Calm Your Portfolio: The Importance of Disciplining Intelligent but Fickle Forecasts in Portfolio Optimization

(Preliminary: please do not circulate.)

This version: October 2022

Pedro Barroso*, Konark Saxena†, Haoxu Wang‡

Abstract

The asset-pricing literature has proposed several return forecasting methods ranging from conventional Fama-MacBeth regressions to advanced methods based on machine learning. The literature on estimation of covariance matrices has proposed many sophisticated estimators based on factor models, shrinkage methods or GARCH. We comprehensively examine whether advances in the two strands of literature can jointly improve the out-of-sample performance of mean-variance efficient (MVE) portfolios. Focusing on the 500 largest stocks, we find that MVE portfolios formed using *improved* inputs do not compare favorably to the passive strategy after accounting for transaction costs. However, their after-cost performance can be substantially improved through risk targeting coupled with transaction cost management. Notably, the transaction-cost-managed risk-targeted portfolios constructed using return forecasts from Fama-MacBeth regressions with the (daily) sample or Galton covariance matrix can attain net Sharpe ratios greater than one and *significantly* outperform the passive counterpart.

Keywords: Markowitz optimization; Fama-MacBeth regressions; machine learning; sample covariances; factor models; shrinkage; GARCH; volatility targeting; transaction cost management;

JEL Classification: G11; G12; G17

*Universidade Católica Portuguesa, Católica Lisbon School of Business and Economics. E-mail: pedro.barroso@ucp.pt

†University of New South Wales. E-mail: k.saxena@unsw.edu.au

‡University of New South Wales. E-mail: haoxu.wang1@unsw.edu.au

1 Introduction

The Nobel prize-winning work of Markowitz (1952) shows how to construct a mean-variance efficient (MVE) portfolio given a vector of expected returns and the covariance matrix of assets. Though theoretically sound, the Markowitz model struggles in out-of-sample (OOS) contexts (Jobson and Korkie 1980; Michaud 1989; Okhrin and Schmid 2006; Kan and Smith 2008). This OOS inefficiency results from estimation error both in the sample mean and in the sample covariance matrix.

It is well established that the sample mean is a noisy estimator for expected returns; moreover, the Markowitz optimizer is more sensitive to errors in means than those in variances or covariances (Best and Grauer 1991; Chopra and Ziemba 1993). Though computationally simple, unbiased and intuitively attractive, the sample covariance matrix is subject to the curse of dimensionality and thus not always a desirable estimator for the covariance matrix, especially when the number of observations in an estimation sample is comparable to or greater than the number of assets. Furthermore, the inverse of the sample covariance matrix is a biased estimator of the true inverse covariance matrix (Bai and Shi 2011).

The starting motivation of our paper is that the asset-pricing literature and the literature on covariance matrix estimation have proposed a variety of sophisticated estimators in place of sample estimates (i.e., historical means and covariances); however, the advances in the two strands of literature are seldom *jointly* studied in portfolio optimization contexts.[1]

For the past 30 years, numerous stock characteristics with return forecasting power have been documented in the asset-pricing literature. Prior studies tend to exploit stock characteristics in portfolio management by constructing characteristic-sorted portfolios or factor portfolios in line with Fama and French (1993). However, the portfolio sorting approach is incompatible with controlling simultaneously for a large number of characteristics. Some studies propose estimating expected returns from a parsimonious factor model consisting of only the market factor or a small set of factors able to span many characteristics.[2] But this method might omit important asset-pricing factors and neglect potential nonlinear predictive power of characteristics for returns as well as interactions among characteristics. Alternatively, several

[1]As an example, Ledoit and Wolf (2017) explain that "It may be possible to cumulate the improvements of the two strands of literature by combining our method for the estimation of the covariance matrix with some other method for the reduction in the estimation risk of the vector of expected returns. This topic is an interesting avenue for future research but it lies outside the scope of the present paper."

[2]E.g., Sharpe (1963) proposes the single-index model to estimate expected returns. Fama and French (2020) examine the performance of their 5-factor and 6-factor models (Fama and French 2015, 2018) in predicting expected returns. Other recent leading models include the q-factor model (Hou, Xue, and Zhang 2015), the mispricing factors (Stambaugh and Yuan 2017), and the behavioral factors of Daniel, Hirshleifer, and Sun (2020).

articles use cross-sectional Fama and MacBeth (1973, Fama-MacBeth) regressions to conduct the expected return estimation (Haugen and Baker 1996; Lewellen 2015; Green, Hand, and Zhang 2017; Bessembinder, Cooper, and Zhang 2018; Fama and French 2020). Compared to the method based on parsimonious factor models, this approach can accommodate a large number of characteristics. Moreover, Fama and French (2020) show that for the same set of characteristics, it compares favorably to the time-series factor-model counterpart in estimating expected returns. Nonetheless, it still fails to account for potential nonlinearities or characteristic interactions.

A recent literature proposes addressing the concerns about conventional methods using machine learning techniques. Kelly, Pruitt, and Su (2019, KPS) and Gu, Kelly, and Xiu (2021, GKX2021) propose conditional latent factor pricing models. Those use observable characteristics to determine exposure to latent risk factors, both of which can potentially accommodate all existing stock characteristics. The former adopts a linear specification to model the characteristic-beta relationship and solve the problem via alternating least squares, while the latter uses a neural network model to account for both nonlinear characteristic-beta relationships and characteristic interactions in beta estimation. Gu, Kelly, and Xiu (2020, GKX2020) and Freyberger, Neuhierl, and Weber (2020, FNW) directly model expected risk premiums of individual stocks as a function of a large number of covariates and advocate the use of complex functional forms such as the nonparametric additive model (FNW) and neural networks (GKX2020) in return forecasting. To justify the economic contribution of a return forecasting method, this strand of literature usually examines the profitability of a long-short portfolio formed on estimated expected returns. Despite the fact that the performance of expected-return-sorted portfolios are often evaluated in terms of their mean-variance efficiency, the expected portfolio risk, which depends on the covariance matrix of assets returns, is not taken into consideration when forming such portfolios.

The literature on covariance matrix estimation has proposed several sophisticated estimators to address the shortcomings of the sample covariance matrix. However, when assessing the quality of a proposed estimator in portfolio optimization contexts, relevant studies typically refrain from estimating expected returns and instead focus on the realized risk of the global minimum variance (GMV) portfolio based on the estimator (e.g., Chan, Karceski, and Lakonishok 1999; Ledoit and Wolf 2017; Hautsch and Voigt 2019). After all, the estimated covariance matrix is the only input of the GMV optimization, so ex post risk of GMV portfolios is a plausible direct metric of the success of the estimation.[3]

[3]Two exceptions are Engle, Ledoit, and Wolf (2019, ELW) and De Nard, Ledoit, and Wolf (2021, DLW) who exploit their proposed estimators in the *full* Markowitz optimization. But they just use the momentum signal to form return forecasts.

Our empirical analysis covers ten estimators for expected returns. The first two are the sample mean and the recently proposed Galton corrected sample mean developed by Barroso and Saxena (2022). The rest include four methods based on Fama-MacBeth regressions and four estimators employing machine leaning techniques, all of which incorporate stock characteristics into estimation.

Our Fama-MacBeth regressions use the large set of 62 characteristics in FNW. We consider two investment universes, one consists of the stocks with market capitalization larger than the NYSE median[4] and the other of all stocks. Estimations with all stocks are with weighted least squares with market equity as weights. Focusing on large stocks only, or weighting errors by market cap, ensures our results are not driven by predictability in small cap stocks that collectively account for only a very small part of the market. Using a large predictor set could result in overfitting of the Fama-MacBeth models. To mitigate this concern, along the lines of how Barroso and Saxena (2022) shrink sample means, we propose two shrinkage Fama-MacBeth estimators, which use historical OOS forecast errors of plain Fama-MacBeth regressions to correct estimates.

The four methods related to machine learning techniques include the nonparametric additive model coupled with the adaptive group *lasso* (FNW), instrumental principal component analysis (IPCA) developed by KPS, "feed-forward" neural networks advocated by GKX2020 and conditional autoencoder model proposed by GKX2021. We confirm that all nine alternative estimators deliver superior return forecasting performance for stocks in our investment universe compared to the sample mean or a naive forecast of zero.

Our main empirical applications cover four covariance matrix estimators: the covariance matrix implied by the single-index model of Sharpe (1963), the daily sample covariance matrix, the Galton covariance matrix of Barroso and Saxena (2022), and the dynamic nonlinear shrinkage estimator of ELW. The first two estimators are commonly used in the literature as well as in practice. The other two estimators, which are newly proposed in the literature, represent linear shrinkage methods and dynamic shrinkage methods, respectively. The covariance matrix implied by the single-index model is always nonsingular regardless of dimension sizes. Singularity and quality of the sample covariance matrix depend on the concentration ratio (i.e., ratio of the number of assets over the number of historical observations). Since we use daily returns for a 500-stock universe and an estimation window of 60 months, the ratio is around 0.4, which is much less than one. Thus the sample covariance matrix is not as unfavorable as in studies computing the sample covariance matrix

[4] Lewellen (2015) define large-cap stocks in the same way.

from monthly returns.[5]

Comparing eight prominent estimators in the literature in terms of OOS standard deviation of the GMV portfolio, the daily sample covariance matrix performs better than the single index model but worse than the LW2003 method.[6] The ELW and Galton estimators are the two top performers. All methods work to reduce risk versus the value-weighted portfolio.

Given these improved optimization inputs, we first examine whether they can result in abnormal OOS performance of MVE portfolios relative to the value-weighted portfolio for an investment universe consisting of top-500 stocks by market equity over the 1987:01-2020:12 period. We find that of 40 tangency portfolios, only ten outperform the passive strategy in a statistically significant way, seven of those use our shrinkage Fama-MacBeth estimators. This shows that it is possible to improve the OOS performance of MVE portfolios through incorporating stock characteristics into optimization. However, compared to shrinking ordinary Fama-MacBeth estimates, using complex functional forms along with machine learning techniques brings limited benefits to mean-variance optimization, at least for large-cap stocks.

Furthermore, owing to the "error-maximization" property of the mean-variance optimizer and instability of the denominator of the tangency portfolio formula, some tangency portfolios and all complete portfolios tend to take extreme positions and thus deliver unrealistic OOS standard deviation. In addition, these portfolios occasionally go bankrupt and require a high degree of turnover. Motivated by Kirby and Ostdiek (2012) who target the estimated expected returns of a benchmark portfolio, we target the estimated variance of the value-weighted portfolio when constructing MVE portfolios. That is, we employ an upper (expected) variance bound as a constraint in the quadratic-utility-maximization problem.[7] The resulting Risk-Targeted (RT) portfolios generally compare favorably to the unconstrained counterparts. Notably, all eight RT portfolios constructed using combinations of characteristic-based expected return estimators with the Galton covariance matrix

[5] That is, for the same 60-month estimation window, the monthly sample covariance matrix is not invertible for a 500-stock universe and is indeed of large dimension for a 50-stock universe.

[6] Using a 500-stock universe, Jagannathan and Ma (2003) document that the GMV portfolio built using the daily sample covariance matrix outperforms that constructed using the covariance matrix implied by the single-index model or the three-factor model of Fama and French (1993) and performs as well as that constructed from the linear shrinkage estimator of Ledoit and Wolf (2003, LW2003). However, for the estimators proposed after the publication of Jagannathan and Ma (2003), their corresponding GMV portfolios all yield lower ex post standard deviation than the GMV portfolio based on the daily sample covariance matrix.

[7] Jagannathan and Ma (2003) argue that despite introducing specification error, a constraint on portfolio weights can reduce sampling error even if the constraint is improper in the population. Apart from weight constraints, they also emphasize the importance of using better return forecasts. They document that tangency portfolios built using sample means in combination with weight constraints or an improved covariance matrix estimator underperform the equal-weighted portfolios in terms of the before-cost Sharpe ratio and suggest that future research should incorporate information beyond historical means into the estimation of expected returns.

deliver reasonable OOS standard deviation and never go bankrupt during the OOS period; furthermore, they can attain impressive Sharpe ratios ranging from 1.10 to 2.04 and significantly outperform the passive counterpart.

Apart from estimation risk, investors face challenges of overcoming transaction costs. The extant literature raises a concern that the presence of trading costs can lead to a drastic performance deterioration for active portfolios. Examining tangency portfolios estimated from various improved inputs, Barroso and Saxena (2022) document that except for tangency portfolios based on their Galton estimates, none can beat the value-weighted and equal-weighted portfolios in terms of the Sharpe ratio net of transaction costs of 10 bps. Avramov, Cheng, and Metzker (2022) examine net-of-costs profitability of the recently documented long-short machine-learning portfolios and point out that "accounting for reasonable transaction costs would make it difficult for most machine learning signals to leave alpha on the table".[8] Ledoit and Wolf (2017, LW2017) demonstrate that for the investment universe consisting of the 500 largest stocks, no GMV portfolio can achieve a higher Sharpe ratio than the $1/N$ portfolio when the bid-ask spread is greater than or equal to 5 bps regardless of GMV estimators employed. All of these studies highlight the importance of managing transaction costs ex ante.

