing error of one model's factors with respect to another's. Building directly on their earlier work, Barillas et al. (2020) develop asymptotically valid tests to compare the squared Sharpe ratios of both nested and, crucially, non-nested models. For nested models, the standard GRS F-test on the excluded-factor alphas is sufficient. For non-nested models, they derive the asymptotic distribution of the difference in sample squared Sharpe ratios, allowing for a formal statistical test of which model is superior[17].

With the factor maximum Sharpe ratio framework becoming the dominant approach in model comparison, subsequent research has addressed several important practical challenges. First, in-sample (ex-post) Sharpe ratios are known to be upward biased due to estimation risk. Fama and French (2018) and Kan et al. (2024) compare models based on the out-of-sample Sharpe ratios of their factors. Kan et al. (2024) further show that models with more factors are subject to greater estimation risk, which leads to a larger deterioration in out-of-sample performance. They propose that a more complex model must not only have a higher population Sharpe ratio, but one that exceeds a "break-even" Sharpe ratio—a higher threshold required to offset its greater estimation risk relative to a simpler benchmark.

[17]Another related metric of model misspecification, introduced by Hansen and Jagannathan (1997), is the HJ-distance, which measures the distance between a model's stochastic discount factor (SDF) and the closest valid SDF that prices all assets correctly. Kan and Robotti (2009) use the HJ-distance for model comparison, and Barillas et al. (2020) show that comparing models using a modified HJ-distance is equivalent to comparing their maximum squared Sharpe ratios.

Second, real-world frictions may materially impact and reverse the conclusions of model comparisons. Fama and French (2015) show that the factor tangency portfolios that generate high maximum Sharpe ratios are not practical for real-world investors since they require extremely high levels of short selling. The paper re-calculates the maximum Sharpe ratios under a no-short-selling constraint, which better reflects the opportunity set of a long-only investor. Detzel et al. (2023) compare the Sharpe ratios accounting for transaction costs of factors.

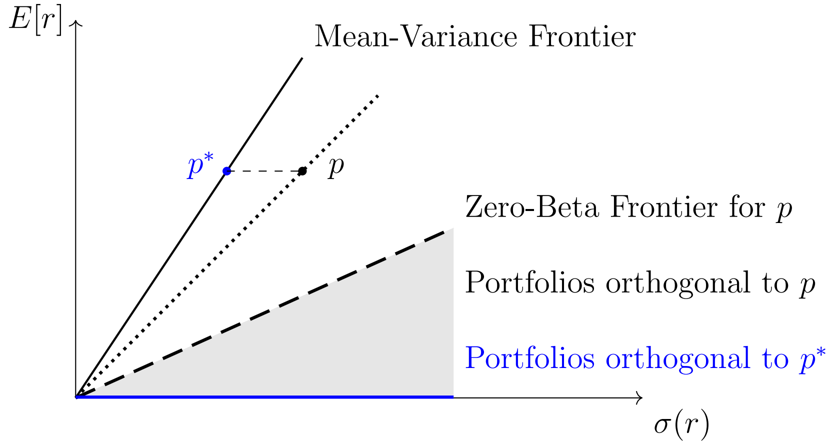
However, Section 3 demonstrates that the factor mean-variance portfolio can result in extreme weights on individual stocks, particularly for ML-based models. This infeasibility raises concerns about the validity of model comparisons. For instance, while the IPCA models yield negative factor Sharpe ratios after accounting for trading costs, the CAPM produces positive values. Does this imply that the CAPM outperforms the IPCA models? Not necessarily. The issue lies in comparing Sharpe ratios of factor portfolios with vastly different compositions. To address this, I propose evaluating factor models using factor-neutral portfolios, which tend to exhibit similarly moderate portfolio compositions across a broad range of factor models.

Wang (2025) proves that the maximum Sharpe ratio of zero-beta portfolios constitute a quantitative measure of model misspecification and it equals to the difference between the Sharpe ratios of the true tangency portfolio and the factor mean-variance portfolio. This relation can be understood in the mean-variance space in Figure 1.

Consider the standard textbook mean-variance framework. The mean-variance frontier for zero-investment portfolios is a straight line passing through the origin[18]. According to asset pricing theory, a multivariate beta equation model corresponds to a mean-variance efficient portfolio on the frontier, denoted by p^* and illustrated in Figure 1. Proposition 3 of Wang (2025) indicates that the factor-neutral portfolios (also known as orthogonal portfolios, originally introduced by Roll, 1980) with respect to p^* lie on the blue horizontal solid line, characterized by zero expected returns. In contrast, a misspecified factor model corresponds to an inefficient portfolio p, also depicted in Figure 1. The zero-beta portfolios with respect to p fall within the shaded area. The black dashed line is referred to as the zero-beta frontier—the set of portfolios that have minimum variance for a given level of mean return and zero beta with respect to p. The Sharpe ratios of these zero-beta portfolios are therefore bounded above by the slope of the zero-beta frontier. The formula of the zero-beta frontier is:

[18] Conventional mean-variance analysis typically focuses on unit-investment portfolios, where portfolio weights sum to 1. In that case, the mean-variance frontier takes the familiar parabolic shape.

Figure. 1. The Zero-Beta Frontier and Zero-Beta Portfolios



Notes: This figure shows the zero-investment mean-variance frontier and the zero-beta (orthogonal) sets for two assets, p^* and p, in the mean-standard deviation space. p^* is on the mean-variance frontier. p has the same mean as p^*. The blue horizontal solid line on the x-axis represents the orthogonal set with respect to p^*. The shaded area represents the orthogonal set with respect to p. The black dashed line represents the zero-beta frontier with respect to p.

$$SR^2(z) = SR^2(p^*) - SR^2(p) \tag{15}$$

Equation (15) shows that the maximum Sharpe ratio attainable by zero-beta portfolios equals the difference in Sharpe ratios between the true tangency portfolio and the model-implied factor mean-variance portfolio. This "Sharpe ratio spread" represents the gain in Sharpe ratio from adding test assets to the factor set, providing a economically interpretable measure of model misspecification. Notice that I have intentionally consider an asset universe that is able to replicate all factors including prespecified Fama-French factors and statistical latent factors. Hence, the tangency portfolio Sharpe ratio $SR^2(p^*)$, which captures both testing portfolios and factors, are shared across factor models, validating the model comparison.

Comparing models based on the maximum Sharpe ratio of factor-neutral portfolios, as introduced in this paper, is fundamentally equivalent to comparing the maximum Sharpe ratio of the model factors. The slope of the zero-beta frontier is determined by the difference between the Sharpe ratio of mean-variance efficient portfolios and the Sharpe ratio of the factor portfolio p. Therefore, a lower maximum Sharpe ratio of zero-beta portfolios implies a higher Sharpe ratio of the factor portfolio given that the Sharpe ratio of mean-variance efficient portfolios is constant depending on the asset universe. In addition, utilizing zero-beta portfolios offers several advantages. First, the method is applicable regardless of whether the model is unconditional or conditional, whether the risk-free rate is observable or not, and whether the factors are traded, non-traded, or purely statistical. As long as zero-beta portfolios can be constructed, model comparison can be performed. In contrast, if the risk-free rate is unobservable, it becomes ambiguous how to calculate Sharpe ratios for the factors, whereas the Sharpe ratios of zero-investment portfolios do not rely on the risk-free rate. Second, as suggested by Fama and French (2018) and Kan et al. (2024), Sharpe ratios of zero-beta portfolios can be compared both in-sample and out-of-sample. Third, following the ideas of Fama and French (2015) and Detzel et al. (2023), leverage and transaction costs can also be incorporated when calculating Sharpe ratios. Regarding the short-selling constraint, although no such restriction is imposed on zero-beta portfolios in this paper, the portfolio weights are generally small, mitigating the leverage concerns highlighted by Fama and French (2018). Unlike Barillas and Shanken (2017), my approach is not independent of the choice of test assets, as the zero-beta portfolios are constructed directly from the selected test assets.

A major advantage of focusing on zero-beta portfolios is their broad applicability across different factor models. This point can be seen by considering several modeling choices. First, while most research assumes the existence of a risk-free asset and proxies it with Treasury yields, recent studies show that Treasury yields are depressed by a convenience yield, making them an imperfect proxy for the frictionless risk-free rate. Zero-beta portfolios, however, are unaffected by this issue, since zero-investment strategies cancel out the unknown risk-free rate. By contrast, models that include a level factor (e.g., the market factor in traditional unconditional models) require a risk-free rate to compute the Sharpe ratio of the factor portfolio (see the dotted line in Figure 1). Second, the literature considers a wide range of risk factors, including traded factors, macroeconomic factors, and statistical factors (e.g., machine learning factors). Traded factors and mimicking portfolios for macroeconomics factors fit naturally into mean–variance analysis, whereas statistical factors can lose interpretability in that framework, and measuring their distance from the mean–variance frontier is inherently ambiguous.[19] Zero-beta portfolios, in contrast, allow us to remain agnostic about the identities of the factors while still maintaining a direct link to the efficiency of the factor portfolio. Third, while most conventional methods are designed for unconditional models, the zero-beta framework can easily be extended to conditional settings. Finally, the following proposition formalizes that the zero-beta frontier remains linear even in out-of-sample analysis and when transaction costs are considered.

Comments on the bounded Sharpe ratio of zero-beta portfolios: The statement

[19] Section **??** shows that the machine learning factors are not economically meaningful due to high turnovers and transaction costs.

of bounded Sharpe ratio of zero-beta portfolios is not inconsistent with the idea that arbitrage opportunities can have unbounded Sharpe ratios. The Sharpe ratios are bounded because they are constrained by the model structure. When I am constructing zero-beta portfolios within the factor model by imposing constraints (e.g., zero exposure to estimated betas and zero-investments), I only use information allowed by the model rather than discover true arbitrage opportunities. The maximum Sharpe ratio of such zero-beta portfolios reflects how misspecified the model is instead of an actual arbitrage opportunity. In summary, while true arbitrage opportunities (in the market) may have infinite or unbounded Sharpe ratios, model-implied zero-beta portfolios can only reflect misspecification within the model, and therefore have bounded Sharpe ratios. Therefore, I use the maximum Sharpe ratio among model-based zero-beta portfolios as a diagnostic tool — not a literal arbitrage detector.

4.2. Statistical Inference

Let $MSR_A^{(z)}$ denote the maximum Sharpe ratio of zero-beta portfolios for model A, and $MSR_B^{(z)}$ the maximum Sharpe ratio of zero-beta portfolios for model B. The null hypothesis for model comparison becomes:

$$\mathcal{H}_0 : MSR_A^{(z)} - MSR_B^{(z)} \geq 0, \quad \mathcal{H}_1 : MSR_A^{(z)} - MSR_B^{(z)} < 0 \tag{16}$$

This is a one-sided test of whether model A's maximum Sharpe ratio of zero-beta portfolios is significantly lower than that of model B. Not rejection of the null implies that model A is not significantly better than model B. Conversely, rejecting the null supports the conclusion that model A is less misspecified than model B.

Statistical inferences are based on the following specialized bootstrapping procedure. First, I compute the observed difference between the maximum Sharpe ratios of model A versus model B, $\Delta_{obs} = MSR_A^{(z)} - MSR_B^{(z)}$. According to the previous propositions, since all zero-beta portfolios on the zero-beta frontier share the same ex-post Sharpe ratios over the full sample, I can choose a random target mean return r to compute the maximum Sharpe ratio based on equation (**??**). Next, I generate bootstrap samples for both portfolio returns by resampling the original return series in blocks of random length, following the stationary block bootstrap method of Politis and Romano (1994). For each bootstrap sample, I calculate the Sharpe ratios and record the differences $\Delta_{boot} = MSR_{A,boot}^{(z)} - MSR_{B,boot}^{(z)}$. The empirical p-value is then computed as the proportion of bootstrap differences that are less than or equal to the observed difference.