For simplicity, previous portfolio optimization studies often assume a flat cost that is constant over time and across stocks. To better measure how transaction costs erode the benefits of a strategy, we build on the comparative results of Abdi and Ranaldo (2017) and use instead the closing quoted spread of Chung and Zhang (2014) and the estimator of Abdi and Ranaldo (2017).[9] This endows our portfolio optimization tests with realistic estimates incorporating well-known cross-sectional and time series heterogeneity in trading costs.

In this realistic setting, all mean-variance optimized portfolios that do not accounting for transaction costs ex ante, including the GMV, tangency, complete and RT portfolios, fail to outperform the passive strategy in a statistically significant way in terms of net-of-costs Sharpe ratios. This motivates us to investigate whether (ex ante) transaction cost management could improve after-cost performance of MVE

[8] In addition, FNW show that decile-spread portfolios formed on expected returns estimated using a nonparametric additive model regularized by the adaptive group *lasso* or a linear model regularized by the adaptive *lasso*, in an universe excluding micro-cap stocks, produce negative net-of-costs Sharpe ratios.

[9] DeMiguel, Garlappi, and Uppal (2009), Kirby and Ostdiek (2012), Hautsch and Voigt (2019), and Barroso and Saxena (2022), for example, assume flat costs. Abdi and Ranaldo (2017) find that the two estimators in our study outperform in terms of correlation with effective spreads estimated from the Trade and Quote (TAQ) data. Brandt, Santa-Clara, and Valkanov (2009), Avramov, Cheng, and Metzker (2022) and Freyberger, Neuhierl, and Weber (2020) model transaction costs as a linear function of relative market capitalization and a time trend. But Novy-Marx and Velikov (2016) show that fails to account for a nonlinear relation between size and transaction costs and also for idiosyncratic volatility, a major determinant of costs.

portfolios built using sophisticated estimators. We incorporate the expected transaction costs into the optimization problem of the RT portfolio. That is, we use a turnover penalization and an upper risk bound in the quadratic utility maximization problem. We show that this problem is a convex quadratically constrained quadratic program, which can be solved by the interior-point method with great efficiency. More importantly, its solution is always globally optimal.

The turnover penalization acts similarly to the *lasso* penalty of Tibshirani (1996) on weights, and effective one-way spreads of each stock observable on the formation date are used to govern the regularization intensity. We keep the risk bound for two reasons. First, the imposition of the bound can render the optimized portfolios comparable to the passive strategy in terms of OOS standard deviation. Second, it can refrain the mean-variance optimizer from producing extreme weights when the market is liquid, since during that period regularization effects of turnover regularization on the optimizer are considerably weaker.

Calming the portfolio with turnover penalties makes a dramatic difference. All 32 transaction-cost-managed RT (TCM-RT) portfolios that exploit stock characteristics produce higher Sharpe ratios than the value-weighted strategy. In particular, those constructed using combinations of a Fama-MacBeth-based estimator with the daily sample covariance matrix can achieve net-of-costs Sharpe ratios between 1.10 and 1.20 that significantly outperform the value-weighted counterpart. Moreover, replacing the sample covariance matrix with the Galton covariance matrix can further reduce the realized risk of TCM-RT portfolios as well as their reliance on short positions. The resultant portfolios using Fama-MacBeth forecasts can achieve net-of-costs Sharpe ratios around 1.10. However, if short sales are prohibited, no TCM-RT portfolio can significantly outperform the passive strategy.

Using the same set of optimization inputs, we study the OOS after-cost performance of another mean-variance optimized portfolio: Reward-to-Risk Timing (RRT) portfolio proposed by Kirby and Ostdiek (2012), which features low turnover and prohibits short selling by design. We find that although all RRT portfolios including the one using sample estimates can economically beat the passive strategy, their outperformance is statistically insignificant. For comparison, we also examine the OOS performance of another popular method that exploits stock characteristics in portfolio management: parametric portfolio policies of Brandt, Santa-Clara, and Valkanov (2009, PPP). We find that the best performing portfolio policy can achieve a Sharpe ratio of 0.85 after costs, while an economically sizable gain, this is not statistically significantly different from the value-weighted portfolio. Also, compared to the Markowitz model, it is hard to impose short-sale constraints in the PPP optimization without destroying its convexity.

The remainder of the article is organized as follows. Section 2 reviews the relevant literature. Section 3 and Section 4 describe return forecasting methods and covariance matrix estimators considered in our empirical applications, respectively. Section 5 presents empirical results. Section 6 concludes the paper.

2 Related Literature

Our work contributes to the literature on estimation of mean-variance efficient portfolios. It is well-known that plain Markowitz portfolio performs poorly out of sample. Jobson and Korkie (1980), Michaud (1989) and Best and Grauer (1991) attribute its undesirable OOS performance to the "error-maximization" property of the Markowitz optimizer. Best and Grauer (1991) show that the optimization results are highly sensitive to changes in means. On top of Best and Grauer (1991), Chopra and Ziemba (1993) document that estimation error in means has more impact on optimization than that in variances or covariances. Estimation error is consequential since, as proven by Okhrin and Schmid (2006), it implies that the expectation of optimal weights in terms of the Sharpe ratio does not even exist. In a related result, Kan and Smith (2008) show that the minimum-variance frontier computed using sample estimates is a considerably biased estimator of the true frontier.

Okhrin and Schmid (2007) examine whether the plain Markowitz portfolio can benefit from using improved estimates and find that shrinkage estimators of Jorion (1986) and LW2003 do exhibit better performance than sample estimators. But they also show this improvement only holds when the number of assets is small.[10] Examining 14 mean-variance portfolio optimization methods across 7 financial datasets, DeMiguel, Garlappi, and Uppal (2009) find that no model can consistently outperform the equal-weighted portfolio (a.k.a. the "1/N" rule) in terms of the Sharpe ratio, certainty-equivalent return, or turnover. By contrast, Kirby and Ostdiek (2012) point out that the results of DeMiguel, Garlappi, and Uppal (2009) are sensitive to research design. If the MVE portfolio is constructed by targeting the estimated expected return of the "1/N" portfolio rather than by plugging estimates into the tangency portfolio formula, the resulting optimized portfolios tend to compare favorably to the "1/N" portfolio. Yet, they also show this outperformance is not robust to transaction costs. To improve after-cost performance of mean-variance portfolios, they develop two optimization methods that can deliver portfolios with low turnover and high after-cost profitability. In particular, they can significantly outperform the "1/N" rule after accounting for transaction costs.

Our work is also related to recent studies on estimation of MVE *currency*

[10] Similar to our exercises, they combine an improved mean estimator: shrinkage means of Jorion (1986)) with an improved covariance matrix: shrinkage covariance matrix of LW2003 in the construction of mean-variance optimized portfolios.

portfolios. Filippou, Rapach, Taylor, and Zhou (2022) construct quadratic-utility-maximizing portfolios by combining currency (excess) return forecasts, which are formed using 70 predictors and various machine learning techniques, with exponentially weighted moving average (EWMA) estimator for the covariance matrix. Chernov, Dahlquist, and Lochstoer (2020) estimate the ex ante MVE portfolio using conditional forecasts of currency excess returns based on three currency characteristics and a covariance matrix estimator that combines the nonlinear method of LW2020 with the EWMA model.

Our paper is related to the literature on estimation of expected stock returns. The asset-pricing literature has identified more than 300 characteristics that can predict the cross-section of stock returns (Harvey, Liu, and Zhu 2016). Different from most of previous research, which examines return forecasting ability of stock characteristics in isolation, Haugen and Baker (1996) estimate expected returns with a set of 41 firm characteristics using cross-sectional Fama-MacBeth regressions and construct decile portfolios based on estimated expected returns. They find that the low-expected-return (high-expected-return) decile produces significant negative (positive) raw and factor-adjusted returns. Using the Fama-MacBeth regressions for 15 relatively long-lived characteristics, Lewellen (2015) shows that the expected return estimates have strong predictive power for actual returns. Besides, he analyzes the distribution and accuracy of those estimates. Employing a larger set of 94 firm characteristics, Green, Hand, and Zhang (2017) find that the profitability of long-short portfolios formed on Fama-MacBeth forecasts varies across size groups and over time. Surprisingly, the portfolio formed on all-but-tiny stocks yields insignificant returns after 2003.

Another strand of the return-forecasting literature advocates the use of machine learning techniques. As mentioned above, seminal studies include KPS, FNW, GKX2020 and GKX2021. Bianchi, Büchner, Hoogteijling, and Tamoni (2021) conduct a comparative analysis of machine learning methods in forecasting bond returns. Relative to these strands of literature, our study differs by focusing on how return forecasts can be used to optimize portfolios and their interplay with the estimation of the covariance matrix.

Our study is related to work on covariance matrix estimation. Prior studies such as Ledoit and Wolf (2003, 2004, 2012, 2017, 2020) and Barroso and Saxena (2022) propose to shrink the sample estimator to a target. Specifically, Ledoit and Wolf (2003, 2004) and Barroso and Saxena (2022) use linear shrinkage methods, whereas Ledoit and Wolf (2012, 2017, 2020) suggest the use of nonlinear shrinkage methods. To reduce the number of parameters to be estimated, some papers rely on observable or latent factor models to estimate the covariance matrix (Chan, Karceski, and Lakonishok 1999; Fan, Fan, and Lv 2008; Fan, Liao, and Mincheva

2013; De Nard, Ledoit, and Wolf 2021). Built upon the seminal work of Engle (1982) and Engle (2002, DCC), Hafner and Reznikova (2012) and ELW develop dynamic estimators for large dimensions. The former estimates the intercept matrix of the DCC model using a linear shrinkage approach, while the latter estimates that using a nonlinear shrinkage approach.

Our study contributes to the literature on the covariance matrix by comparing recently proposed methods in a unified setting, in conjunction with forecasting methods for mean returns. It also assesses the implementability of the resulting portfolio rules controlling for stock-level and time-varying transaction costs.

Finally, our work contributes to the literature on portfolio management in the presence of transaction costs. Early work in this field focuses on the optimal allocation between risky and risk-free assets after accounting for proportional trading costs and proposes a no-trade zone for the optimal policy (Magill, Constantinides et al. 1976; Taksar, Klass, and Assaf 1988; Davis and Norman 1990). DeMiguel and Olivares-Nadal (2018) show that the portfolio problem with p-norm trading costs can be regarded as a robust optimization problem, a regularized regression problem, or a Bayesian portfolio problem. In particular, turnover penalization with proportional (quadratic) transaction costs of individual assets is akin to a *lasso* (ridge) penalization. Similarly, Hautsch and Voigt (2019) demonstrate that a mean-variance problem with quadratic transaction costs is equivalent to a mean-variance problem with shrinkage estimators in the absence of transaction costs, and a GMV problem with proportional transaction costs is equivalent to a GMV problem with a regularized covariance matrix. Novy-Marx and Velikov (2016) examine the after-trading-costs performance of 31 anomalies in isolation. Barroso and Detzel (2021) investigate the after-cost profitability of volatility-managed anomaly portfolios. DeMiguel, Martin-Utrera, Nogales, and Uppal (2020) study how the presence of transaction costs impacts the number of jointly significant characteristics using the PPP method.

Our study contributes to this literature by considering a comprehensive set of estimation methods for expected returns and the covariance matrix coupled with realistic transaction cost estimates. From this emerges a vivid illustration of the practical relevance of the growing literature on transaction costs and frictions. It posits that, in the universe of large stocks, managing transaction costs is of crucial importance to obtain *any* meaningful gains from sophisticated covariance and expected return estimation methods.

3 Return Forecasting Methods

3.1 Linear model

3.1.1 Fama-MacBeth (FM) estimator

Two popular linear models used to estimate expected returns are cross-sectional Fama and MacBeth (1973, hereafter FM) regressions and time-series factor regressions. Specifically, the former forms forecasts using the latest known values of stock characteristics along with estimated coefficients on characteristics obtained from historical cross-sectional regressions of returns on lagged characteristics; the latter makes return predictions using estimated loadings of individual stocks on asset-pricing factors coupled with historical means of factor premiums. Fama and French (2020) show that for the same set of characteristics, the FM approach compares favorably to the factor-model counterpart. Hence, our empirical analysis focuses on the former.

For each month between $m = 1$ and $m = t$, we conduct a cross-sectional regression of stock excess returns on lagged stock characteristics (including a constant), which is given by

$$r_{i,m} = \theta_{0,m} + \sum_{j=1}^{J} \theta_{j,m} c_{i,j,m-1} + \epsilon_{i,m} \tag{3.1}$$

where $c_{i,j,m-1}$ denotes the value of characteristic j for stock i at the end of month $m-1$ and $r_{i,m}$ denotes return of stock i in month m. We then compute the time-series average of estimated coefficients on each characteristic to smooth coefficient estimates. On one hand, smoothing can absorb time fixed effects. On the other hand, return forecasts based on smoothed coefficients are less prone to overfitting than those based on estimated coefficients from a pooled regression or from a single cross-sectional regression using data from the most recent month. With the latest known characteristic data as of month t and smoothed coefficient estimates, we form return forecasts ($\hat{\mu}_{t+1|t}$) as

$$\hat{\mu}_{t+1|t} = \tilde{\theta}_{0,t} + \sum_{j=1}^{J} \tilde{\theta}_{j,t} c_{i,j,t} \tag{3.2}$$

$$\text{where} \quad \tilde{\theta}_{j,t} = \frac{1}{t} \sum_{m=1}^{t} \hat{\theta}_{j,m} \quad \text{for} \quad j = 0, \ldots, J, \tag{3.3}$$

where $\hat{\theta}_{j,m}$ denotes the estimated coefficient on characteristic j in month m. Note that we use an *expanding* window to obtain the smoothed coefficient estimates.

3.1.2 Shrinkage FM estimators

To mitigate noise in FM estimates, Lewellen (2015) suggests a shrinkage FM estimator that shrinks FM estimates to their cross-sectional mean, which is given by

$$\tilde{\mu}_{i,t+1|t} = (1-s) \times \text{CSMean}(\hat{\mu}_{i,t+1|t}) + s \times \hat{\mu}_{i,t+1|t}, \tag{3.4}$$

where $\hat{\mu}_{i,t+1|t}$ denotes a FM estimate obtained from Eq. (3.2) and s governs the shrinkage intensity. Similarly, Barroso and Saxena (2022) propose a general shrinkage estimator (hereafter Galton) that can be used as a correction on top of any estimator for expected returns. A Galton-corrected FM estimate (hereafter GFM) can be expressed as

$$\tilde{\mu}_{i,t+1|t} = s_0 \times 1 + s_1 \times \hat{\mu}_{i,t+1|t}. \tag{3.5}$$

They implicitly assume a shrinkage target of 1, which is also used in DeMiguel, Martin-Utrera, and Nogales (2013) and corresponds to the estimated expected return in the GMV framework.

Both methods estimate (monthly) optimal shrinkage parameters by running a cross-sectional regression of realized returns on FM estimates. Lewellen (2015) shows that the estimated slope of FM estimates ($\hat{\beta}_{FM}$) is the optimal (in-sample or pseudo OOS) weight on $\hat{\mu}_{i,t+1|t}$ in the sense that the difference in mean-squared-forecast-errors (MSFE) between the null forecast ($\text{CSMean}(\hat{\mu}_{i,t+1|t})$) and the shrinkage forecast ($\tilde{\mu}_{i,t+1|t}$) is maximized at that value (see his footnote 3 for more details).

The GFM approach uses not only $\hat{\beta}_{FM}$ but also estimated regression intercept ($\hat{\beta}_{intercept}$) to form shrinkage forecasts. Hence, the GFM forecasts minimize (pseudo OOS) MSFE and have the same mean as realized returns. A GFM forecast can be reformulated as

$$\tilde{\mu}_{i,t+1|t} = \frac{\hat{\beta}_{intercept}}{\text{CSMean}(\hat{\mu}_{\text{i,t+1|t}})} \times \text{CSMean}(\hat{\mu}_{\text{i,t+1|t}}) + \hat{\beta}_{FM} \times \hat{\mu}_{i,t+1|t}.$$

It shows that when $\frac{\beta_{intercept}}{\text{CSMean}(\hat{\mu}_{\text{i,t+1|t}})} = 1 - \hat{\beta}_{FM}$, the GFM estimate becomes the shrinkage estimate of Lewellen (2015).

3.1.3 GFM estimator

We employ the GFM estimator in our main analysis and investigate whether it can improve the performance of the FM estimator in portfolio optimization contexts. The shrinkage parameters of GFM are estimated with the time-series average of historical monthly shrinkage parameters. That is, we get smoothed shrinkage estimators by performing cross-sectional Fama-Macbeth regressions of historical realized returns on historical FM forecasts. For each month between month $m = 121$

and month $m = t$ (i.e., we require at least 10-year data to form an FM estimate), we perform a cross-sectional regression, that is,

$$r_{i,m} = g_{0,m} + g_{1,m}\hat{\mu}_{i,m|m-1} + \epsilon_{i,m} \tag{3.6}$$

The estimated regression coefficients are denoted by $\hat{g}_{0,m}$ and $\hat{g}_{1,m}$. We then form an GFM forecast using the specification:

$$\tilde{\mu}_{i,t+1|t} = \tilde{g}_{0,t} + \tilde{g}_{j,t}\hat{\mu}_{i,t+1|t} \tag{3.7}$$

$$\text{where} \quad \tilde{g}_{j,t} = \frac{1}{t-120}\sum_{m=121}^{t}\hat{g}_{j,m} \quad \text{for} \quad j = 0, 1. \tag{3.8}$$

Besides shrinkage, the Galton procedure can be interpreted as learning from pseudo OOS forecast errors of the FM estimator, which is akin to the validation procedure in machine learning models. Moreover, it is more general than Lewellen (2015)'s shrinkage estimator in the sense that the GFM forecasts are not restricted to be convex combinations of FM forecasts and a shrinkage target.

3.1.4 Conditional linear latent factor model

KPS propose an asset pricing model in which stock excess returns are assumed to follow a latent factor structure. The restricted version of their model is given by

$$r_{i,t} = \beta(c_{i,t-1}) \times f_t + \epsilon_{i,t}, \tag{3.9}$$

where the intercept is restricted to be zero, f_t denotes a $K \times 1$ vector of returns on K latent factors in month t and $\beta(c_{i,t-1})$ denotes a $1 \times K$ vector of factor loadings. Factor loadings are defined as a linear combination of observable stock characteristics, that is,

$$\beta(c_{i,t-1}) = c_{i,t-1}^T \Gamma_\beta, \tag{3.10}$$

where $c_{i,t-1}$ is a $P \times 1$ vector of characteristics and Γ_β denotes a $P \times K$ matrix that defines mapping from P observable characteristic to K betas ($P \gg K$). Different from Kozak, Nagel, and Santosh (2018) who propose a latent factor model based on principal component analysis of anomaly returns, KPS use observable stock characteristics as instrumental variables to determine loadings on latent common factors. So, the model is named as Instrumented Principal Component Analysis (IPCA). A return forecast implied by the IPCA model is given by

$$\hat{\mu}_{i,t+1|t} = c_{i,t}^T \hat{\Gamma}_{\beta,t}\hat{f}_{t+1|t}, \tag{3.11}$$

where $\hat{\Gamma}_{\beta,t}$ denotes the mapping matrix estimated from data up to month t and $\hat{f}_{t+1|t}$ denotes expected returns of latent factors that are estimated with means of the estimated factor realizations up to month t.[11]

3.2 Additive regression model

It is possible that stock characteristics are nonlinearly associated with future returns. FNW account for nonlinearities using an additive regression model, which is effectively a linear regression model with nonlinearly transformed covariates. FNW formulate the model as

$$r_{i,t} = g(c_{i,t-1}; \theta_t, m(\cdot)) = \sum_{j=1}^{J} p(c_{i,t-1,j})'\theta_t^{(j)} + \epsilon_{i,t}, \tag{3.12}$$

where $c_{i,t-1,j}$ denotes the value of characteristic j for stock i in month $t-1$, $p(\cdot)$ denotes a vector of basis functions applied to each characteristic and $\theta_t^{(j)}$ denotes a vector of coefficients associated with characteristic j.[12] To mitigate overfitting resulting from series expansion, FNW use the adaptive group *lasso* of Huang, Horowitz, and Wei (2010) for model selection and estimation. The corresponding optimization problem is given by

$$\arg\min_{\theta_t} \frac{1}{2N} \left\| r - c^T\theta \right\|_2^2 + \lambda \sum_j w_j \left\| \theta_t^{(j)} \right\|_2, \tag{3.13}$$

$$\text{where} \quad \left\| \theta_t^{(j)} \right\|_2 = \sqrt{\sum_{s \in I_j} \theta_{t,s}^2} \quad \text{and} \quad w_j = \frac{1}{||\tilde{\theta}_t^{(j)}||_2}. \tag{3.14}$$

r denotes an $N \times 1$ vector of (demeaned) excess returns; c denotes a $\tilde{J} \times N$ matrix of (demeaned) characteristics where $\tilde{J} = S \times J$; S and J denote the number of spline terms and characteristics, respectively; θ_t is a $\tilde{J} \times 1$ vector of coefficients; $\tilde{\theta}_t^{(j)}$ denotes the optimal coefficients associated with characteristic j when $w_j = 1$ (i.e., the penalty term in ordinary group *lasso*). λ is a hyperparameter.[13] We choose λ using Bayesian Information Criterion of Schwarz et al. (1978) in line with Yuan and Lin (2006).

[11] We appreciate authors of KPS for making their replication code available.

[12] Following FNW, we use quadratic truncated power basis functions to represent $m(\cdot)$. E.g., for quadratic splines with 2 knots at ξ_1 and ξ_2, according to page 144 of Hastie, Tibshirani, and Friedman (2009), the truncated power basis functions are $p_1(x) = 1$, $p_2(x) = x$, $p_3(x) = x^2$, $p_4(x) = (x - \xi_1)_+^2$, $p_5(x) = (x - \xi_2)_+^2$.

[13] We select the optimal λ from a grid of 100 candidates ranging from $0.0001\lambda_{min}$ to λ_{min} where λ_{min} is the minimum λ giving a solution of all zeroes.

3.3 Artificial intelligence

3.3.1 "Feed-forward" neural network

Compared to more advanced nonlinear models such as artificial neural networks, additive models cannot account for interactions among return predictors. For example, the return forecasting ability of some characteristics varies across size groups.[14] We follow GKX2020 to use a "feed-forward" neural network. Specifically, it consists of input, hidden and output layers. For each node in hidden layers, we apply a nonlinear activation such as rectified linear unit (ReLU) to the aggregated inputs from the last layer. The objective function is mean squared forecast errors with ℓ_1 penalization. To mitigate overfitting, model estimation is accompanied by four regularization methods: learning rate shrinkage, early stopping, batch normalization, and ensembles. The candidate hyperparameter values we consider are the same as those used by GKX2020.[15] We use the Adam optimizer to estimate the model.

3.3.2 Conditional autoencoder neural network

Another return forecasting method based on artificial neural networks is the conditional autoencoder (CA) model of GKX2021, which is on top of the IPCA model and standard autoencoder neutral networks. Under a latent factor structure, it uses a standard autoencoder for the factor extraction and a "Feed-Forward" neutral network for the estimation of conditional betas. Akin to the IPCA model, the CA model estimates risk exposures of individual stocks with lagged characteristics. The model estimates premiums of latent factors with linear combinations of returns on characteristic-managed portfolios. We use the same model hyperparameters as GKX2021 with the exception of batch size.[16]

[14]As shown in FNW, it is possible to interact each non-size characteristic with firm size in an additive regression model to account for this effect. However, it is infeasible to account for multiple interaction effects.

[15]ℓ_1 penalty $\in$ $\{10^{-5}, 10^{-4}, 10^{-3}\}$; Learning rate$\in$ $\{0.001, 0.01\}$; Batch Size = 10000; Epochs=100; Patience=5; Ensemble=10.

[16]Figure 2 of GKX2021 implies that batches should be drawn at monthly level. Specifically, before each estimation, we convert all observations into three 3-D tensors with dimensions of (# historical months, # stocks, # features), (# historical months, # managed portfolios, 1), and (# historical months, # stocks, 1), respectively. Invalid observations are masked (details can be found on `https://www.tensorflow.org/guide/keras/masking_and_padding`). The first two tensors correspond to model inputs (i.e., stock characteristics and characteristic-managed portfolio returns, respectively), while the last one is model output (i.e. realized risk premium). The batch size we consider is 1. Since batches are defined over the first dimension, the number of stock observations a batch has is the number of valid stocks in a month.

4 Covariance Matrix Estimation

4.1 Sample covariance matrix

The sample covariance matrix is defined as

$$\hat{\Sigma}^{sam}_{t+1|t} = \frac{1}{H-1}\sum_{h=0}^{H-1}(r_{t-h} - \bar{r}) \times (r_{t-h} - \bar{r})', \tag{4.1}$$

where H denotes the length of an estimation window and $\bar{r}$ denotes a vector of historical means of returns. It is an unbiased estimator for the population counterpart and also the maximum likelihood estimator under normality. However, when the number of stocks exceeds the number of observations, it is rank deficient and thus not invertible. Even if it is not singular, its inverse is a biased estimator for the inverse of the population counterpart.[17] To overcome its shortcomings, the literature has proposed several alternative estimators using factor models or shrinkage methods.

4.2 Estimator implied by a factor model

Let's impose following factor structure on the data-generating process

$$r_{i,h} = \beta_i' f_h + \epsilon_{i,h}, \tag{4.2}$$

where $r_{i,h}$ is the excess return on stock i at time h, β_i is a $K \times 1$ vector of factor loadings for stock i, f_h is a $K \times 1$ vector of factor returns at time h, and $\epsilon_{i,h}$ denotes the residual return. The covariance matrix of residuals is assumed to be a diagonal matrix with residual variances on its main diagonals. Under this assumption, it is an exact factor model (EFM) (De Nard, Ledoit, and Wolf 2021). The covariance matrix implied by an EFM estimated at time t is given by

$$\hat{\Sigma}^{fac}_{t+1|t} = \hat{\beta}_t' \hat{\Sigma}_{f,t+1|t} \hat{\beta}_t + diag(\hat{\Sigma}_{\epsilon,t+1|t}), \tag{4.3}$$

where $\hat{\beta}_t$ denotes a $K \times N$ matrix of estimated factor loadings at time t, $\hat{\Sigma}_{f,t+1|t}$ denotes the sample covariance matrix of K factors computed at time t and $diag(\hat{\Sigma}_{\epsilon,t+1|t})$ is a diagonal matrix with the sample variances of residuals on its main diagonals..