Table 4: Within-Model Class Comparison Results

# Factors	1	3	6	9
Panel A: In-Sample Comparison				
FF	2.97	2.96	2.91	2.89
PCA	2.97	2.96	**2.87***	**2.87***
IPCA	3.45	3.34	**3.12***	**3.01***
AE2	3.45	3.45	**2.70***	**2.18***
Panel B: Out-of-Sample Comparison				
FF	1.17	1.17	1.08	1.07
PCA	1.17	1.17	**1.07***	**0.99***
IPCA	1.15	1.10	**0.92***	**0.88***
AE2	1.05	**0.92***	**0.86***	**0.72***

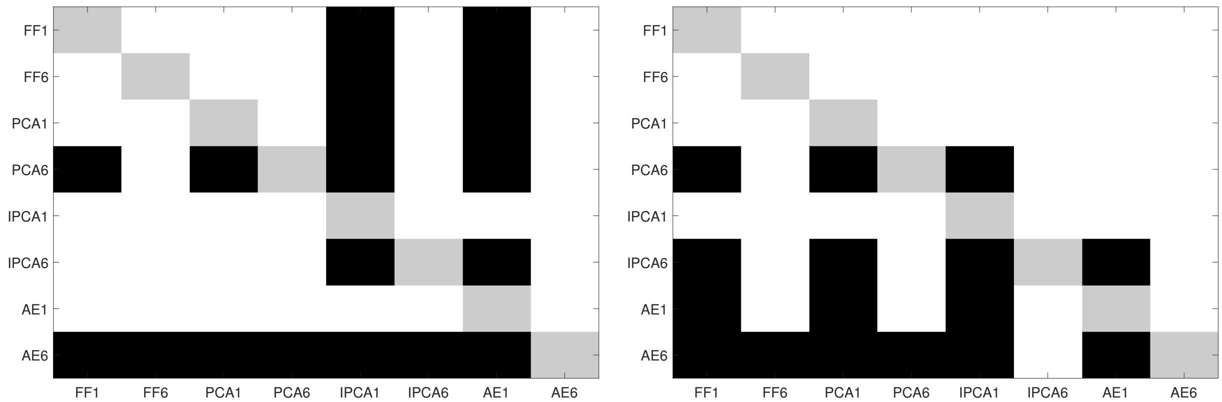
Notes: This table reports in-sample and out-of-sample annualized maximum Sharpe ratios attainable by zero-investment, zero-beta portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors. Out-of-sample calculations incorporate transaction costs. Within each model class, the 1-factor case serves as the benchmark; an asterisk (*) indicates that the Sharpe ratio is significantly lower than the benchmark at the 5% level, based on p-values computed using the stationary block-bootstrap method of Politis and Romano (1994).

4.3. Empirical Results

As discussed in Section 4.2, the maximum Sharpe ratio attainable by zero-investment, zero-beta portfolios can serve as an indicator of model misspecification. Recall from Proposition **??** that the zero-beta frontier remains linear for both in-sample and out-of-sample comparisons. Consequently, the maximum Sharpe ratio is easy to visualize, as it corresponds to the slope of the zero-beta frontier in mean–standard deviation space. Figures B.18 and B.19 plot the in-sample and out-of-sample zero-beta frontiers for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors. A flatter zero-beta frontier indicates a smaller distance between the factors and the efficiency frontier, signaling a better-specified model. Visual inspection suggests that adding more factors in the FF model does not improve specification relative to the 1-factor case (CAPM), whereas additional factors in PCA and IPCA appear beneficial. Notably, the 6-factor and 9-factor AE models may dominate the smaller-factor versions. Formal analysis follows.

Focusing on within-model class comparisons, Table 4 reports in-sample and out-of-sample annualized maximum Sharpe ratios attainable by zero-investment, zero-beta portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 6, and 9 factors. Out-of-sample calculations

Figure. 2. Cross-Model Class Comparison Results



(a) In-Sample (b) Out-of-Sample

Notes: Panel (a) shows the significance of in-sample model comparison. Panel (b) shows the significance of out-of-sample model comparison. I compare eight models: the 1-factor and 6-factor versions of FF, PCA, IPCA, and AE. In the plots, models on the horizontal axis serve as the benchmark; a black block indicates that the model on the vertical axis delivers a significantly lower attainable maximum Sharpe ratio.

incorporate transaction costs. Within each model class, the 1-factor case serves as the benchmark; an asterisk (*) indicates that the Sharpe ratio is significantly lower than the benchmark at the 5% level, based on p-values computed using the stationary block-bootstrap method of Politis and Romano (1994). A significantly lower Sharpe ratio indicates a better-specified model.

In-sample results show that increasing the number of factors in the FF model does not significantly improve specification relative to the 1-factor case (CAPM). For PCA, IPCA, and AE, the 3-factor specifications do not improve upon the 1-factor case, but the 6- and 9-factor models clearly dominate the smaller-factor versions. Out-of-sample, attainable Sharpe ratios fall substantially—consistent with Fama and French (2018), Kan et al. (2024), and others—yet the conclusions mirror those of the in-sample analysis.

Next, I turn to cross-model class comparisons. This generality is a key advantage of the zero-beta portfolio approach. The Sharpe ratios are reported in Table 4, and Figure 2 visualizes the significance of model improvements using a heat map. I compare eight models: the 1-factor and 6-factor versions of FF, PCA, IPCA, and AE. In the plots, models on the horizontal axis serve as the benchmark; a black block indicates that the model on the vertical axis delivers a significantly lower attainable maximum Sharpe ratio, hence a better-specified model. Focusing first on Panel (a), the CAPM (FF1) and the 1-factor PCA significantly improve upon both IPCA1 and AE1, suggesting that simple factor structures outperform these machine-learning models in the 1-factor case. However, when additional factors are introduced, the relative ranking changes. The 6-factor version of AE displays substantial gains: AE6 dominates all other models. This highlights the benefit of richer factor structures, particularly when using more flexible machine-learning approaches. Turning to Panel (b), which is arguably more informative for model comparisons from an investment perspective. The initial advantage of FF1 and PCA1 disappears; they no longer dominate the 1-factor machine-learning models. Instead, AE1 outperforms all the other 1-factor specifications, showing robustness of the machine learning approach in sparse factor environments. Among the 6-factor models, AE6 again performs strongly, significantly improving upon most alternatives, though it does not dominate IPCA6. This suggests that both IPCA and AE provide strong and competitive specifications when many factors are used, with AE delivering the broadest improvements. Overall, the cross-model comparison suggests that while traditional factor models remain competitive in simpler 1-factor settings, machine-learning approaches become clearly advantageous as the factor dimension grows. In particular, AE stands out as the most robust model, delivering substantial improvements both in-sample and out-of-sample.

5. Conclusion

This paper reexamines the economic viability of factor investing in the era of machine learning (ML). I show that while ML-based factor models deliver impressive in-sample and even out-of-sample Sharpe ratios, their practical performance deteriorates sharply once realistic transaction costs are incorporated. In contrast, factor-neutral investing—constructed as optimal factor-neutral portfolios that exploit model mispricing—offers a robust and implementable alternative. The factor-neutrality constraint functions as a natural shrinkage mechanism, mitigating turnover, reducing leverage, and enhancing resilience to trading frictions.

Empirically, factor-neutral portfolios achieve economically significant investment performance. These strategies maintain modest exposures, exhibit less pronounced diseconomies of scale, and derive value from persistent alpha signals rather than rapidly decaying factor signals. Their profitability also varies systematically with macroeconomic conditions, particularly the stock–bond covariance, suggesting renewed relevance in the post-COVID environment.

Overall, the findings imply that the true economic value of ML-based factor models lies not in the factors they estimate, but in the structure of their errors. Factor-neutral investing provides a general, scalable, and practically implementable framework that bridges the gap between statistical model performance and real-world portfolio profitability.

References

Abdi, F., Ranaldo, A., 2017. A simple estimation of bid-ask spreads from daily close, high, and low prices. The Review of Financial Studies 30, 4437–4480.

Asness, C., Frazzini, A., Gormsen, N. J., Pedersen, L. H., 2020. Betting against correlation: Testing theories of the low-risk effect. Journal of Financial Economics 135, 629–652.

Avellaneda, M., Lee, J.-H., 2010. Statistical arbitrage in the us equities market. Quantitative Finance 10, 761–782.

Bali, T. G., Brown, S. J., Murray, S., Tang, Y., 2017. A lottery-demand-based explanation of the beta anomaly. Journal of Financial and Quantitative Analysis 52, 2369–2397.

Barillas, F., Kan, R., Robotti, C., Shanken, J., 2020. Model comparison with sharpe ratios. Journal of Financial and Quantitative Analysis 55, 1840–1874.

Barillas, F., Shanken, J., 2017. Which alpha? The Review of Financial Studies 30, 1316–1338.

Black, F., 1972. Capital market equilibrium with restricted borrowing. The Journal of business 45, 444–455.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of finance 52, 57–82.

Chamberlain, G., Rothschild, M., 1983. Arbitrage, factor structure, and mean-variance analysis on large asset markets.

Chen, A. Y., Velikov, M., 2023. Zeroing in on the expected returns of anomalies. Journal of Financial and Quantitative Analysis 58, 968–1004.

Chen, L., Pelger, M., Zhu, J., 2024. Deep learning in asset pricing. Management Science 70, 714–750.

Connor, G., 1984. A unified beta pricing theory. Journal of Economic Theory 34, 13–31.

Connor, G., Korajczyk, R. A., 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. Journal of financial economics 15, 373–394.

Connor, G., Korajczyk, R. A., 1988. Risk and return in an equilibrium apt: Application of a new test methodology. Journal of financial economics 21, 255–289.

Cooper, I., Ma, L., Maio, P., Philip, D., 2021. Multifactor models and their consistency with the apt. The Review of Asset Pricing Studies 11, 402–444.

Corwin, S. A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. The journal of finance 67, 719–760.

DeMiguel, V., Martín-utrera, A., Uppal, R., 2024. A multifactor perspective on volatility-managed portfolios. The Journal of Finance 79, 3859–3891.

Detzel, A., Novy-Marx, R., Velikov, M., 2023. Model comparison with transaction costs. The Journal of Finance 78, 1743–1775.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33, 3–56.

Fama, E. F., French, K. R., 2015. Incremental variables and the investment opportunity set. Journal of Financial Economics 117, 470–488.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of financial economics 128, 234–252.

Feng, G., He, J., Polson, N. G., Xu, J., 2024. Deep learning in characteristics-sorted factor models. Journal of Financial and Quantitative Analysis 59, 3001–3036.

Ferson, W. E., Harvey, C. R., 1991. The variation of economic risk premiums. Journal of political economy 99, 385–415.

Frazzini, A., Israel, R., Moskowitz, T. J., 2018. Trading costs, vol. 3229719. SSRN.

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. Journal of financial economics 111, 1–25.

Fung, W., Hsieh, D. A., 2001. The risk in hedge fund strategies: Theory and evidence from trend followers. The review of financial studies 14, 313–341.

Gagliardini, P., Ossola, E., Scaillet, O., 2016. Time-varying risk premium in large cross-sectional equity data sets. Econometrica 84, 985–1046.