[17] Bai and Shi (2011) shows that when the number of stocks is $H/2+2$, $E(\hat{\Sigma}^{-1}_{t+1|t}) = 2\Sigma^{-1}_{t+1}$ under normality.

4.2.1 Single index model

When Eq. (4.2) uses the market index only, it corresponds to the single-index model of Sharpe (1963).[18] The conditional covariance matrix implied by the model is estimated as

$$\hat{\Sigma}^{capm}_{t+1|t} = \hat{\beta}_t' \hat{\beta}_t \hat{\sigma}^2_{mkt,t+1|t} + diag(\hat{\Sigma}_{\epsilon,t+1|t}), \tag{4.4}$$

where $\hat{\beta}_t$ is the estimated market beta at time t and $\hat{\sigma}^2_{mkt,t+1|t}$ is the sample variance of excess returns on the market portfolio.

4.2.2 Latent factor model

When f_h is assumed to be unobservable, Eq. (4.2) is a latent factor model. We can estimate the corresponding covariance matrix by normalizing the covariance matrix of factors to an identity matrix. Under this restriction, we can estimate the covariance matrix by principal component analysis (PCA) or maximum likelihood method (Bai and Shi 2011). The covariance matrix based on PCA is given by

$$\hat{\Sigma}^{pca}_{t+1|t} = \hat{B}_t \hat{B}_t' + \hat{\Sigma}_{\epsilon,t+1|t}, \tag{4.5}$$

where $\hat{B}_t = (\hat{\lambda}^{0.5}_{1,t}\hat{q}_{1,t}, ..., \hat{\lambda}^{0.5}_{K,t}\hat{q}_{K,t})$, $\hat{\Sigma}_{\epsilon,t+1|t} = \hat{\Sigma}^{sam}_{t+1|t} - \hat{B}_t\hat{B}_t'$, K denotes the number of principle components (factors), $\hat{\lambda}_{1,t} \geq \hat{\lambda}_{2,t}... \geq \hat{\lambda}_{K,t} \geq 0$ are the ordered eigenvalues of $\hat{\Sigma}^{sam}_{t+1|t}$ and $\hat{q}_{k,t}$ is the corresponding orthonormal eigenvector. To reduce sampling errors or overcome rank-deficiency, we can restrict $\hat{\Sigma}_{\epsilon,t+1|t}$ to be a diagonal matrix or impose sparsity on it. For example, Fan, Liao, and Mincheva (2013) apply soft thresholding to it and the resultant covariance matrix estimator is named as Principal Orthogonal complEment Thresholding (POET). The estimator has three desirable properties: 1) it relies on past return data only; 2) it accounts for covariances of residual returns; 3) it is always positive definite as long as a proper thresholding constant is used.

4.3 Linear shrinkage estimators

A linear shrinkage estimator shrinks $\hat{\Sigma}^{sam}_{t+1|t}$ to a target. Despite introducing biases, it can reduce the variability of estimation. The idea can be traced back to the breakthrough work of Stein et al. (1956) and James, Stein et al. (1961). They show that shrinking the sample mean towards a target can reduce mean squared errors. The challenges regarding the estimator lie in how to find a proper target and optimal shrinkage parameters.

[18] Since we use the market factor of the CAPM model (Sharpe 1964; Lintner 1965; Black 1972), we name the estimator as EFM_{capm} or $\hat{\Sigma}^{fac}_{t+1|t}$ if covariances of residuals are assumed to zeroes.

4.3.1 Linear shrinkage estimators based on Frobenius loss function

A linear shrinkage estimator is of form:

$$\hat{\Sigma}^{shr}_{t+1|t} = \gamma_0 \Omega + \gamma_1 \hat{\Sigma}^{sam}_{t+1|t}, \tag{4.6}$$

where γ_0 and γ_1 are shrinkage parameters governing the shrinkage intensity and Ω denotes a shrinkage target. Ledoit and Wolf (2003) and Ledoit and Wolf (2004) respectively use the covariance matrix implied by the single index model and that based on a constant correlation model as shrinkage targets. Both studies assume a convex combination of $\hat{\Sigma}^{sam}_{t+1|t}$ and Ω, and calculate optimal shrinkage parameters by minimizing the Frobenius norm of the differences between $\hat{\Sigma}^{shr}_{t+1|t}$ and the population counterpart.

4.3.2 Regression-based linear shrinkage estimator

Based on stylized facts that sample estimates of variance, pairwise correlation or pairwise covariance regress to the mean, Barroso and Saxena (2022) propose shrinkage estimators for the three variables (a.k.a. Galton estimators). Let $X^H_{k,m-E}$ denote a sample estimate of variable X (variance, pairwise covariance or pairwise correlation) for stock i calculated from returns over the period month $m-E-H-1$ to month $m-E$ (i.e., an estimation window of H months) and $X^E_{i,m}$ denote a subsequent realization calculated from returns over the period month $m-E+1$ to month t. The Galton regression model is formulated as

$$X^E_{k,m} = g_{0,m} + g_{1,t} X^H_{k,m-E} + \epsilon_{k,m}. \tag{4.7}$$

The Galton parameters in month t ($g_{0,t}$ and $g_{1,t}$) are the time-series averages of $g_{0,m}$ and $g_{1,m}$ obtained from historical Galton regressions. When estimating the Galton covariance matrix, we perform the Galton regression for variances and pairwise correlations in isolation. The Galton covariance estimator is calculated as

$$\hat{\Sigma}^{gal}_{t+1|t} = diag\Big((\hat{X}^{gal,var}_{t+1|t})^{\frac{1}{2}}\Big) \hat{X}^{gal,corr}_{t+1|t} diag\Big((\hat{X}^{gal,var}_{t+1|t})^{\frac{1}{2}}\Big),$$

where $\hat{X}^{gal,var}_{t+1|t}$ denotes a vector of Galton variances calculated at the end of month t and $\hat{X}^{gal,corr}_{t+1|t}$ denotes the Galton correlation matrix at the end of month t. To ensure that all diagonal elements of $\hat{X}^{gal,corr}_{t+1|t}$ are ones, $\hat{X}^{gal,corr}_{t+1|t}$ is calculated as

$$\hat{X}^{gal,corr}_{t+1|t} = (1 - \tilde{g}_{0,t} - \tilde{g}_{1,t}) I + \tilde{g}_{0,t} \mathbb{1} + \tilde{g}_{1,t} \hat{\Sigma}^{sam,corr}_{t+1|t},$$

where I is an identity matrix, $\mathbb{1}$ is an all-ones matrix, and $\hat{\Sigma}^{sam,corr}_{t+1|t}$ is the sample correlation matrix; $\tilde{g}_{0,t}$ and $\tilde{g}_{1,t}$ denote Galton parameters for pairwise correlation as of month t. Suppose $\hat{\Sigma}^{sam,corr}_{t+1|t}$ is positive definite, $\hat{X}^{gal,corr}_{t+1|t}$ is positive definite only if the Galton parameters are non-negative and their sum is less than or equal to one (except for the scenario of $\tilde{g}_{1,t} = 0$ and $\tilde{g}_{0,t} = 1$). Note that both conditions always hold in our empirical exercises.

4.4 Nonlinear shrinkage estimator

The spectral decomposition of the sample covariance matrix is given by:

$$\hat{\Sigma}^{sam}_{t+1|t} = \hat{Q}_t \hat{\Lambda}_t \hat{Q}'_t, \tag{4.8}$$

where $\hat{\Lambda}_t$ is a diagonal matrix whose diagonal elements are ordered eigenvalues, and $\hat{Q}_t$ is an *orthonormal* matrix whose columns are corresponding eigenvectors. Based on the decomposition, Ledoit and Wolf (2012) creatively propose a nonlinear shrinkage estimator

$$\hat{\Sigma}^{nl}_{t+1|t} = \hat{Q}_t \hat{\Delta}_t \hat{Q}', \tag{4.9}$$

where Δ_t is a diagonal matrix whose elements are functions of eigenvalues of $\hat{\Sigma}^{sam,corr}_{t+1|t}$, that is, $\delta_i = \phi(\lambda_i)$ where ϕ is a shrinkage function that pushes up relatively small eigenvalues and pushes down relatively large eigenvalues. On one hand, adjusting eigenvalues properly can mitigate the overfitting of $\hat{\Sigma}^{sam}_{t+1|t}$. On the other hand, it is always positive definite even when $\hat{\Sigma}^{sam}_{t+1|t}$ is singular. Based on random matrix theory, Ledoit and Wolf (2012) propose to compute δ_i numerically, but it is computationally intensive. Instead, we use the analytical formula of Ledoit and Wolf (2020) to get $\hat{\Sigma}^{nl}_{t+1|t}$. LW2020 show that the analytical method is typically 1000 times faster than the numerical counterpart without hurting accuracy.

4.5 Dynamic covariance estimator

The last estimator we consider is a dynamic covariance matrix estimator in the spirit of Engle (1982) and Engle (2002). Engle, Ledoit, and Wolf (2019) propose a dynamic estimator for a large number of assets. They first model conditional variances using Generalized AutoRegressive Conditional Heteroskedasticity (GARCH), that is,

$$\sigma^2_{i,h} = \omega_i + a_i \epsilon^2_{i,h-1} + b_i \sigma^2_{i,h-1} \tag{4.10}$$

where (ω_i, a_i, b_i) are GARCH(1,1) parameters for stock i, and $\epsilon^2_{i,h-1}$ denotes stock i's (demeaned) return at time $h-1$. The model is estimated from daily return data.

They then model dynamics of the correlation matrix using the multivariate

GARCH (1,1), that is,

$$Q_h = \Omega + \alpha \nu_{h-1} \nu'_{h-1} + \beta Q_{h-1}. \tag{4.11}$$

where ν_h is a vector of devolatized residuals calculated as $\nu_h = D_h^{-1} \epsilon_h$ and D_h denotes a diagonal matrix whose main diagonal consists of fitted conditional volatility. Instead of estimating Ω that has $\frac{N(N+1)}{2}$ unknown elements, ELW use the dynamic constant correlation (DCC) model of Engle (2002) in which Ω is replaced with $(1 - \alpha - \beta)C$. C is the unconditional correlation matrix of ν_h, which can be estimated as $\hat{C} = \frac{1}{H} \sum_{t=1}^{H} \nu_h \nu'_h$. However, the DCC model is not compatible with a large dimension. As argued by ELW, on one hand, it is not always feasible to estimate a DCC model for large dimensions. On the other hand, the DCC matrix for large dimensions tend to deliver unfavorable performance. The ELW estimator addresses the two issues by using the composite-likelihood method of Pakel, Shephard, Sheppard, and Engle (2020) and applying the numerical nonlinear shrinkage method of LW2012 to $\hat{C}$, respectively. Again, we use the analytical formula of LW2020 to shrink $\hat{C}$ nonlinearly in our empirical exercises.

5 Empirical Results

5.1 Data

5.1.1 Investment universe

At the end of each calendar year, we first select NYSE, AMEX, and NASDAQ common stocks (CRSP share code of 10 or 11) with a full return history over the past 60 months and a full return history over the subsequent 12 months. We then select the 500 *largest* stocks by market capitalization. These account for 77% of total market cap on average. The investment universe keeps unchanged for the following 12 months. We build the sample using the monthly stock file from the Center for Research in Security Prices (CRSP). Our sample selection method is standard in the literature (e.g. LW2017, DLW and Barroso and Saxena (2022)).[19] Our OOS period begins in January 1987 and ends in December 2020.[20]

[19] The restriction that stocks must have valid observations over following 12 months is subject to look-ahead bias. However, as mentioned in LW2017, this forward-looking restriction is commonly applied in the portfolio optimization literature.

[20] Our start date reflects the need of having enough historical data to implement machine learning methods and the Galton correction. It is the same start date as in closely related studies (Gu, Kelly, and Xiu 2020, 2021; Avramov, Cheng, and Metzker 2022).

5.1.2 Stock characteristics

We use the same set of stock characteristics as FNW and strictly follow their procedure to construct a characteristic dataset.[21] Our dataset includes all common stocks traded on NYSE, AMEX or NASDAQ. We obtain stock market data from CRSP and accounting information from the Compustat Annual Fundamental Files. The stock market data are assumed to be publicly available immediately after being generated, while the accounting data for a fiscal year ending in calendar year $t-1$ are assumed to be publicly observable in June of calendar year t. A stock is defined as a *large* stock if its market value is above the NYSE median at the beginning of a month (Lewellen 2015). Our characteristic data begin in July 1964 and end in December 2020. In line with Freyberger, Neuhierl, and Weber (2020), Kelly, Pruitt, and Su (2019), Gu, Kelly, and Xiu (2020), Gu, Kelly, and Xiu (2021), and Chen, Pelger, and Zhu (2020), we use ranked-normalized characteristics to form return forecasts. See Appendix A for details about data transformation and missing data handling. Table A1 presents a list of 62 characteristics by category. Note that some characteristics have weak predictive power for returns over the full sample period or that the predictive power of some characteristics concentrates in small- and micro-cap stocks. We do not remove such characteristics to capture the idea that an unsophisticated investor ex ante does not know whether a stock characteristic has significant predictive power or whether there exist interaction effects between firm size and a characteristic in estimating expected returns.

[Insert Table A1 near here]

5.2 Statistical inference

We use the i.i.d. studentized bootstrap method of Ledoit and Wolf (2008) and Ledoit and Wolf (2011) to compute a two-sided p-value for the null hypothesis of equal variances or (net) Sharpe ratios between an investment strategy and the *value-weighted* portfolio.[22] We resample individual pairs of returns from observed pairs with replacement to get bootstrap data. A studentized test statistic of the difference in Sharpe ratios or natural logarithm of variances for the original data is given by[23]

$$d = \frac{|\hat{\Delta}|}{s(\hat{\Delta})},$$

[21] We are grateful to Dr. Michael Weber for sharing his SAS code used to generate characteristic data.

[22] DeMiguel et al. (2020) also use this method to test the difference of Sharpe ratios between their OOS regularized PPP portfolios and a benchmark portfolio

[23] Note that Ledoit and Wolf (2008) and Ledoit and Wolf (2011) use a *symmetric* studentized bootstrap confidence interval.

where $\hat{\Delta}$ denotes difference of Sharpe ratios or (log) variances and $s(\hat{\Delta})$ denotes its standard error computed using the delta method (see Appendix B for more details about the calculation of standard errors).

For the bootstrap data, the centered studentized statistic estimated from the m^{th} bootstrap sample is given by

$$\tilde{d}^{*,m} = \frac{|\hat{\Delta}^{*,m} - \hat{\Delta}|}{s(\hat{\Delta}^{*,m})},$$

where $\hat{\Delta}^{*,m}$ denotes the difference of Sharpe ratios or (log) variance between two strategies for the m^{th} bootstrap sample and $s(\hat{\Delta}^{*,m})$ denotes its standard error computed using the delta method. The two-sided p-value for the null hypothesis is given by

$$p = \frac{I(\tilde{d}^{*,m} \geq d) + 1}{M + 1}, \tag{5.1}$$

where $I(\cdot)$ represents an indicator function and M is the number of bootstrap resamples.

5.3 Transaction costs

The turnover of strategy j (TO_t^j) at the end of month t is defined as

$$TO_t^j = \sum_i |w_{i,t}^j - \tilde{w}_{i,t-1}^j|, \tag{5.2}$$

where $w_{i,t}^j$ is the weight of stock i in portfolio j after rebalancing at the end of month t and $\tilde{w}_{i,t-1}^j$ is the weight of stock i in portfolio j before rebalancing at the end of month t, that is,

$$\tilde{w}_{i,t-1}^j = \frac{w_{i,t-1}(1 + r_{i,t})}{\sum_i w_{i,t-1}^j (1 + r_{i,t})}.$$

The sum of weights after rebalancing is always 1. Asset i can be a risk-free asset that is assumed to incur zero transaction costs.[24] When computing turnover as well as transaction costs of strategy j incurred at the end of month t, we consider all stocks that are used to form strategy j at the end of month t or at the end of month $t-1$. That is, if stock i that is not used in strategy j at the end of month $t-1$ is part of the strategy at the end of month t, $\tilde{w}_{i,t-1} = 0$; if stock i is included in strategy j at the end of month $t-1$ but excluded from the strategy at the end of month t, $w_{i,t}^j = 0$.

We define the net-of-costs return and Sharpe ratio in the same way as Brandt,

[24] Hence, to better reflect potential transaction costs incurred, the portfolio turnover reported in our tables does not take the turnover of the risk-free asset into account.

Santa-Clara, and Valkanov (2009, see their Equation (19)) and Hautsch and Voigt (2019, see their Equations (41) through (44)). The net-of-costs return of portfolio j in month $t+1$ is defined as

$$r_{t+1}^{j,net} = \sum_i w_{i,t}^j \times r_{i,t+1} - \sum_i \kappa_{i,t} |w_{i,t}^j - \tilde{w}_{i,t-1}^j|, \tag{5.3}$$

where the first part is gross return of portfolio j in month $t+1$ and the second part captures the trading costs incurred to rebalance from $\tilde{w}_{i,t-1}^j$ to $w_{i,t}^j$ at the end of month t. Net-of-costs Sharpe ratio is given by

$$\text{SR}^{j,net} = \frac{\text{Mean}(r_t^{j,net} - rf_t)}{\text{Std}(r_t^{j,net} - rf_t)},$$

where rf_t is the risk-free rate in month t.

We use one half of Chung and Zhang (2014)'s quoted spread supplemented with Abdi and Ranaldo (2017)'s CHL method to estimate effective one-way spreads of stocks ($\kappa_{i,t}$). See Appendix C for details about the estimation of $\kappa_{i,t}$. Fig. 1 depicts the dynamics of monthly minimum, maximum and average $\kappa_{i,t}$ for the top-500 stocks by market value over the OOS period. The figure shows that our estimates can reflect both cross-sectional and time-series variation in effective spreads. Moreover, our estimates accurately reflect major liquidity-related market events. For example, average spreads become significantly tighter after decimalization in 2001.

[Insert Figure 1 near here]

5.4 Forecasting performance of expected return estimators

In this section, we examine OOS predictive performance of 10 expected return estimators including

- $\textbf{SAM}_{\textbf{mean}}$: the average return over the past 60 months.
- $\textbf{GAL}_{\textbf{mean}}$: the Galton-corrected sample mean.
- $\textbf{FM}_{\textbf{large}}$ and $\textbf{GFM}_{\textbf{large}}$: ordinary and Galton-corrected Fama-MacBeth estimators as defined in Section 3.1.1 and Section 3.1.3, respectively. Subscript "large" indicates that Fama-MacBeth regressions are implemented using large-cap stocks only.
- $\textbf{FM}_{\textbf{wls}}$ and $\textbf{GFM}_{\textbf{wls}}$: ordinary and Galton-corrected Fama-MacBeth estimators as defined in Section 3.1.1 and Section 3.1.3, respectively. Subscript "wls"

indicates that Fama-MacBeth regressions are implemented using weighted least squares with market capitalization as weights.

- **$\mathbf{AGLASSO_{large}}$**: an estimator based on the nonparametric model of FNW. We first perform model/characteristic selection for large-cap stocks using data up to the beginning of the OOS period. At the end of each formation month during the OOS period, we estimate the model on the selected characteristics using data of large-cap stocks over the past 120 months. Then, we use the estimated coefficients along with the latest values of selected characteristics to predict one-month-ahead returns. The adaptive group *lasso*, which is described in Section 3.2, is used for both characteristic selection and model estimation.[25]

- **$\mathbf{IPCA_{K=5,large}}$**: return forecasts implied by the the restricted IPCA model with 5 latent factors, which is introduced in Section 3.1.4. We use data of large-cap stocks to perform recursive backward-looking estimation of the model.

- **$\mathbf{CA2_{K=5}}$**: the conditional autoencoder (CA) model described in Section 3.3.2. We use the CA model consisting of 2 hidden layers and 5 latent factors. We follow GKX2021 to update model parameters yearly (at the end of June). We use a 12-year rolling validation window and an expanding estimation window, that is, we enlarge the training data by one year every time we refit.[26]

- **NN3**: the "Feed-Forward" Neural Network (NN). We choose the best performing NN architecture in GKX2020, which consists of three hidden layers with 32, 16 and 8 neurons, respectively. We update model parameters at the end of June. The definitions of training and validation samples are the same as those for the $CA2_{K=5}$ model.

Except for SAM_{mean} and GAL_{mean}, which use daily returns over the past 60 months to form return forecasts, all estimators exploit stock characteristics in return forecasting.[27] Since $CA2_{K=5}$ and NN3 already account for characteristic interactions by design, including the effect of size on other characteristics, we do not restrict the

[25] We use R package "oem" developed by Xiong, Dai, Huling, and Qian (2016) to conduct model selection and estimation.

[26] The numbers of hidden layers and latent factors are hyperparameters. We could dynamically find the optimal values of the two hyperparameters using a validation procure as we do for other hyperparameters; however, it is computationally intensive to do so. Hence, we follow prior studies to use fixed values over the entire OOS period. The 5-factor model is the specification examined in all empirical applications of KPS and GKX2021. Also, it is in line with the 5-factor model of Fama and French (2015). The $CA2_{K=5}$ model is examined in all empirical exercises of GKX2021 and examined by Avramov, Cheng, and Metzker (2022) as well. In addition, the lengths of estimation and validation windows are hyperparameters. We follow GKX2020 to use a validation window of 12 years. We fit the $CA2_{K=5}$ and NN3 models using *TensorFlow 2.6.1* via *Python*.

[27] For estimates of expected returns and covariance matrices computed from daily data, we convert them to monthly estimates in portfolio optimization.

sample to large-cap stocks or use value-weighted forecast errors when estimating their parameters.

We follow the recent literature (e.g., KPS; FNW; GKX2020; GKX2021) to measure the forecasting performance of an expected return estimator for expected returns using the OOS predictive R-squared, which is defined as

$$R_{oos}^2 = 1 - \frac{\sum_{(i,t)\in oos}(r_{i,t+1} - \hat{\mu}_{i,t+1|t})^2}{\sum_{(i,t)\in oos} r_{i,t+1}^2}. \tag{5.4}$$

It compares model predictions ($\hat{\mu}_{i,t+1|t}$) against a naive forecast of zero and pools prediction errors across stocks and over time. Table 1 shows R_{oos}^2 of each estimator for stocks included in our investment universe.

[Insert Table 1 near here]

Consistent with prior studies, SAM_{mean} is inferior to a naive forecast of zero. However, as shown in Barroso and Saxena (2022), the predictive performance of SAM_{mean} can be dramatically improved by learning from its own past OOS forecast errors. GAL_{mean} achieves an R_{oos}^2 of 0.50%, which is 2.25% higher than that of SAM_{mean} and outperforms the naive counterpart. The forecasting performance of GAL_{mean} is comparable to that of the two Fama-MacBeth estimators: FM_{large} and FM_{wls}. Shrinking Fama-Macbeth forecasts using the Galton correction improves predictions. $\text{GFM}_{\text{large}}$ and GFM_{wls} raise R_{oos}^2 to 0.61% and 0.64%, respectively.

The last four rows report the predictive performance of four machine learning methods. $\text{AGLASSO}_{\text{large}}$ produces an R_{oos}^2 of 0.25%. This suggests that accounting for nonlinearities through basis expansions of predictors can result in deterioration of predictive performance compared to the Fama-MacBeth counterparts.[28] The two methods based on a conditional latent factor structure ($\text{IPCA}_{\text{K=5,large}}$ and $\text{CA2}_{\text{K=5}}$) exhibit similar forecasting performance with R_{oos}^2 of 0.58% and 0.61%, respectively. NN3 is the best performing model with an R_{oos}^2 of 0.66%.

The evidence suggests that all alternative estimators we consider provide superior forecasting performance compared to sample means. Also, despite achieving higher R_{oos}^2 than the two ordinary Fama-MacBeth estimators, complex methods such as $\text{CA2}_{\text{K=5}}$ and NN3 are not necessarily superior, at least for large-cap stocks, since they rely on hyperparameters, which may reflect informed choices after looking at the data, and have relatively high computational costs.

[28] GKX2020 examine the predictive performance of an additive regression model expanded by quadratic splines in conjunction with the group *lasso* and Huber loss. They report that R_{oos}^2 for the top-1000 stocks by market value over the 1987-2016 period is 0.14%.

5.5 Mean-variance efficient portfolio

Next, we explore whether the superior forecasting performance of the alternative estimators relative to the sample mean can translate into better OOS performance of the (conditional) mean-variance efficient (MVE) portfolio. We follow Kirby and Ostdiek (2012) to assume that at the end of month t, an investor with mean-variance preferences selects a $N \times 1$ vector of risky-asset weights to maximize his/her expected utility. The problem is formulated as

$$\max_{w_t} \; w_t' \hat{\mu}_{t+1|t} - \frac{\gamma}{2} w_t' \hat{\Sigma}_{t+1|t} w_t, \tag{5.5}$$

where $\hat{\mu}_{t+1|t}$ denotes a vector of estimates of the expected excess returns on the N risky assets, $\hat{\Sigma}_{t+1|t}$ denotes an estimate of the covariance matrix of the excess risky-asset returns and γ denotes the risk aversion coefficient. The investor invests $1 - \sum_i w_{i,t}$ in the risk-free asset at time t. The risk-free asset is assumed to have identical borrowing and lending rates as well as zero transaction costs. We assume the investor estimates expected returns and the covariance matrix independently. The problem has an analytical solution:

$$\hat{w}_t = \frac{\hat{\Sigma}^{-1}_{t+1|t} \hat{\mu}_{t+1|t}}{\gamma}, \tag{5.6}$$

which is known as risky-asset weights in the complete portfolio.