Gatev, E., Goetzmann, W. N., Rouwenhorst, K. G., 2006. Pairs trading: Performance of a relative-value arbitrage rule. The review of financial studies 19, 797–827.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society pp. 1121–1152.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Gu, S., Kelly, B., Xiu, D., 2021. Autoencoder asset pricing models. Journal of Econometrics 222, 429–450.

Guijarro-Ordonez, J., Pelger, M., Zanotti, G., 2021. Deep learning statistical arbitrage. arXiv preprint arXiv:2106.04028 .

Haddad, V., Kozak, S., Santosh, S., 2020. Factor timing. The Review of Financial Studies 33, 1980–2018.

Hansen, L. P., Jagannathan, R., 1997. Assessing specification errors in stochastic discount factor models. The Journal of Finance 52, 557–590.

Hasbrouck, J., 2009. Trading costs and returns for us equities: Estimating effective costs from daily data. The Journal of Finance 64, 1445–1477.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. The Review of Financial Studies 28, 650–705.

Huberman, G., 1982. A simple approach to arbitrage pricing theory. Journal of Economic Theory 28, 183–191.

Ingersoll Jr, J. E., 1984. Some results in the theory of arbitrage pricing. The Journal of Finance 39, 1021–1039.

Jensen, T. I., Kelly, B., Pedersen, L. H., 2023. Is there a replication crisis in finance? The Journal of Finance 78, 2465–2518.

Jensen, T. I., Kelly, B. T., Malamud, S., Pedersen, L. H., 2024. Machine learning and the implementable efficient frontier. Swiss Finance Institute Research Paper .

Kan, R., Robotti, C., 2009. Model comparison using the hansen-jagannathan distance. The Review of Financial Studies 22, 3449–3490.

Kan, R., Wang, X., Zheng, X., 2024. In-sample and out-of-sample sharpe ratios of multi-factor asset pricing models. Journal of Financial Economics 155, 103837.

Kan, R., Wang, X., Zhou, G., 2022. Optimal portfolio choice with estimation risk: No risk-free asset case. Management Science 68, 2047–2068.

Karolyi, G. A., Van Nieuwerburgh, S., 2020. New methods for the cross-section of returns. The Review of Financial Studies 33, 1879–1890.

Kelly, B. T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524.

Kim, S., Korajczyk, R. A., Neuhierl, A., 2021. Arbitrage portfolios. The Review of Financial Studies 34, 2813–2856.

Kyle, A. S., Obizhaeva, A. A., 2016. Market microstructure invariance: Empirical hypotheses. Econometrica 84, 1345–1404.

Lettau, M., Pelger, M., 2020. Factors that fit the time series and cross-section of stock returns. The Review of Financial Studies 33, 2274–2325.

Liu, J., Stambaugh, R. F., Yuan, Y., 2018. Absolving beta of volatility's effects. Journal of Financial Economics 128, 1–15.

Moreira, A., Muir, T., 2017. Volatility-managed portfolios. The Journal of Finance 72, 1611–1644.

Politis, D. N., Romano, J. P., 1994. The stationary bootstrap. Journal of the American Statistical association 89, 1303–1313.

Roll, R., 1980. Orthogonal portfolios. Journal of Financial and Quantitative analysis 15, 1005–1023.

Ross, S. A., 1976. The arbitrage theory of capital asset pricing. In: *Journal of Economic Theory*, Elsevier, pp. 341–360.

Schneider, P., Wagner, C., Zechner, J., 2020. Low-risk anomalies? The Journal of Finance 75, 2673–2718.

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. The review of financial studies 30, 1270–1315.

Wang, R., 2025. Resolving the zero-beta rate puzzle. Job Market Paper .

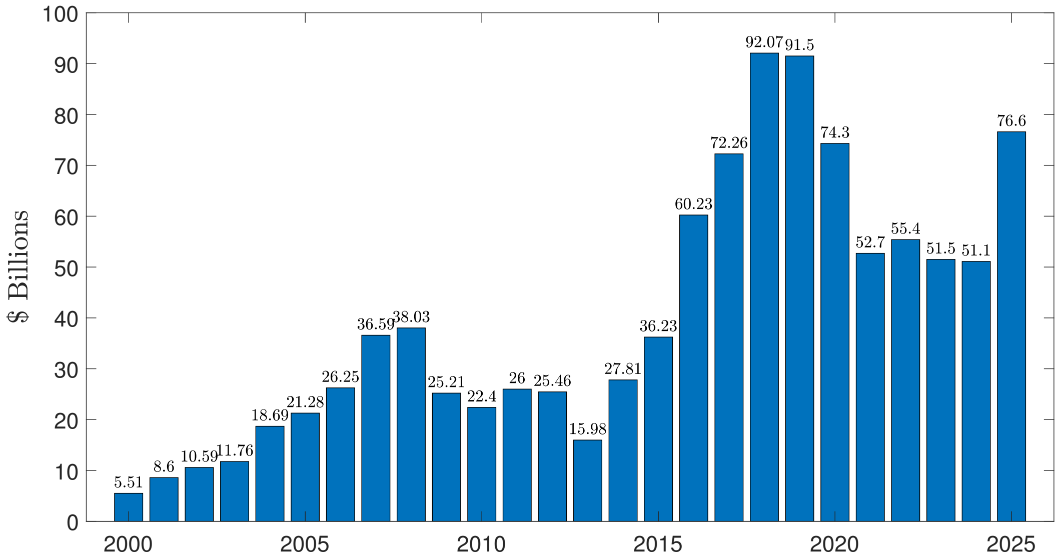
Appendix A. Data

A.1. Stock Characteristics

Appendix B. Additional Empirical Results

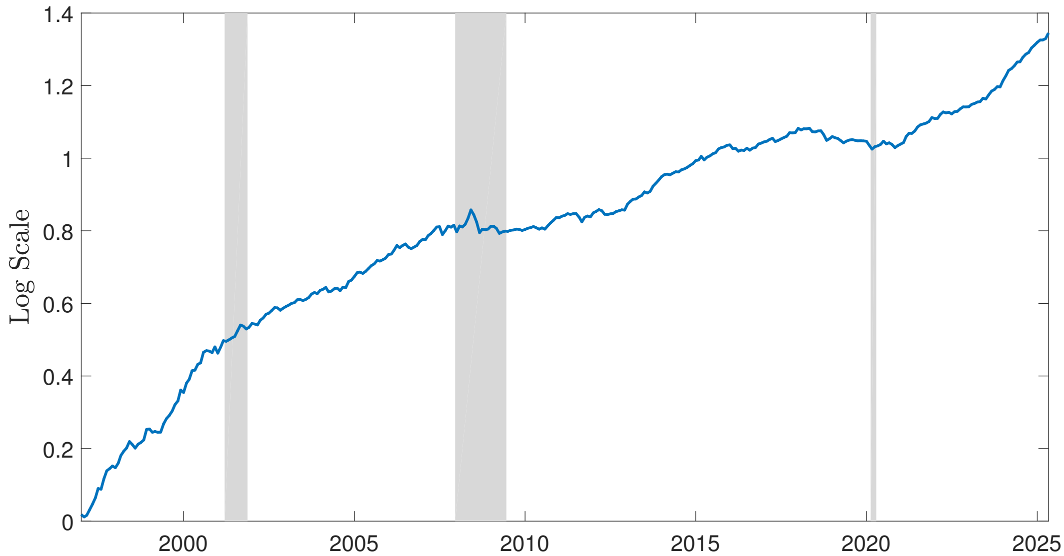
B.1. Equity Market Neutral Hedge Funds

Figure. B.1. Asset Under Management of Equity Market Neutral Hedge Funds



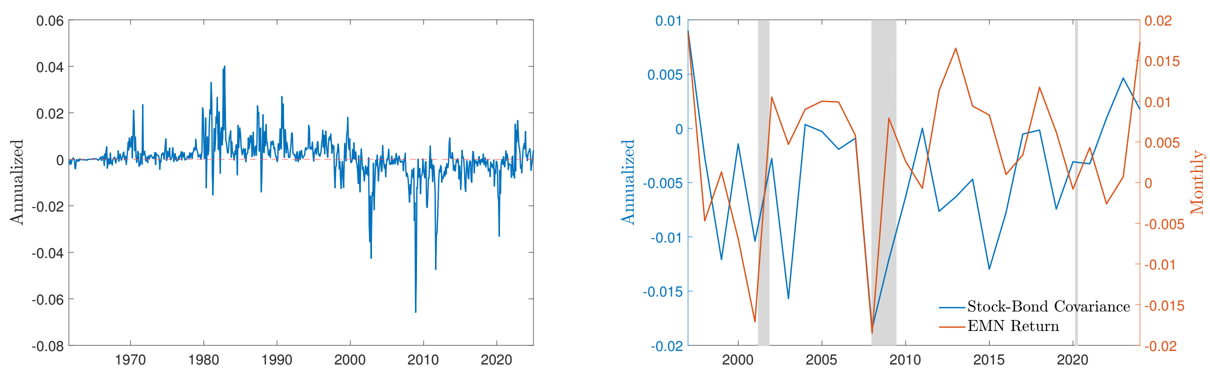
Notes: This figure shows the asset under management (AUM) of Equity Market Neutral Hedge Funds from the first quarter of 2000 to the first quarter of 2025. Source: BarclayHedge.

Figure. B.2. Barclay Equity Market Neutral Index



Notes: This figure shows the Barclay Equity Market Neutral Index, calculated and updated real-time as the cumulative equal-weighted average monthly (net) return of hedge funds in the Equity Market Neutral category. Source: BarclayHedge.

Figure. B.3. Stock-Bond Covariance and EMN Returns



(a) Monthly Stock-Bond Covariance

(b) Annual Covariance and EMN Returns

Notes: Panel (a) shows the monthly stock-bond covariance between nominal 10-year constant maturity bond returns and S&P 500 returns using a 30 trading-day rolling window. Panel (b) shows the annual trend in this covariance together with Barclay EMN returns.

The correlation between the annual trend in stock–bond covariance and Barclay EMN returns is 0.40. A regression of monthly EMN returns on monthly stock–bond covariance (including a constant) yields a significant coefficient of 0.23 with a t-statistic of 4.76. Alternatively, Using percentage changes in AUM as a proxy for returns at the quarterly frequency produces a coefficient of 3.93 with a t-statistic of 4.06. These results indicate that the performance of equity market-neutral strategies is linked to the hedging role of bonds. When stock–bond covariance rises, bonds become less reliable as hedges for equities, making market-neutral strategies appear more attractive as alternative diversification tools. The traditional 60/40 allocation rule gained popularity partly because stock–bond covariance had been negative since 2000. However, this pattern has recently reversed, with covariance turning positive after Covid. This shift suggests that market-neutral strategies may attract greater investor interest going forward.

B.2. In-Sample Factor Model Performance

Table B.1: In-Sample Model R^2

Models	Test Assets	Metrics	# Factors			
			1	3	6	9
FF	Individual Stocks	Total R^2	10.9	17.0	18.7	19.5
		Pred R^2	0.93	0.90	0.89	0.87
	Portfolios	Total R^2	89.4	94.1	95.2	95.4
		Pred R^2	3.42	3.42	3.42	3.42
PCA	Individual Stocks	Total R^2	4.5	10.3	13.1	14.3
		Pred R^2	0.01	0.28	0.34	0.34
	Portfolios	Total R^2	94.1	97.4	98.1	98.3
		Pred R^2	3.42	3.42	3.42	3.42
IPCA	Individual Stocks	Total R^2	12.8	15.1	15.8	16.0
		Pred R^2	0.81	0.79	0.78	0.77
	Portfolios	Total R^2	78.7	94.0	95.0	95.5
		Pred R^2	3.01	3.10	3.18	3.17
AE	Individual Stocks	Total R^2	12.4	13.3	13.5	13.5
		Pred R^2	1.12	1.02	1.09	1.16
	Portfolios	Total R^2	84.1	92.9	92.9	93.3
		Pred R^2	3.52	3.50	3.51	3.37

Notes: This table reports the in-sample total $R2$ and predictive R^2 in percentages (%) for FF, PCA, ICA, and AE with 1, 3, 6, and 9 factors.

Table B.2: Out-of-Sample Model R^2

Models	Test Assets	Metrics	# Factors			
			1	3	6	9
FF	Individual Stocks	Total R^2	6.5	7.4	2.7	0.3
		Pred R^2	-0.23	-0.21	-0.24	-0.25
	Portfolios	Total R^2	86.6	92.1	93.6	93.9
		Pred R^2	2.89	2.88	2.89	2.89
PCA	Individual Stocks	Total R^2	7.4	6.9	7.1	7.2
		Pred R^2	-1.09	-1.06	-1.07	-1.07
	Portfolios	Total R^2	92.3	96.2	97.1	97.5
		Pred R^2	2.88	2.88	2.88	2.88
IPCA	Individual Stocks	Total R^2	11.0	13.3	13.9	14.1
		Pred R^2	0.54	0.55	0.54	0.52
	Portfolios	Total R^2	73.6	93.0	93.8	94.4
		Pred R^2	3.15	3.22	3.23	3.27
AE	Individual Stocks	Total R^2	10.2	11.3	11.3	11.3
		Pred R^2	0.66	0.78	0.71	0.81
	Portfolios	Total R^2	79.5	91.7	93.3	92.5
		Pred R^2	3.23	3.02	3.27	3.13

Notes: This table reports the out-of-sample total $R2$ and predictive R^2 in percentages (%) for FF, PCA, ICA, and AE with 1, 3, 6, and 9 factors.

B.3. Out-of-Sample Factor Model Performance

B.4. Multi-Factor Volatility-Managed Portfolios

Table B.3: Investment Performance of Moreira-Muir Volatility-Managed Portfolios

Models	# Factors	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
		Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
FF	1	0.53	0.51	8.0	7.8	15.1	15.1	54.3	55.1	62.9	62.9
	3	0.46	0.38	8.1	6.8	17.9	17.9	57.2	59.2	56.9	56.7
	5	0.64	0.50	17.2	13.6	26.9	27.1	79.6	82.5	61.7	60.5
	6	0.63	0.51	15.7	12.7	24.9	25.1	73.3	76.4	64.0	62.6
PCA	1	0.46	0.44	7.1	6.8	15.3	15.4	51.1	51.6	60.0	60.0
	3	0.58	0.52	10.2	9.3	17.7	17.8	71.9	73.0	66.0	66.0
	5	0.59	0.49	12.2	10.3	20.8	20.9	63.5	65.8	60.7	59.8
	6	0.93	0.79	15.3	13.1	16.4	16.7	47.6	51.4	65.5	65.0
IPCA	1	0.42	-4.76	6.5	--	15.5	++	44.6	++	57.6	0.0
	3	1.82	-3.48	35.2	--	19.3	++	58.4	++	74.3	0.0
	5	3.99	-4.66	85.7	--	21.5	++	19.6	++	91.0	0.0
	6	3.88	-4.69	87.8	--	22.6	++	21.0	++	91.2	0.0
AE	1	0.42	-4.76	6.5	--	15.5	++	44.6	++	57.6	0.0
	3	1.82	-3.48	35.2	--	19.3	++	58.4	++	74.3	0.0
	5	3.99	-4.66	85.7	--	21.5	++	19.6	++	91.0	0.0
	6	3.88	-4.69	87.8	--	22.6	++	21.0	++	91.2	0.0

Notes: The table reports the performance of Moreira and Muir (2017) volatility-managed portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 5, and 6 factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%), and hit rates (%). I scale the mean-variance portfolio weights each month to target a 15% annualized volatility using historical estimates of mean returns and the covariance matrix. Performance is shown both before transaction costs ("gross") and after transaction costs ("net"). "++" indicates a number above 100 (%), and "--" indicates a number below -100 (%).

Table B.4: Investment Performance of DMU Volatility-Managed Portfolios

Models	# Factors	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
		Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
FF	1	0.53	0.51	8.0	7.8	15.1	15.1	54.3	55.1	62.9	62.9
	3	0.46	0.38	8.1	6.8	17.9	17.9	57.2	59.2	56.9	56.7
	5	0.64	0.50	17.2	13.6	26.9	27.1	79.6	82.5	61.7	60.5
	6	0.63	0.51	15.7	12.7	24.9	25.1	73.3	76.4	64.0	62.6
PCA	1	0.46	0.44	7.1	6.8	15.3	15.4	51.1	51.6	60.0	60.0
	3	0.58	0.52	10.2	9.3	17.7	17.8	71.9	73.0	66.0	66.0
	5	0.59	0.49	12.2	10.3	20.8	20.9	63.5	65.8	60.7	59.8
	6	0.93	0.79	15.3	13.1	16.4	16.7	47.6	51.4	65.5	65.0
IPCA	1	0.42	-4.76	6.5	--	15.5	++	44.6	++	57.6	0.0
	3	1.82	-3.48	35.2	--	19.3	++	58.4	++	74.3	0.0
	5	3.99	-4.66	85.7	--	21.5	++	19.6	++	91.0	0.0
	6	3.88	-4.69	87.8	--	22.6	++	21.0	++	91.2	0.0
AE	1	0.42	-4.76	6.5	--	15.5	++	44.6	++	57.6	0.0
	3	1.82	-3.48	35.2	--	19.3	++	58.4	++	74.3	0.0
	5	3.99	-4.66	85.7	--	21.5	++	19.6	++	91.0	0.0
	6	3.88	-4.69	87.8	--	22.6	++	21.0	++	91.2	0.0

Notes: The table reports the performance of DeMiguel et al. (2024) volatility-managed portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 5, and 6 factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%), and hit rates (%). I scale the mean-variance portfolio weights each month to target a 15% annualized volatility using historical estimates of mean returns and the covariance matrix. Performance is shown both before transaction costs ("gross") and after transaction costs ("net"). "++" indicates a number above 100 (%), and "--" indicates a number below -100 (%).

Multi-factor portfolios with vs without an intercept.

B.5. Prediction-Sorted Portfolios

Table B.5: Investment Performance of Prediction-Sorted Portfolios

Models	# Factors	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
		Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
FF	1	-0.54	-0.80	-9.3	-14.4	17.2	18.0	98.6	99.8	47.1	43.8
	3	-0.46	-0.75	-7.4	-12.6	16.1	16.8	97.6	99.6	45.2	41.9
	5	-0.47	-0.76	-7.4	-12.6	15.9	16.6	97.6	99.6	45.7	41.7
	6	-0.50	-0.81	-7.7	-12.9	15.3	16.0	97.7	99.6	44.8	41.2
PCA	1	-0.62	-0.86	-12.8	-18.9	20.7	21.9	99.7	100.0	48.3	43.3
	3	-0.62	-0.86	-12.8	-18.9	20.7	21.9	99.7	100.0	48.3	43.3
	5	-0.62	-0.86	-12.8	-18.9	20.7	21.9	99.7	100.0	48.3	43.3
	6	-0.62	-0.86	-12.8	-18.9	20.7	21.9	99.7	100.0	48.3	43.3
IPCA	1	0.97	0.70	18.2	12.3	18.9	17.5	44.8	52.5	64.3	59.5
	3	1.08	0.81	19.6	13.4	18.1	16.7	39.8	48.3	64.3	61.2
	5	1.04	0.76	18.5	12.4	17.8	16.3	34.6	43.8	65.0	59.5
	6	0.96	0.68	17.4	11.4	18.1	16.7	43.3	51.3	64.5	60.7
AE	1	1.30	0.73	15.8	8.4	12.2	11.4	25.0	31.9	66.4	60.7
	3	1.06	0.47	13.1	5.5	12.4	11.6	33.2	43.1	66.7	60.2
	5	1.16	0.60	15.0	7.4	13.0	12.3	29.1	36.9	64.8	57.4
	6	1.13	0.57	14.2	6.5	12.5	11.4	20.4	24.0	63.8	57.6

Notes: The table reports the performance of prediction-sorted portfolios for the FF, PCA, IPCA, and AE models with 1, 3, 5, and 6 factors. Reported statistics include annualized Sharpe ratios, mean returns (%), standard deviations (%), maximum drawdowns (%), and hit rates (%). Performance is shown both before transaction costs ("gross") and after transaction costs ("net").

B.6. Model-Free Portfolios

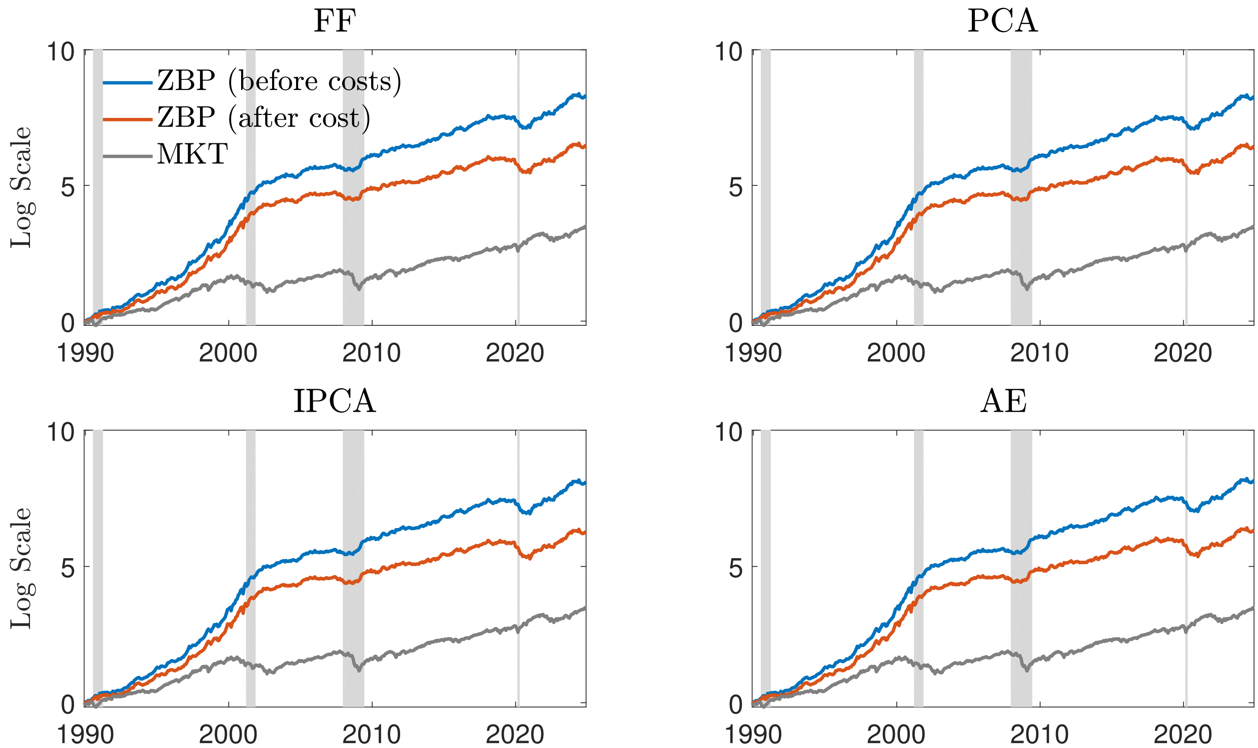
Table B.6: Investment Performance of Other Benchmark Portfolios

Portfolio Rules	Sharpe Ratio		Mean (%)		STD (%)		MDD (%)		Hit Rate (%)	
	Gross	Net	Gross	Net	Gross	Net	Gross	Net	Gross	Net
MKT	0.53	0.52	11.1	11.0	15.5	15.5	51.6	51.9	64.5	64.5
MVP	0.79	0.39	9.6	4.9	12.1	12.5	36.0	40.3	61.0	56.7
GMV	0.69	0.49	7.5	5.5	10.9	11.3	26.8	30.3	61.2	58.8
EW	0.48	0.46	8.5	8.1	17.7	17.7	56.7	57.2	60.0	60.0
RP	0.50	0.48	8.6	8.3	17.2	17.2	56.3	56.7	61.0	60.2
KWZ	1.31	1.02	27.2	21.7	20.8	21.2	49.0	53.6	68.6	65.0

Notes: The table reports the performance of conventional model-free portfolios, including the market portfolio (MKT), mean-variance portfolio (MVP), global minimum-variance portfolio (GMV), equal-weighted portfolio (EW), risk-parity portfolio (RP), Kan et al. (2022) optimal portfolio (KWZ). I scale the MVP and KWZ portfolio weights each month to target a 15% annualized volatility using historical estimates of mean returns and the covariance matrix. Investment performance is reported both before transaction costs ("gross") and after transaction costs ("net").