5.5.1 Global Minimum Variance portfolio

When $\hat{\mu}_{t+1|t}$ is proportional to an all-ones vector (i.e., all stocks are expected to earn the same rate of return) and funds are fully invested in risky assets ($\sum_i w_{i,t} = 1$), the resultant optimized portfolio is the Global Minimum Variance (GMV) portfolio whose weights are given by

$$\omega_t^{GMV} = \frac{\hat{\Sigma}^{-1}_{t+1|t} \mathbb{1}}{\mathbb{1}' \hat{\Sigma}^{-1}_{t+1|t} \mathbb{1}}. \tag{5.7}$$

Since a GMV portfolio is constructed using an estimated covariance matrix only, its OOS standard deviation is a standard measure in the literature for the quality of a proposed estimator. Table 2 presents the comparison of covariance matrix estimators introduced in Section 4 in terms of the OOS standard deviation of their corresponding GMV portfolios as well as other portfolio performance metrics.

SAM$_{\text{cov}}$: the sample covariance matrix.

EFM$_{\text{capm}}$: the covariance matrix implied by the single-index model (CAPM).

GAL$_{\text{cov}}$: the Galton covariance matrix.

POET: the covariance matrix implied by a factor model with 5 principal component factors. The covariance matrix of residuals is estimated using the POET method of Fan, Liao, and Mincheva (2013).[29]

$\mathbf{LW_{capm}}$: the LW2003 estimator, which linearly shrinks $\mathrm{SAM_{cov}}$ towards $\mathrm{EFM_{capm}}$.

$\mathbf{LW_{cc}}$: the LW2004 estimator, which linearly shrinks $\mathrm{SAM_{cov}}$ to the constant-correlation model.

$\mathbf{LW_{nl}}$: the LW2020 estimator, which nonlinearly shrinks $\mathrm{SAM_{cov}}$ through adjustment of its eigenvalues.

$\mathbf{DCC_{nl}}$: a multivariate-GARCH estimator proposed by ELW who combines the DCC-GARCH model with $\mathrm{LW_{nl}}$. We use the "average-forecasting" method of DLW to forecast the one-month-ahead covariance matrix.

With the exception of $\mathrm{EFM_{capm}}$, which is estimated using monthly returns over the past 60 months, all estimators are based on daily returns over the past 60 months.

[Insert Table 2 near here]

For comparison, we also consider two naive strategies: the value-weighted (VW) and equal-weighted (EW) portfolios. The column labeled $\hat{\sigma}$ presents the OOS standard deviation (% per year) of each strategy. As expected, all GMV portfolios outperform the two naive strategies in terms of $\hat{\sigma}$. Their outperformance over the VW strategy is statistically significant. So, estimated covariance matrices are useful to reduce risk in real time. The best performing estimator is $\mathrm{DCC_{nl}}$ with an annualized volatility of 10.18%.

An alternative test of the success of a covariance matrix estimate is its ability to forecast the risk of (optimal) portfolios (Barroso and Saxena 2022). The column labeled $\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$ presents the risk accuracy of each method, which is measured by the ratio of ex post standard deviation to its average ex ante estimate. All estimators underestimate the risk of the GMV portfolio out of sample. $\mathrm{GAL_{cov}}$ has the lowest ratio ($\frac{\sigma(r)}{\bar{\sigma}_{exp}(r)} = 1.38$), whereas $\mathrm{EFM_{capm}}$ dramatically underestimates the risk ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}} = 5.60$).

The next three columns present sum of negative weights, active share and bankruptcy rate of each portfolio, respectively. Unlike the GMV portfolio, the VW and EW portfolios do not require short sales. The GMV portfolio estimated using $\mathrm{SAM_{cov}}$ deviates most from the passive counterpart with an active share of 334.05%. Notably, no GMV portfolio goes bankrupt during the OOS period.

Despite the fact that the objective of the GMV optimization is only to minimize portfolio risk, all GMV portfolios, except for those based on $\mathrm{SAM_{cov}}$ and $\mathrm{EFM_{capm}}$, yield higher Sharpe ratios than the VW strategy. But these differences are all statistically insignificant. While GMV strategies based on $\mathrm{SAM_{cov}}$ and $\mathrm{EFM_{capm}}$

[29]We are grateful to the authors for making their Matlab code available to us.

produce slightly lower Sharpe ratios than the VW strategy, their underperformance is also statistically insignificant.

Compared to the VW portfolio, the GMV portfolio requires more frequent rebalancing. Not surprisingly, all GMV portfolios have higher turnover than the naive counterparts. Among the eight estimators, EFM_{capm} has the lowest GMV turnover. The final column displays the net-of-costs Sharpe ratio (NSR) of each strategy. All GMV portfolios underperform the VW portfolio in terms of NSR though the differences are statistically insignificant at the 5% level. The GMV strategy using GAL_{cov} produces the highest NSR of 0.53.

5.5.2 Tangency portfolios

In this section, we study the OOS performance of the tangency portfolio (hereafter TP) constructed using different combinations of estimators for expected returns and the covariance matrix. The TP consists of risky assets only. Its weights are effectively *relative* risky-asset weights in the complete portfolio (hereafter CP), which are given by

$$w_t^{tan} = \frac{\hat{\Sigma}_{t+1|t}^{-1}\hat{\mu}_{t+1|t}}{\mathbb{1}'\hat{\Sigma}_{t+1|t}^{-1}\hat{\mu}_{t+1|t}}. \tag{5.8}$$

Following DeMiguel, Garlappi, and Uppal (2009), if a TP is conditionally inefficient (i.e. the denominator is negative), we assume the investor invest -100% in the TP and 200% in the risk-free asset. The mean-variance analysis suggests that, without estimation risk, the Sharpe ratio of the TP is equal to that of any CP formed using the same set of stocks.[30] However, the two types of MVE strategies can produce very different OOS results in the presence of not only estimation risk but also transaction costs. If $\mathbb{1}'\hat{\Sigma}_{t+1|t}^{-1}\hat{\mu}_{t+1|t}$ is close to zero, Eq. (5.8) tends to generate extreme TP weights. On one hand, the CP is immune from this problem if γ is set properly (e.g., $\gamma = 5$). On the other hand, if γ is substantially less than $\mathbb{1}'\hat{\Sigma}_{t+1|t}^{-1}\hat{\mu}_{t+1|t}$, the CP will be highly leveraged (i.e., put large negative weights on the risk-free asset) resulting in considerably high OOS standard deviation and return. In Table 3 and Table 4, we report the OOS performance of the TP strategy.[31] The OOS performance of the CP strategy for $\gamma = 5$ is reported in Table A2 and Table A3.

Since the mean-variance optimization is more sensitive to estimation error in expected returns than that in covariances, our empirical applications focus on (pairwise) combinations between a set of 10 return forecasting methods considered in

[30] The TP can be regarded as the CP of an investor with $\gamma = \mathbb{1}'\hat{\Sigma}_{t+1|t}^{-1}\hat{\mu}_{t+1|t}$.

[31] DeMiguel, Garlappi, and Uppal (2009) and Barroso and Saxena (2022) focus on the OOS performance of the TP as well. Jagannathan and Ma (2003) study the TP rather than the CP in the empirical application considering the estimation of expected returns.

Table 1 and a set of 4 (of eight) covariance matrix estimators.[32] The first two covariance matrix estimators are EFM_{capm} and SAM_{cov}, both of which are commonly used in the literature and compatible with large dimensions. The other two estimators are GAL_{cov} and DCC_{nl}. As shown in Table 2, they are the best performers in terms of the realized risk of the GMV portfolio. In addition, GAL_{cov} is the best performing method in forecasting risk of the GMV portfolio.

[Insert Table 3 near here]

Panel B of Table 3 presents the OOS performance of the TP strategies implemented using EFM_{capm} in combination with different return forecasting methods. All combinations substantially underestimate the risk of the TP with $\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$ ranging between 4.42 and 17.13 and overestimate the profitability of the TP as shown in the column labeled $\frac{\hat{\mu}}{\bar{\mu}_{exp}}$, which denotes the ratio of the average realized return over the average expected return.

The three TP strategies implemented using the Galton-corrected return forecasts (GAL_{mean}, $\text{GFM}_{\text{large}}$ and GFM_{wls}) and the two using expected returns implied by the conditional latent factor model ($\text{IPCA}_{\text{K=5,large}}$ and $\text{CA2}_{\text{K=5}}$) never go bankrupt during the OOS period. The bankruptcy of other TP strategies can be attributable to the instability of $|\mathbb{1}'\hat{\Sigma}^{-1}_{t+1|t}\hat{\mu}_{t+1|t}|$. For example, the TP combining FM_{wls} with EFM_{capm} has a $|\mathbb{1}'\hat{\Sigma}^{-1}_{t+1|t}\hat{\mu}_{t+1|t}|$ of 0.13 at the end of 1990:10 resulting in a cross-sectional standard deviation of portfolio weights of 1320%. The two TP strategies implemented using the two GFM methods (together with EFM_{capm}) can achieve higher Sharpe ratios (1.05 and 1.06) than the VW strategy and their outperformance is statistically significant. The TP obtained using GAL_{mean}, $\text{IPCA}_{\text{K=5,large}}$ or $\text{CA2}_{\text{K=5}}$ (in combination with EFM_{capm}) can achieve a comparable or even higher Sharpe ratio than the VW counterpart; however, the difference is not statistically significant.

The performance of the TP strategies deteriorates in the presence of transaction costs. The two TP strategies combining EFM_{capm} with GAL_{mean} and GFM_{wls} have the lowest turnover (TO = 22.45% per month) and highest net-of-costs Sharpe ratio (NSR = 0.60), respectively. Of the 10 TP strategies, only two have higher net-of-costs Sharpe ratios than the VW strategy. Both are implemented using GFM estimators for expected returns. However, their outperformance is not statistically significant.

In Panel C, we use SAM_{cov} in place of EFM_{capm} to estimate the covariance matrix. As before, the TP optimizer displays an optimistic bias, overestimating returns and underestimating risk. Compared to the TP strategies based on EFM_{capm}, those

[32]For the sake of brevity, we do not report the OOS performance of all 80 combinations. Results for the remaining combinations are available upon request.

based on SAM_{cov} tend to have more optimistic expectations for risk. Consistent with the finding in the literature, even with daily data, plain Markowitz portfolio is dismal with a bankruptcy rate of 0.74%, a turnover of 11560.70% and a Sharpe ratio of 0.02. Of the 10 TP strategies, only three never suffer from bankruptcy. The TP strategies combining SAM_{cov} with $\text{GFM}_{\text{large}}$ and $\text{IPCA}_{\text{K=5,large}}$ outperform the VW strategy by statistically and economically significant margins with Sharpe ratios of 1.64 and 1.15, respectively. The enormous realized risk of the TP combining SAM_{cov} with GFM_{wls} is mainly caused by extreme allocations at the end of 2010:06. $|\mathbb{1}'\hat{\Sigma}^{-1}_{t+1|t}\hat{\mu}_{t+1|t}|$ then is just 0.01 leading to a large cross-sectional standard deviation of weights (16376%). The "TO" column shows that all TP strategies using SAM_{cov} have higher turnover than their EFM_{capm} counterparts. More importantly, all of them deliver lower net-of-costs Sharpe ratios than the VW strategy.

[Insert Table 4 near here]

In Table 4, we examine the TP strategies obtained using combinations of the same 10 expected return estimators with GAL_{cov} (Panel B) or DCC_{nl} (Panel C). When combined with GAL_{cov} and DCC_{nl}, $\text{GFM}_{\text{large}}$, GFM_{wls} and $\text{IPCA}_{\text{K=5,large}}$ compare favorably to others. The 6 TP strategies based on them can deliver remarkable Sharpe ratios ranging between 1.14 and 1.65 which are significantly higher than the Sharpe ratio attained by the VW strategy. So, estimation error can be overcame. Still, after accounting for transaction costs, even high-quality estimators for covariance matrices do not contribute much to the OOS performance of the TP. Only the TP constructed using $\text{GFM}_{\text{large}}$ coupled with GAL_{cov} has a higher net-of-costs Sharpe ratio than the VW strategy (0.65 vs 0.57), but that difference is statistically insignificant.

Overall, our results suggest that it is possible to improve the performance of the sample-based TP (a.k.a plain Markowitz portfolio) by using more sophisticated estimators. $\text{GFM}_{\text{large}}$ is the most robust return forecasting method in the sense that all of the four TP strategies related to it significantly outperform the VW strategy in terms of the Sharpe ratio. However, after accounting for transaction costs, no TP can outperform the VW counterpart regardless of optimization inputs.