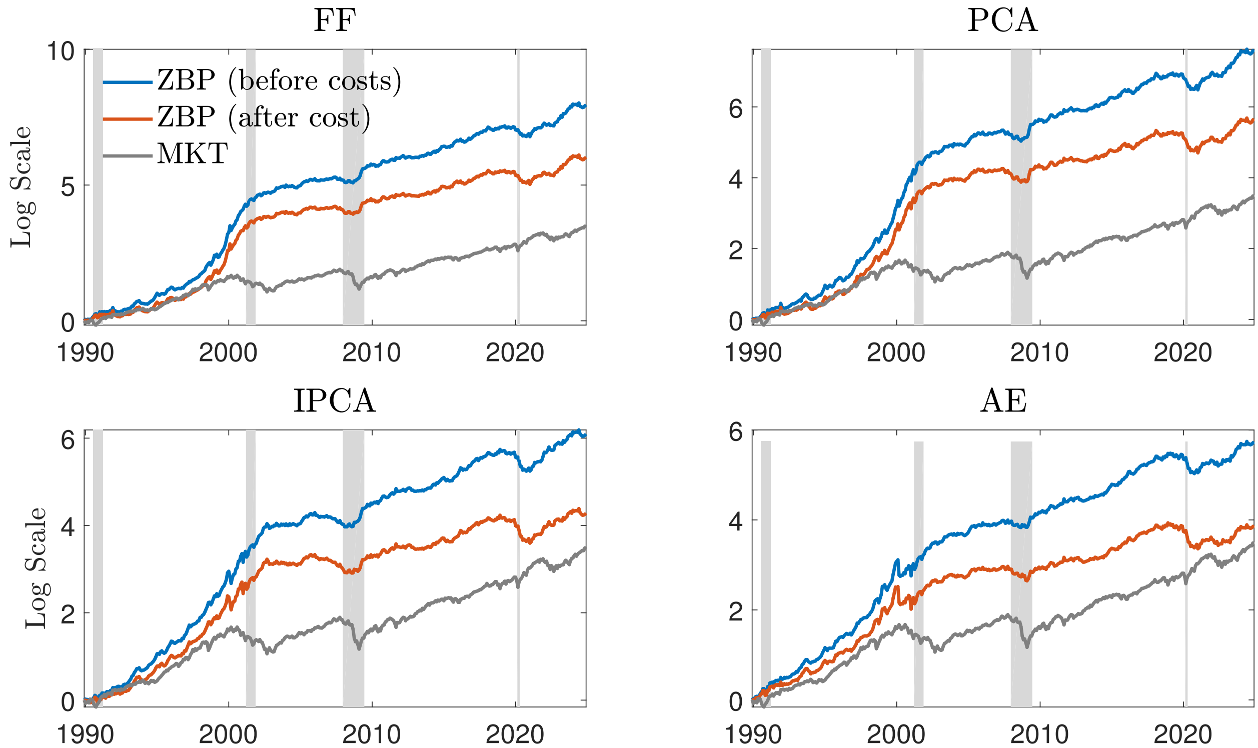
B.7. Price Path of Factor-Neutral Portfolios

Figure. B.4. Price Path of Factor-Neutral Portfolios (Single Factor)



Notes: This figure shows the logarithmic price paths (cumulative returns) of factor-neutral portfolios before and after transaction costs, alongside the market portfolio. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with a single factors.

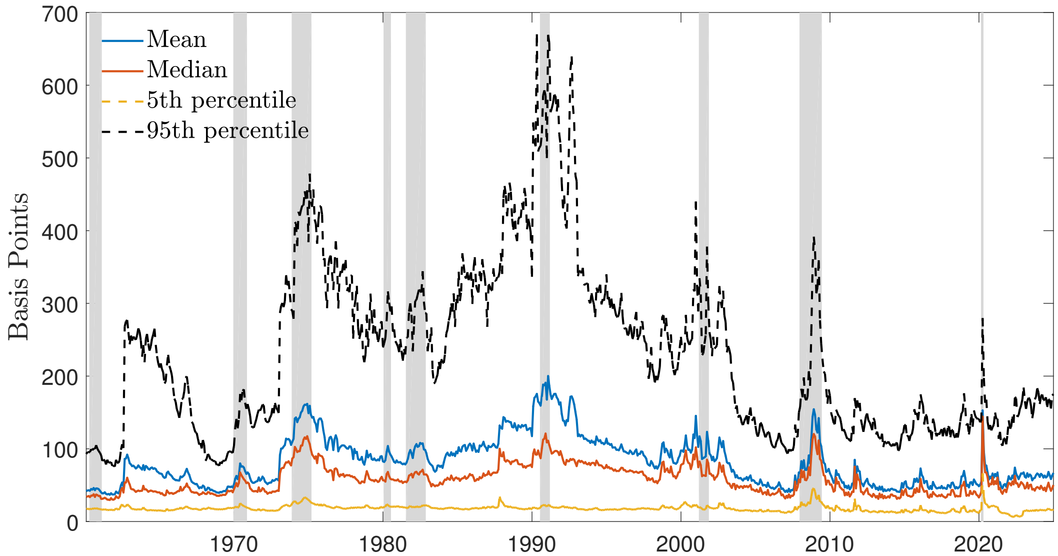
Figure. B.5. Price Path of Factor-Neutral Portfolios (Six Factors)



Notes: This figure shows the logarithmic price paths (cumulative returns) of factor-neutral portfolios before and after transaction costs, alongside the market portfolio. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with six factors.

B.8. Individual-Stock Level Transaction Costs

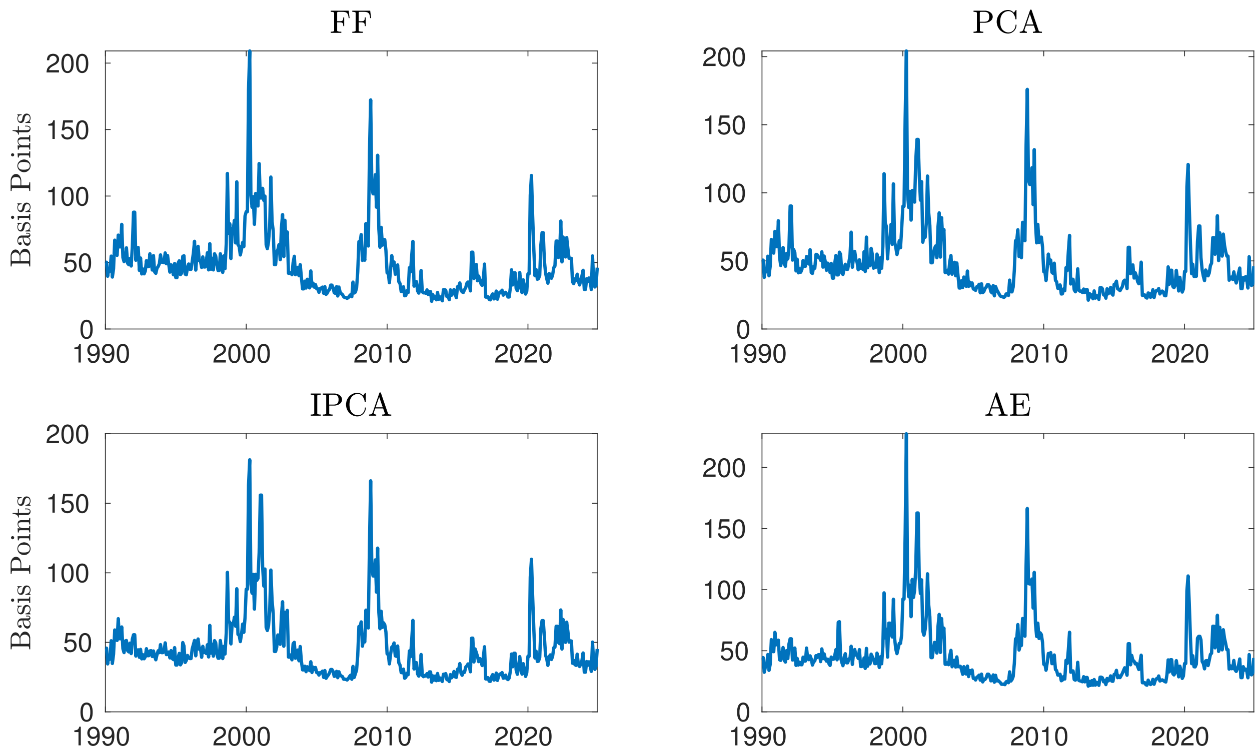
Figure. B.6. Individual-Stock Level Transaction Costs



Notes: This figure shows the time variation of the mean, median, 5th percentile, and 95th percentile of individual transaction costs from Jan 1960 to Dec 2024, measured using the average low-frequency effective spreads described in Chen and Velikov (2023).

B.9. Zero-Beta Portfolio Transaction Costs

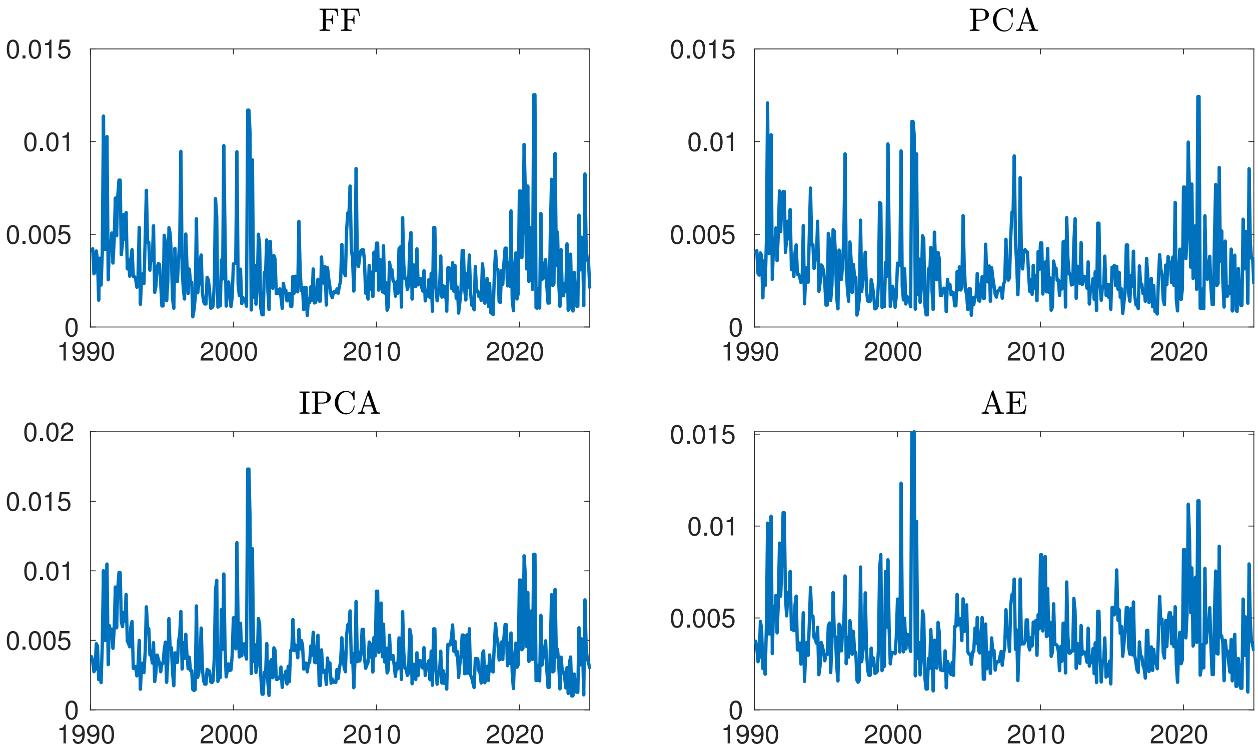
Figure. B.7. Factor-Neutral Portfolio Transaction Costs (Six Factors)



Notes: This figure shows the transaction costs of the optimal factor-neutral portfolios from January 1990 to December 2024, measured using the average low-frequency effective spreads in Chen and Velikov (2023). The models considered are FF, PCA, IPCA, and AE with six factors.

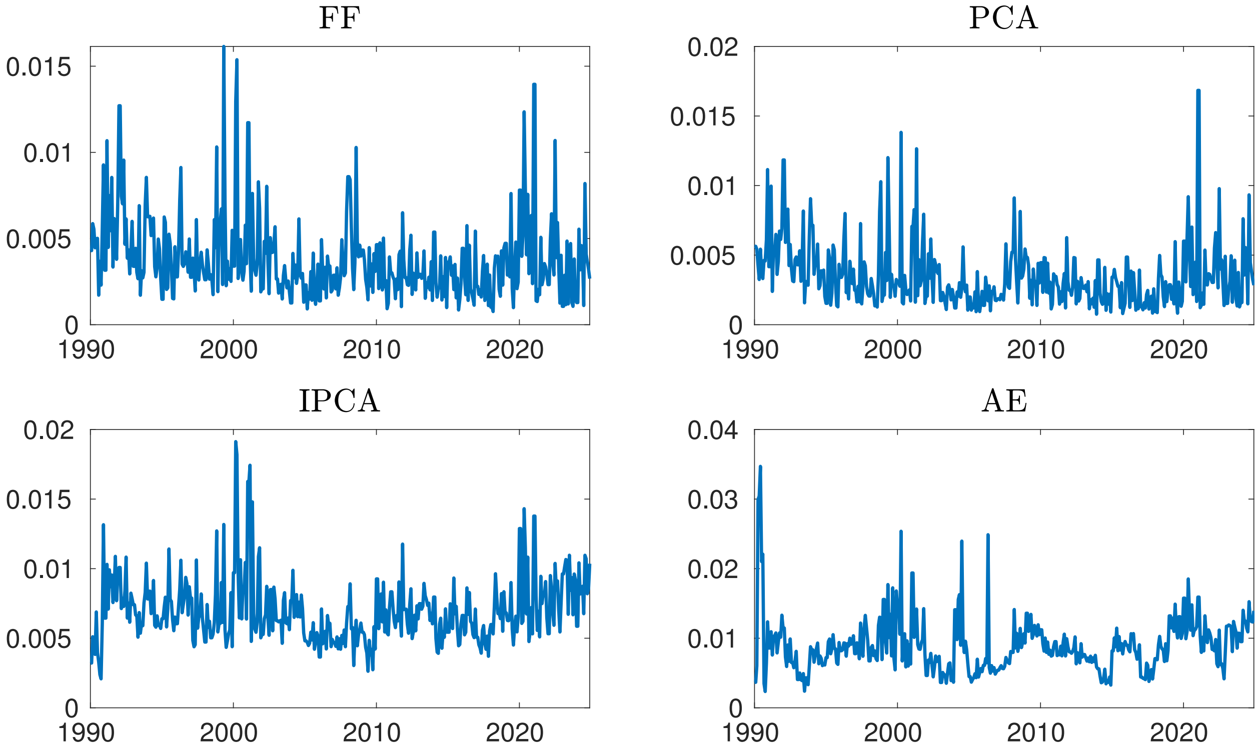
B.10. Median Absolute Factor-Neutral Portfolio Weight Changes

Figure. B.8. Median Absolute Factor-Neutral Portfolio Weight Changes (Single Factors)



Notes: This figure shows the median absolute weight changes of optimal zero-beta portfolios across the 272 characteristic-sorted portfolios from January 1990 to December 2024. The models considered are FF, PCA, IPCA, and AE with a single factors.

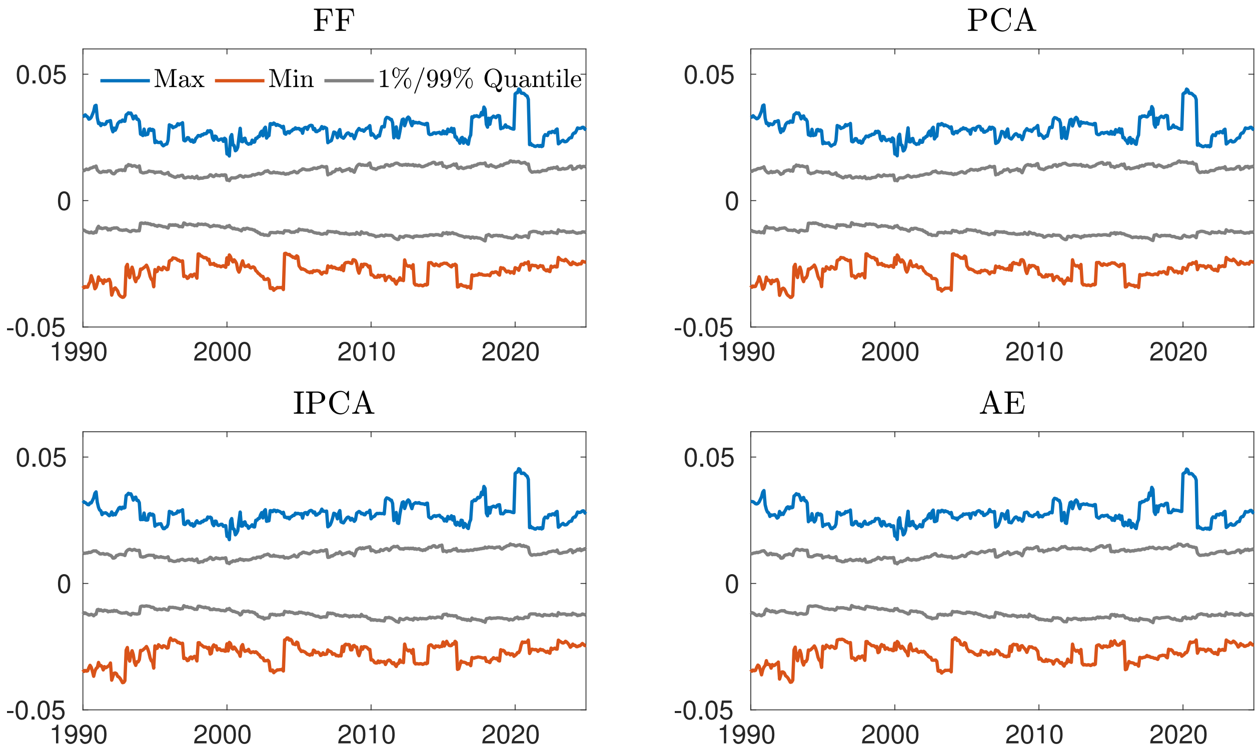
Figure. B.9. Median Absolute Factor-Neutral Portfolio Weight Changes (Six Factors)



Notes: This figure shows the median absolute weight changes of optimal factor-neutral portfolios across the 272 characteristic-sorted portfolios from January 1990 to December 2024. The models considered are FF, PCA, IPCA, and AE with six factors.

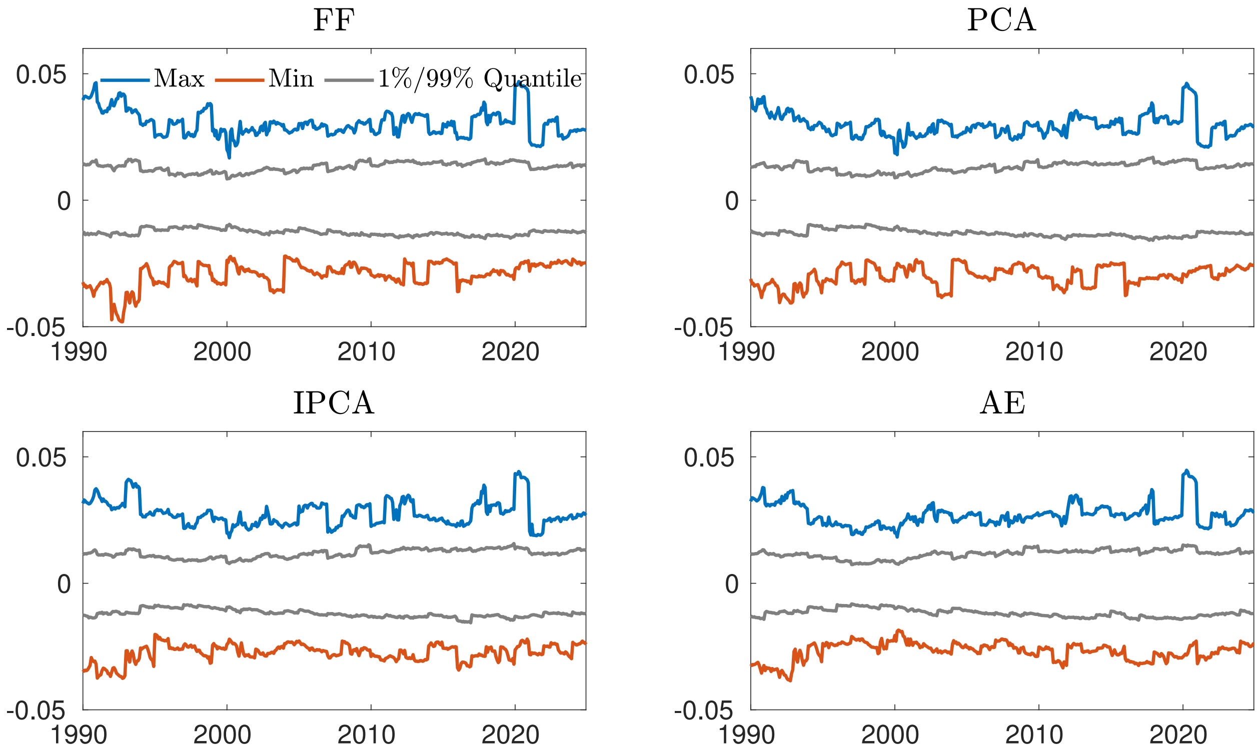
B.11. Factor-Neutral Portfolio Weights on Individual Stocks

Figure. B.10. Factor-Neutral Portfolio Weights on Individual Stocks (Single Factors)



Notes: This figure shows the maximum, minimum, 1% quantile, and 99% quantile of the zero-beta portfolio weights on individual stocks from January 1990 to December 2024. The models considered are FF, PCA, IPCA, and AE with a single factors.