5.5.3 Risk targeting

The previous section suggests that the instability of the denominator of the TP formula coupled with excessive optimism of the mean-variance optimizer can expose the optimized portfolio to high bankruptcy risk. Similarly, high leverage along with the excessive optimism can render the CP prone to bankruptcy and unimplementable in practice.[33] In the TP analysis, we assume the investor always executes the TP strategy regardless of how extreme the optimal allocations are. In this section, we build on Kirby and Ostdiek (2012) and Bessler and Wolff (2015), and assume the investor constrains the mean-variance optimizer by targeting the estimated variance of the VW strategy.[34] That is, we set an upper bound for the expected portfolio risk in Eq. (5.5), which is given by

$$\max_{w_t} \quad w_t'\hat{\mu}_{t+1|t} - \frac{\gamma}{2} w_t'\hat{\Sigma}_{t+1|t} w_t \tag{5.9a}$$

$$\text{subject to} \quad w_t'\hat{\Sigma}_{t+1|t} w_t \leq \sigma^2_{vw,t+1|t} = \hat{w}_{vw,t}'\hat{\Sigma}_{t+1|t}\hat{w}_{vw,t}. \tag{5.9b}$$

When the constraint is binding, the expected risk of the optimized portfolio equals the expected risk of the VW portfolio. We refer to the constrained MVE portfolio from Eq. (5.9) as the Risk-Targeted (RT) portfolio. Note that we keep $\frac{\gamma}{2} w_t'\hat{\Sigma}_{t+1|t} w_t$ in the objective function so that the CP is still the optimal choice when its estimated variance is less than the risk target.

[Insert Table 5 near here]

Table 5 presents the OOS performance of the RT portfolios constructed using combinations of 10 return forecasting methods with EFM_{capm} (Panel B) or SAM_{cov} (Panel C). We adopt a γ of 5. The average volatility of the VW portfolio estimated ex ante using EFM_{capm} and SAM_{cov} are 14.59% and 17.40%, respectively. Like the TP and CP strategies, the RT strategy tends to be overly optimistic ex ante, as shown in the columns labeled $\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$ and $\frac{\hat{\mu}}{\bar{\mu}_{exp}}$. Compared to the CP and TP counterparts, the RT strategies are less prone to bankruptcy. Of the 20 portfolios, only the one combining $\text{IPCA}_{\text{K=5,large}}$ with EFM_{capm} goes bankrupt in one month, Except for it, all RT portfolios formed using return forecasting methods that exploit stock characteristics attain higher Sharpe rations than the VW counterpart and the differences are statistically significant. However, turnover can erode their advantages in the presence of trading costs. Only three RT strategies can produce higher net-of-costs Sharpe ratios than the VW strategy, but their outperformance is economically small (max NSR = 0.68) and statistically insignificant.

[33] As shown in Table A2 and Table A3, all CP strategies deliver high turnover and most of them have non-zero bankruptcy rates. Specifically, of the 40 CP strategies, none can achieve a higher after-cost Sharpe ratio than the VW strategy and only 4 never go bankrupt, all of which use the Galton covariance matrix.

[34] We use the estimated variance of excess returns on the VW portfolio as the upper bound of portfolio risk. The risk target is computed using an estimated covariance matrix and weights of the VW portfolio at the end of month t, both of which are ex ante information. Bessler and Wolff (2015) also formulate the mean-variance investment problem using a quadratic utility function with an upper variance bound. Similarly, Kirby and Ostdiek (2012) target the estimated expected excess return of the equal-weighted portfolio when constructing their optimal constrained portfolios.

[Insert Table 6 near here]

Table 6 presents the OOS performance of the RT strategies executed using the two improved covariance matrix estimators: GAL_{cov} (Panel B) and DCC_{nl} (Panel C). The average volatility of the VW portfolio estimated ex ante using GAL_{cov} and DCC_{nl} are 16.75% and 16.06%, respectively. As before, only one (out of twenty RT portfolios) has a non-zero bankruptcy rate. Notably, all 16 RT strategies implemented using characteristic-based return forecasts outperform the VW portfolio in terms of the Sharpe ratio and their outperformance is economically large, with Sharpe ratios as high as 2.04, and statistically significant. The RT portfolios constructed using GAL_{cov} deliver reasonable volatility and have lower $\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$ than those formed using the other three covariance matrix estimators regardless of whichever mean estimator is used. However, like all strategies examined before, all 20 RT strategies in Table 6 fail to outperform the VW counterpart in a significant way after accounting for transaction costs. This contrast is meaningful from an economic perspective: before costs are considered, several strategies achieve OOS performance that seem to violate any plausible "good deal" boundaries (Cochrane and Saa-Requejo 2000), and hence appear as near-arbitrage opportunities; but none outperforms significantly a simple value-weighted benchmark after costs. Therefore, a representation of the risk-return trade-off available to investors in the market is incomplete if it fails to account for such costs.[35]

5.5.4 Transaction cost management

Given the unfavorable net-of-costs Sharpe ratios earned by the 3 MVE strategies: the TP, CP and RT strategies, we next examine whether accounting for transaction costs ex ante can improve the after-cost performance of the mean-variance optimization. We incorporate the (expected) transaction costs into the optimization problem of the RT portfolio (Eq. (5.9)), which is given by

$$\max_{w_t} \quad w_t'\hat{\mu}_{t+1|t} - \sum_i \kappa_{i,t}|w_{i,t} - \tilde{w}_{i,t-1}| - \frac{\gamma}{2} w_t'\hat{\Sigma}_{t+1|t} w_t \tag{5.10a}$$

$$\text{subject to} \quad w_t'\hat{\Sigma}_{t+1|t} w_t \leq \sigma^2_{vw,t+1|t} = \hat{w}'_{vw,t}\hat{\Sigma}_{t+1|t}\hat{w}_{vw,t}. \tag{5.10b}$$

Statistically, the turnover regularization term, $\sum_i \kappa_{i,t}|w_{i,t} - \tilde{w}_{i,t-1}|$, acts similarly as *lasso* penalization on portfolio weights with effective one-way spreads of stocks governing the regularization intensity. We refer to the resultant optimized portfolio

[35] In a very recent paper, of August 2022, Jensen, Kelly, Malamud, and Pedersen (2022) propose the concept of "implementable efficient frontier" as the after-cost frontier available to investors. Our results so far, obtained with a broad range of return and covariance forecasting methods, would suggest that feasible frontier is approximately spanned by the value-weighted portfolio.

as the transaction-cost-managed RT (TCM-RT) strategy. The main reason why we keep the risk upper bound is that it can prevent the optimized portfolio from being highly leveraged when the market is liquid (i.e., $\kappa_{i,t}$ is close to zero). In addition, the risk constraint can be interpreted as a generalized *ridge* penalization on portfolio weights.[36]

The optimization problem is a *convex* quadratically constrained quadratic program (QCQP). We follow Chen, Lezmi, Roncalli, and Xu (2019) to rewrite transaction cost of each stock as $\kappa_{i,t}\Delta w_{i,t}^{+} + \kappa_{i,t}\Delta w_{i,t}^{-}$ where $\Delta w_{i,t}^{-}$ and $\Delta w_{i,t}^{+}$ respectively denote the sale and purchase of stock i at time t. By definition, $w_{i,t} = \tilde{w}_{i,t-1} + \Delta w_{i,t}^{+} - \Delta w_{i,t}^{-}$. We rewrite Eq. (5.10) as

$$\min_{\theta_t} \quad \frac{\gamma}{2}\theta_t' B_t^* \theta_t - \theta_t' M_t \tag{5.11a}$$

$$\text{subject to} \quad \theta_t' B_t^* \theta_t \leq \sigma^2_{vw,t+1|t}, \tag{5.11b}$$

$$\Delta w_t^{+}, \Delta w_t^{-} \geq \mathbf{0}_n \tag{5.11c}$$

$$A\theta_t = \tilde{w}_{t-1} \tag{5.11d}$$

where $\theta_t = \begin{bmatrix} w_t \\ \Delta w_t^{+} \\ \Delta w_t^{-} \end{bmatrix}$, $M_t = \begin{bmatrix} \hat{\mu}_{t+1|t} \\ -\kappa_t \\ -\kappa_t \end{bmatrix}$, $A = \begin{bmatrix} \boldsymbol{I}_{n\times n} & -\boldsymbol{I}_{n\times n} & \boldsymbol{I}_{n\times n} \end{bmatrix}$ and $B_t^* = \begin{bmatrix} \hat{\Sigma}_{t+1|t} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix}$. B_t^* is a positive semi-definite matrix, and therefore the optimization problem is a *convex* QCQP with $3n$ variables, which can be solved using the interior-point method of Nesterov and Todd (1997) with great efficiency.[37] The estimated optimal portfolio choice is given by $\hat{w}_t = (\boldsymbol{I}_{n\times n}, \mathbf{0}_{n\times n}, \mathbf{0}_{n\times n})\hat{\theta}_t$. Since

[36] Table A4 and Table A5 report the OOS performance of transaction-cost-managed (TCM) portfolios without targeting the estimated expected risk of the VW portfolio in optimization. They show that transaction cost management alone is not adequate to prevent the mean-variance optimizer from generating highly leveraged portfolios though most of the TCM portfolios do have lower turnover than their complete counterparts. The undesirable performance of the TCM strategies can be partially attributable to the improvement in market liquidity, which in turns results in tight effective spreads of the large-cap stocks. With $\kappa_{i,t}$ being closer to zero, the effects of turnover regularization on weights become weaker, and thus the optimizer tends to deliver portfolios like the CP. Note that our investment problem is slightly different from that of Hautsch and Voigt (2019). They implement transaction-cost-managed mean-variance strategies using risky assets only. Instead of using a budget constraint (BC), we use a variance constraint so that our optimized portfolios tend to be less volatile and more comparable to the passive strategy. See Table A6 and Table A7 for the OOS performance of the TCM portfolios constructed using risky assets only (TCM-BC).

[37] Note that any convex QCQP can be reformulated as a second-order cone program (a.k.a conic quadratic optimization). We solve the problem using the academic version of MOSEK (2022) via Python package CVXPY of Diamond and Boyd (2016) and Agrawal, Verschueren, Diamond, and Boyd (2018). CVXPY can formulate Eq. (5.10) in conic form and send the resulting conic problem to MOSEK (2022). Another feature of CVXPY is that it can check convexity and feasibility of an optimization problem.

the optimization problem is convex, the solution is globally optimal.

[Insert Table 7 near here]

Table 7 shows the OOS performance of the 20 TCM-RT portfolios constructed using combinations of 10 return forecasting methods with EFM_{capm} (Panel B) or SAM_{cov} (Panel C). Notably, compared to the CP strategies obtained using the same inputs (Table A2), the TCM-RT portfolios deliver much lower realized risk and turnover owing to the imposition of turnover penalization and the variance upper bound. No TCM-RT portfolio goes bankrupt during the OOS period. More importantly, the VW portfolio cannot significantly outperform any TCM-RT strategies, except for that estimated using SAM_{mean} and SAM_{cov}, in terms of before- and after-cost Sharpe ratios. Of the 20 TCM-RT strategies, 17 can produce higher after-cost Sharpe ratios than the VW portfolio. In particular, the 4 strategies constructed using combinations of SAM_{cov} with expected return estimators based on Fama-MacBeth regressions can achieve net-of-costs Sharpe ratios that are not only economically meaningful (NSR $= 1.16, 1.13, 1.12$ and 1.06) but also significantly different from the Sharpe ratio achieved by the VW strategy (NSR $= 0.57$).

It is worth mentioning that with the reduction in effective costs incurred for large-cap stocks, turnover regularization exerts weaker effects on portfolio weights. As a consequence, as shown in the columns labeled TO_{first} and $\text{TO}_{\text{second}}$, TCM-RT portfolios have much higher turnover in the second half of the OOS period during which the market is more liquid.

[Insert Table 8 near here]

Table 8 presents the OOS performance of the 20 TCM-RT strategies obtained using combinations of 10 return forecasting methods with two improved covariance matrix estimators: GAL_{cov} (Panel B) and DCC_{nl} (Panel C). Of 20 portfolios, only the hybrid of GAL_{mean} with DCC_{nl} goes bankrupt in one month. With the exception of the two portfolios based on SAM_{mean}, all TCM-RT portfolios can produce impressive Sharpe ratios ranging from 0.63 to 1.77, some of which are significantly different from the Sharpe ratio attained by the VW strategy. Moreover, the differences in Sharpe ratios between the VW strategy and the two portfolios constructed using SAM_{mean} are statistically insignificant. The final column shows that after accounting for transaction costs, only the 4 TCM-RT strategies that combine GAL_{cov} with Fama-MacBeth-related estimators (FM_{large}, FM_{wls}, $\text{GFM}_{\text{large}}$ and GFM_{wls}) can significantly outperform the VW portfolio (Net Sharpe Ratio $= 1.16, 1.13, 1.12$ and 1.06). Furthermore, in comparison to the best performing TCM-RT strategies in Table 7, the four strategies do less short selling and are therefore less susceptible to the related costs.

5.5.5 Short-sale constraints

Most investors have small, if any, short selling positions. On top of that, the short legs of some well-known anomalies have very high short selling costs (Drechsler and Drechsler 2018). These costs are hard to quantify for representative historical samples, such as ours, and therefore typically not accounted for in the portfolio optimization literature. Due to this, we consider a conservative case in which mean-variance investors are short-sale-constrained and ask if they can still benefit from the TCM-RT strategy implemented with improved inputs. That is, we add the constraint, $w_{i,t} \geq 0$, into Eq. (5.10).