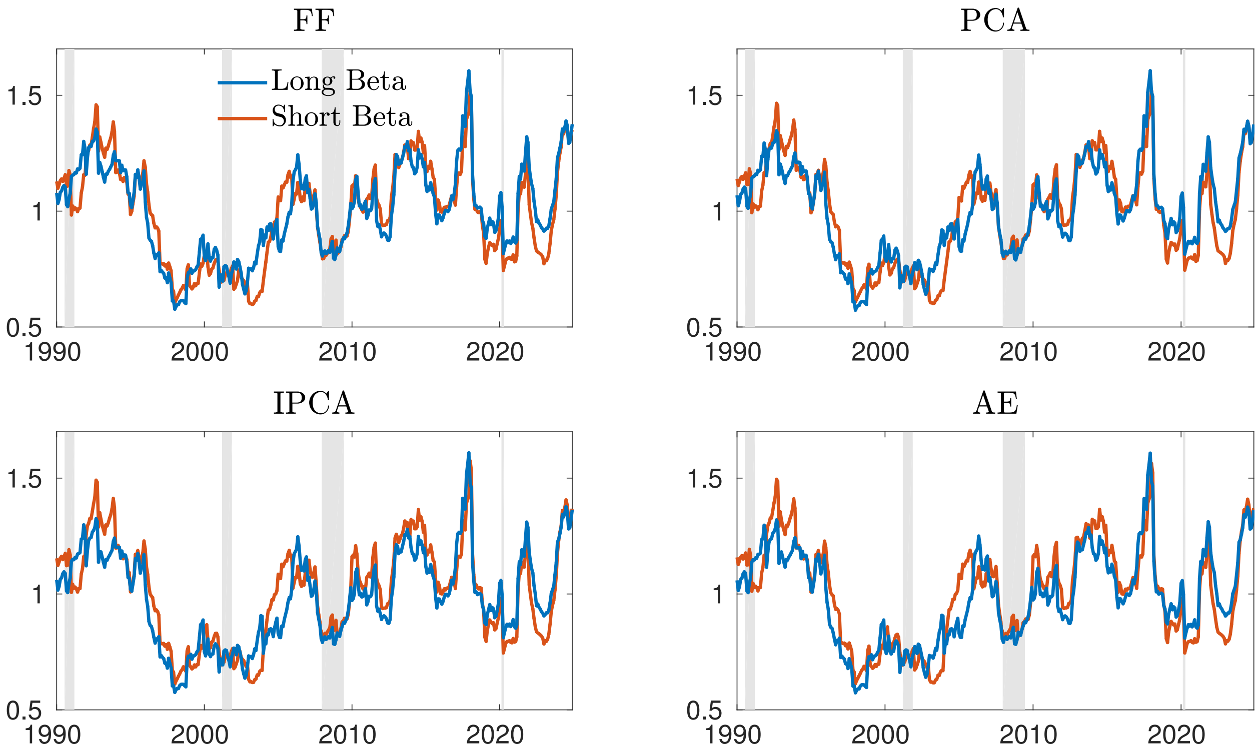
Figure. B.11. Zero-Beta Portfolio Weights on Individual Stocks (Six Factors)



Notes: This figure shows the maximum, minimum, 1% quantile, and 99% quantile of the factor-neutral portfolio weights on individual stocks from January 1990 to December 2024. The models considered are FF, PCA, IPCA, and AE with six factors.

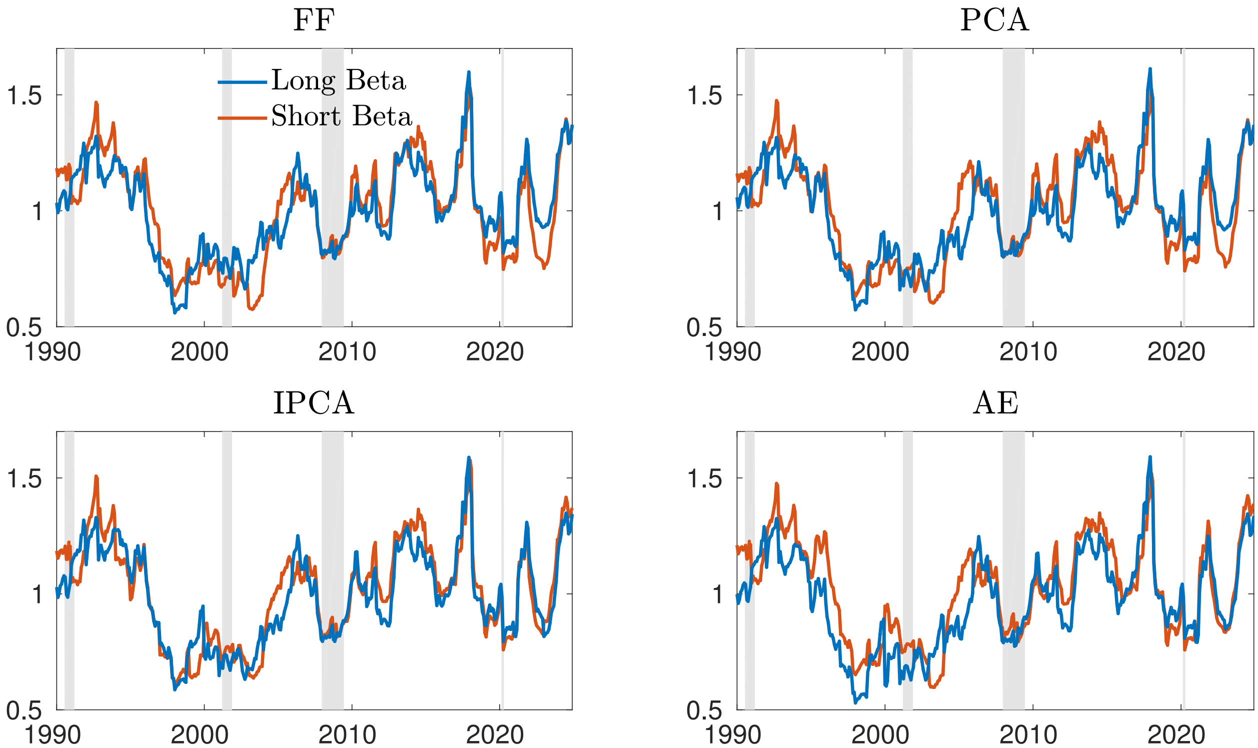
B.12. Market Betas of Long vs Short Positions

Figure. B.12. Market Betas of Long vs Short Positions (Single Factor)



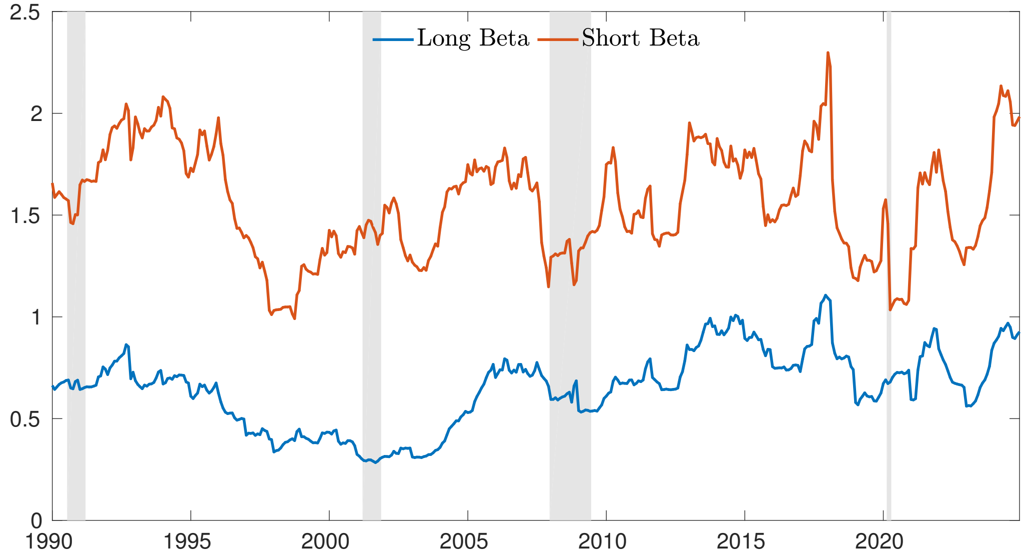
Notes: This figure shows the weighted average market betas of long and short positions for the optimal factor-neutral portfolios. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with a single factors.

Figure. B.13. Market Betas of Long vs Short Positions (Six Factor)



Notes: This figure shows the weighted average market betas of long and short positions for the optimal factor-neutral portfolios. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with three factors.

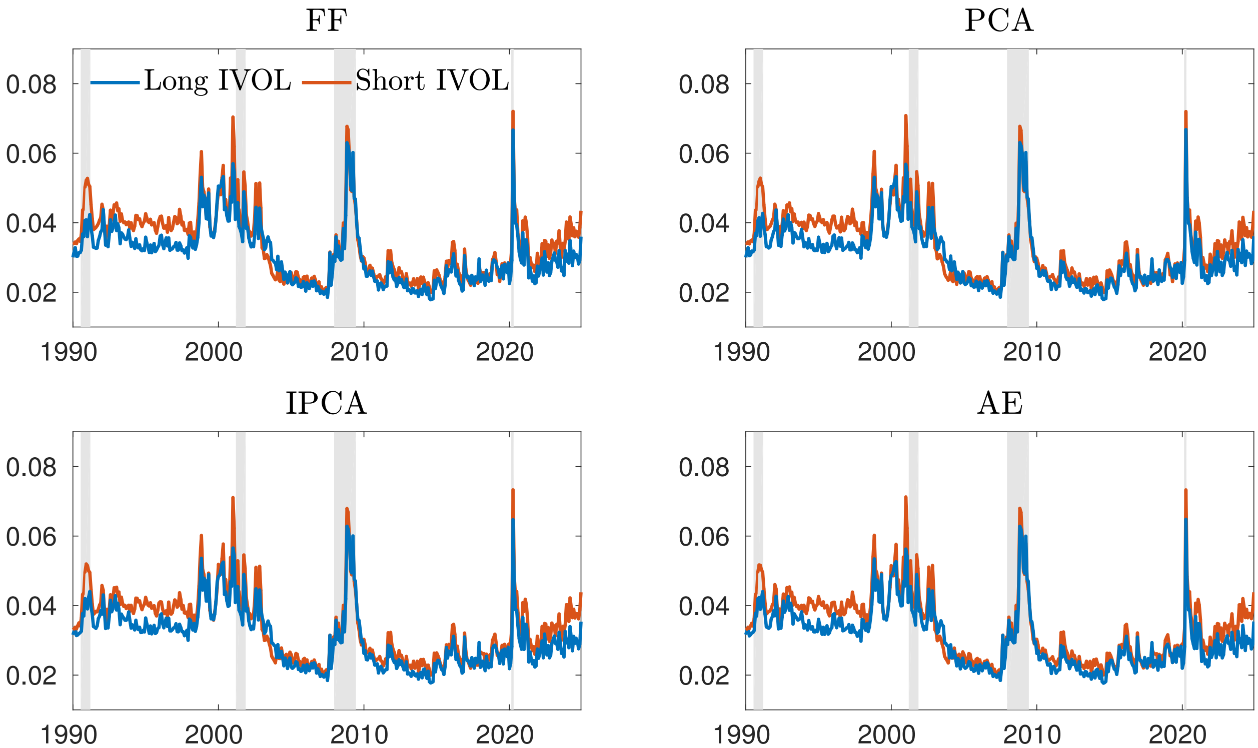
Figure. B.14. Market Betas of Long vs Short Positions (BAB)



Notes: This figure shows the weighted average market betas of long and short positions for the BAB factor portfolio. Gray-shaded areas denote NBER recessions.

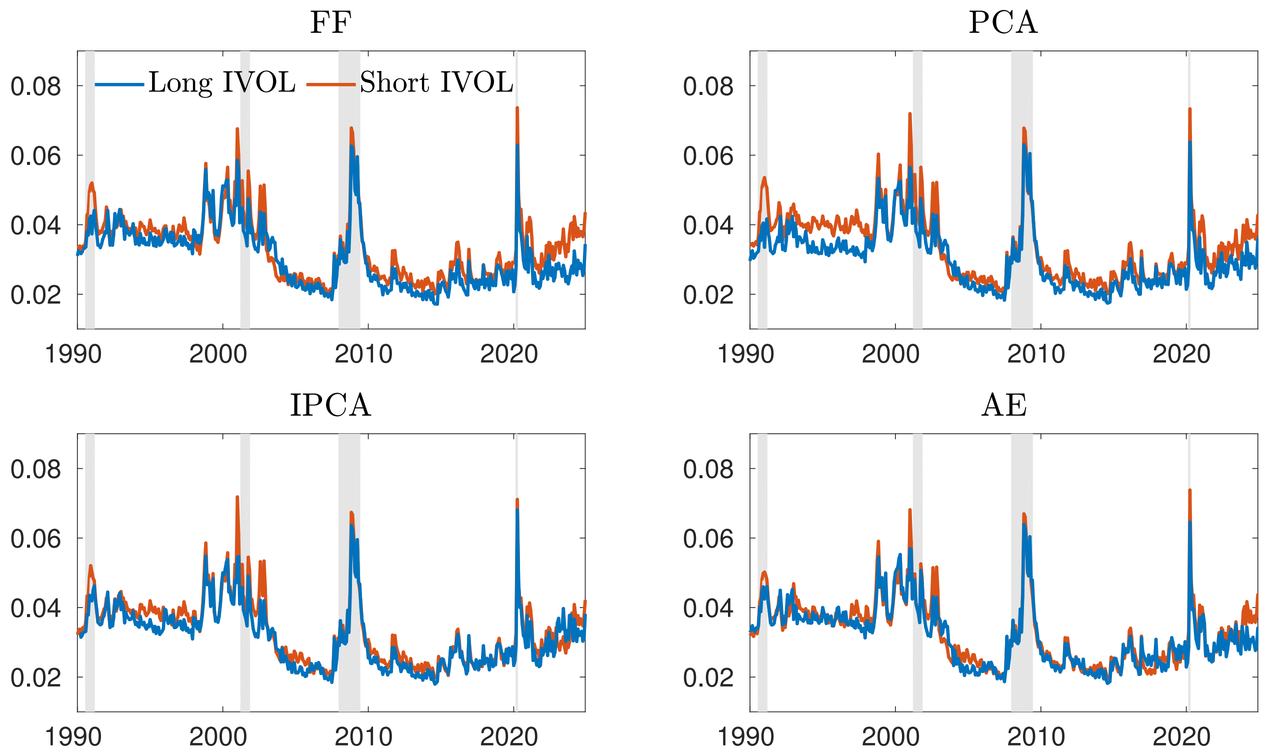
B.13. Idiosyncratic Risks of Long vs Short Positions

Figure. B.15. Idiosyncratic Risks of Long vs Short Positions (Single Factor)



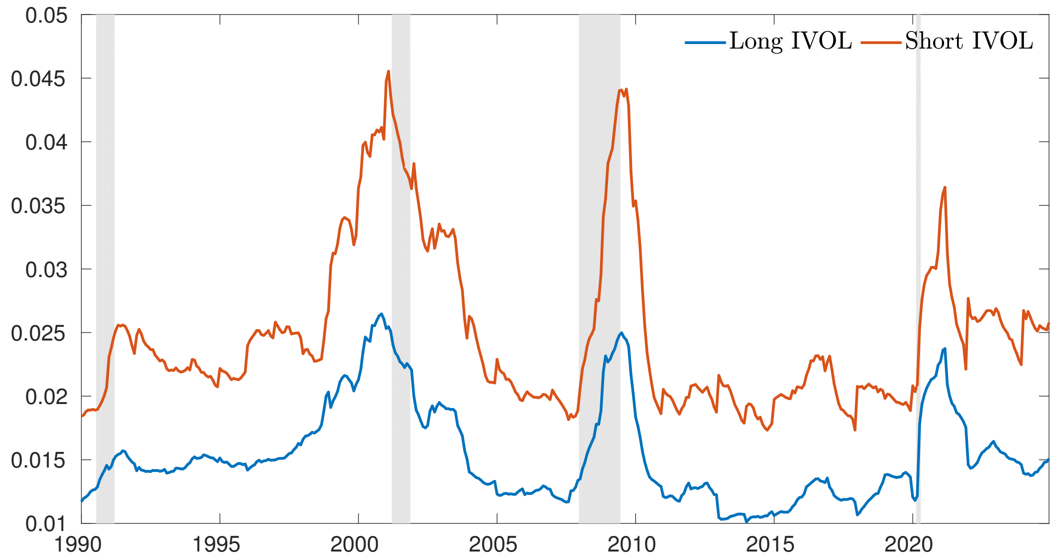
Notes: This figure shows the weighted average idiosyncratic risks of long and short positions for the optimal factor-neutral portfolios. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with a single factors.

Figure. B.16. Idiosyncratic Risks of Long vs Short Positions (Six Factor)



Notes: This figure shows the weighted average idiosyncratic risks of long and short positions for the optimal factor-neutral portfolios. Gray-shaded areas denote NBER recessions. The models considered are FF, PCA, IPCA, and AE with three factors.

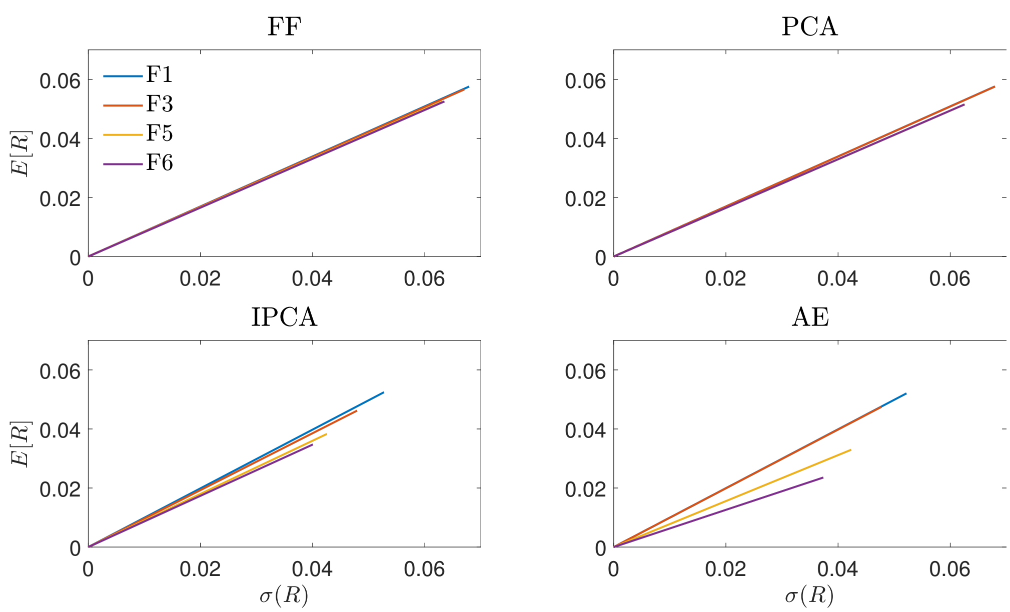
Figure. B.17. Idiosyncratic Risks of Long vs Short Positions (BAB)



Notes: This figure shows the weighted average idiosyncratic risks of long and short positions for the BAB factor portfolio. Gray-shaded areas denote NBER recessions.

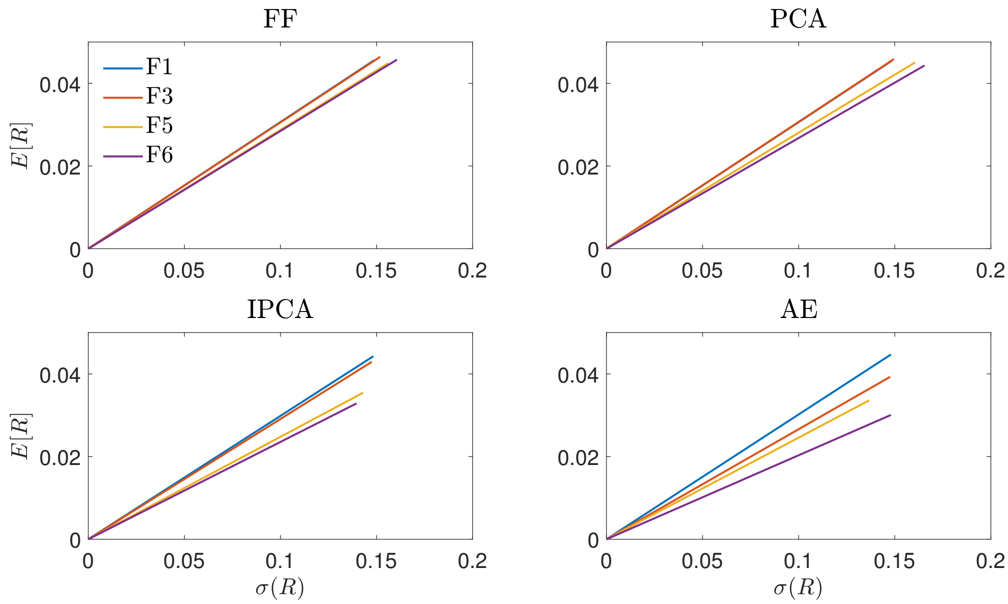
B.14. Visualization of Model Comparison

Figure. B.18. Zero-Investment, Zero-Beta Frontiers (In-Sample)



Notes: This figure shows the in-sample zero-investment, zero-beta frontiers for the FF, PCA, IPCA, and AE models with 1, 3, 5, and 6 factors. The slope of each zero-beta frontier corresponds to the in-sample maximum Sharpe ratio attainable by zero-beta portfolios.

Figure. B.19. Zero-Investment, Zero-Beta Frontiers (Out-of-Sample)



Notes: This figure shows the out-of-sample zero-investment, zero-beta frontiers for the FF, PCA, IPCA, and AE models with 1, 3, 5, and 6 factors. The slope of each zero-beta frontier corresponds to the out-of-sample maximum Sharpe ratio attainable by zero-beta portfolios.