[Insert Table 9 near here]

Table 9 documents the OOS performance of the TCM-RT long-only portfolios constructed using combinations of 10 return forecasting methods with EFM_{capm} (Panel B) or SAM_{cov} (Panel C).[38] Not surprisingly, prohibiting short selling leads to substantial reduction in turnover. But it also results in reduction in profitability for most of the combinations. That is, of the 20 TCM-RT long-only portfolios, 17 perform worse than the unconstrained counterparts in terms of Sharpe ratios; more importantly, none can significantly outperforms the VW portfolio in terms of the net-of-costs Sharpe ratio. It is interesting to note that the imposition of short-sale constraints can considerably improve the performance of the unconstrained TCM-RT strategy using sample estimates with the Sharpe ratio increasing from 0.06 to 0.54. In addition, the short-sale-constrained TCM-RT portfolios constructed using the hybrid of SAM_{mean} with EFM_{capm} and the hybrid of GAL_{mean} with SAM_{cov} deliver higher Sharpe ratios than their TCM-RT counterparts. Overall, the evidence is not supportive of the TCM-RT optimization for investors with short-sale constraints. This does not exclude the possibility of investors gaining from portfolio optimization, OOS and after transaction costs, as they *might* be able to short-sell at low costs. But it shows that if we remove the possibility of short selling (for free) such gains become much subdued.

5.5.6 Smaller investment universe

In Table A10 and Table A11, we use a smaller investment universe that consists of the 50 largest common stocks by market equity. The two tables suggest that no TCM-RT portfolio can outperform the VW portfolio in a statistically significant way though some of them can produce net-of-costs Sharpe ratios greater than 0.7.

[38] Since optimized portfolios based on GAL_{cov} and DCC_{nl} deliver similar performance to those based on EFM_{capm} and SAM_{cov}, for the sake of brevity, their results are reported in Panels B and C of Table A8, respectively.

Interestingly, the TCM-RT strategies obtained using the Galton covariance matrix consistently overestimate the OOS risk, whereas the others still underestimate that.

5.5.7 Kirby and Ostdiek (2012)'s low-turnover strategy

Instead of directly managing transaction costs, Kirby and Ostdiek (2012) propose a Reward-to-Risk Timing (RRT) strategy featuring no short sale, low turnover and simple construction procedure. The strategy refrains from estimating covariances and assumes investors with a strong prior belief that all stocks will earn non-negative excess returns. The estimated optimal weight for stock i at the end of month t is given by

$$\hat{w}_{i,t} = \frac{(\hat{\mu}^{+}_{i,t+1|t}/\hat{\sigma}_{i,t+1|t})}{\sum_i(\hat{\mu}^{+}_{i,t+1|t}/\hat{\sigma}_{i,t+1|t})}, \tag{5.12}$$

where $\hat{\mu}^{+}_{i,t+1|t} = \max(\hat{\mu}_{i,t+1|t}, 0)$, $\hat{\mu}_{i,t+1|t}$ denotes an estimate of the expected return, $\hat{\sigma}_{i,t+1|t}$ denotes an estimate of the variance. The RRT strategy is effectively a tangency portfolio consisting of stocks with positive estimated expected returns and a diagonal covariance matrix.

[Insert Table 10 near here]

Table 10 presents the OOS performance of the RRT portfolios constructed using combinations of 10 return forecasting methods with the estimated variances extracted from the main diagonal of EFM_{capm} (Panel B) or SAM_{cov} (Panel C). The two estimators for variances are effectively the sample variances computed from monthly or daily returns, respectively.[39] As expected, all RRT strategies exhibit low turnover and perform better than the value-weighted strategy in terms of the Sharpe ratio. In particular, there are 14 strategies (out of 20) whose outperformance over the VW strategy is statistically significant. Economically, all RRT strategies outperform the VW strategy in terms of the net-of-costs Sharpe ratios; however, their outperformance is statistically insignificant. Table 10 also indicates that investors cannot derive much benefit from using daily sample variances in place of monthly counterparts to implement the RRT strategy. Moreover, monthly sample variances tend to perform as well as their Galton and GARCH counterparts when used to construct the RRT portfolio (see Table A9 for results about GAL_{cov} and DCC_{nl}).

[39] Note that the variance estimated by the single index model is slightly different from the sample variance due to the minor difference in degrees of freedom.

5.6 Parametric Portfolio Policies (PPP)

Another popular method to simultaneously exploit multiple stock characteristics in portfolio optimization is Parametric Portfolio Policies (PPP) developed by Brandt, Santa-Clara, and Valkanov (2009, BSV). Different from the MVE strategies which are implemented using expected returns estimated from characteristic data as well as an estimated covariance matrix, the PPP portfolio is constructed by directly parameterizing portfolio weights as a function of stock characteristics, which is given by

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t}\theta^T \hat{z}_{i,t}, \tag{5.13}$$

where $\bar{w}_{i,t}$ is the weight of stock i in a benchmark portfolio such as the value-weighted portfolio at the end of month t, θ denotes a vector of coefficients that needs to be estimated, and $\hat{z}_{i,t}$ denotes a vector of the latest known values of normalized characteristics with a mean of zero for stock i at the end of month t. The investor's problem is to find an optimal vector of weights that can maximize the conditional expected utility of the portfolio's return ($r_{p,t+1}$) net of trading costs, that is,

$$\max_{w_t} E_t\big[U(r_{p,t+1})\big] = E_t\left[U\left(\sum_{i=1}^{N_t} w_{i,t}r_{i,t+1} - \kappa_{i,t}|w_{i,t} - \tilde{w}_{i,t-1}|\right)\right]. \tag{5.14}$$

The coefficients (θ) are assumed to be fixed across assets and over time so that Eq. (5.14) can be rewritten as an unconditional optimization problem whose sample analog is given by

$$\max_{\theta} \frac{1}{T}\sum_{t=0}^{T-1}\left[U\left(\sum_{i=1}^{N_t}(\bar{w}_{i,t} + \frac{1}{N_t}\theta^T \hat{z}_{i,t})r_{i,t+1} - \kappa_{i,t}|w_{i,t} - \tilde{w}_{i,t-1}|\right)\right]. \tag{5.15}$$

When effective spreads of all stocks are zeroes, it corresponds to the vanilla version of the PPP method. Consistent with BSV, we assume an investor with a CRRA preference and a risk aversion coefficient of 5. The estimation sample consists of the 500 largest stocks by market equity with non-missing characteristics. We fit the model yearly at the end of June using an expanding window. Specifically, we use the data up to June 1986 to get the first set of the estimated coefficients which is in turn used to construct OOS PPP portfolios over the holding period 1986:07 to 1987:06 along with monthly lagged characteristics. At the end of June in the subsequent year, we re-estimate the coefficients using the enlarging estimation sample. The investment universe is the same as that used in other OOS applications. To further mitigate overfitting as a result of a large predictor set, DeMiguel et al. (2020) use a *lasso* constraint in the PPP optimization, which is given by

$$\|\theta\|_1 \leq \delta,$$

where δ is a hyperparameter that governs the intensity of coefficient regularization. When $\delta = 0$, the resulting PPP portfolio is the benchmark portfolio; when δ is large enough, the constraint is inactive, and therefore the unconstrained PPP is recovered. Following DeMiguel et al. (2020), we use a 5-fold cross-validation to select optimal δ from 32 candidate values that are evenly spaced over an interval between 0 and 500 and use a rolling window of 100 months to estimate the coefficients.

The proportional transaction cost parameter used in BSV and DeMiguel et al. (2020) is a function of time and firm size, which is given by

$$\kappa^*_{i,t} = y_t z_{i,t}, \tag{5.16}$$

where y_t is assumed to decrease linearly from a value of 5.02 in July 1964 to 1.0 in January 2002 and remain unchanged afterwards.[40] $z_{i,t} = 0.006 - 0.0025 \times lme_{i,t}$, where $lme_{i,t}$ is the ranked-normalized market capitalization of stock i in month t ranging from zero to one. This function implies that the proportional trading costs for the largest stock are approximately 156 basis points in 1987:01 and remain constant at 35 basis points after 2002.

[Insert Figure 2 near here]

Fig. 2 depicts the dynamics of proportional transaction-cost parameters estimated by BSV and those estimated by us. It shows that the BSV approach tends to assume higher average costs than ours. This should erode more before-cost gains and also penalize marginal turnover more heavily in the model estimation.

[Insert Table 11 near here]

Table 11 presents the OOS performance of portfolio policies. We use the VW portfolio as the benchmark strategy for all policies. PPP3 denotes policies with three characteristics: value, momentum and size characteristics, which are the same as in the original study. PPP62 denotes policies exploiting the full set of 62 characteristics listed in Table A1. $\text{PPP62}_{\text{lasso}}$ denotes policies that exploits the 62 characteristics and are regularized by the *lasso* constraint. We consider three versions of the PPP optimization: PPP without transaction cost management (Panel B),

[40] It is based on the assumption of BSV that y_t in January 1974 are four times larger than in January 2002.

transaction-cost-managed PPP built using our effective-spread estimates (Panel C) and transaction-cost-managed PPP built using effective-spread estimates of BSV (Panel D).

Panel B shows that PPP3 that is not optimized with trading costs yields lower before- and after-cost Sharpe ratios than the benchmark strategy. Transaction-cost-unmanaged PPP62 produces a very unfavorable net-of-costs Sharpe ratio of −0.46. Panel C suggests that performance gains as a result of transaction cost management for PPP3 are marginal, whereas those for PPP62 are impressive. Transaction-cost-managed PPP62 has much less turnover than its unmanaged counterpart (8120.60% vs 703.75%) and outperforms the VW strategy in terms of before- and after-cost Sharpe ratios though the differences are statistically insignificant. Despite further reducing the turnover of PPP62, the *lasso* constraint can hurt the profitability of transaction-cost-managed PPP62. The before- and after-cost Sharpe ratios delivered by $\text{PPP62}_{\text{lasso}}$ are just 0.47 and 0.27, respectively.

Panel D reports the OOS performance of transaction-cost-managed policies estimated using the trading-cost estimator of Brandt, Santa-Clara, and Valkanov (2009). With a stronger turnover regularization, transaction-cost-managed PPP62 outperforms the VW portfolio in terms of both before-cost and after-cost Sharpe ratios. The before-cost outperformance is statistically significant with a p-value of just 5% (SR = 1.00 and p-value = 0.05). Though economically meaningful, the after-cost Sharpe ratio of 0.84 is not significantly different from that of the VW portfolio (p-value = 0.21). For the other two portfolio policies, PPP3 is still comparable to the passive counterpart though the policy has short positions in stocks, while $\text{PPP62}_{\text{lasso}}$ remains dismal. Overall, for the universe of large-cap stocks, no portfolio policy can outperform the VW strategy in a statistically significant way, which is consistent with the findings of DeMiguel et al. (2020).[41]

6 Conclusion

It is well-documented that the sample-based MVE portfolio delivers undesirable OOS performance. We examine whether improvements in estimation of expected returns and the covariance matrix can translate into outperformance of the MVE strategies over the value-weighted portfolio, especially after accounting for transaction costs. We use an OOS period from 1987:01 to 2020:12 and a relatively large investment universe consisting of the 500 largest stocks by market capitalization.

[41] In table IA.30, they document that the regularized parametric portfolio policy with 51 characteristics constructed using large stocks (i.e., stocks in the top size quintile) fails to outperform the passive strategy. Also, the large-stock policy based on the size, book-to-market, and momentum characteristics and the large-stock policy based on the size, book-to-market, asset growth, and gross profitability characteristics yield lower net-of-costs Sharpe ratios than the passive strategy.

Besides conventional return forecasting methods based on Fama-MacBeth regressions and the recently proposed methods based on machine learning techniques, we propose two shrinkage Fama-MacBeth estimators which apply the Galton correction of Barroso and Saxena (2022) to ordinary Fama-MacBeth estimates. Like other alternative mean estimators, they can achieve much higher OOS predictive R^2 than the sample mean. Our empirical applications focus on four covariance matrix estimators, all of which can handle large dimensions. The first two are respectively the covariance matrix implied by the single-index model and the daily sample covariance matrix, both of which are widely used and can be easily obtained. The other two are the best performing methods among the 8 prominent covariance matrix estimators we examine in terms of the realized risk of the GMV portfolio.

We find that of the 80 TP strategies we examine, only 10 can deliver significantly higher Sharpe ratios than the VW portfolio and that only 19 never go bankrupt during the OOS period. More importantly, no tangency portfolio can outperform the VW counterpart after accounting for trading costs. Risk targeting can substantially reduce the possibility of bankruptcy and improve the before-cost performance of the MVE strategy. Of the 80 RT strategies, only two have non-zero bankruptcy rate and 30 can outperform the VW strategy by economically and statistically significant margins. However, no RT strategy can outperform the passive counterpart in a statistically significant way after accounting for trading costs.

The after-cost performance of the RT strategy can be substantially improved by incorporating transaction costs into the optimization problem. The 4 TCM-RT strategies implemented using combinations of a Fama-MacBeth-related mean estimator with the (daily) sample or Galton covariance matrix can deliver impressive after-cost Sharpe ratios, which are significantly different from the Sharpe ratio earned by the VW portfolio. However, in the presence of short-sale constraints, the TCM-RT strategy does not exhibit any advantages over the buy-and-hold counterpart, at least for the set large-cap U.S. stocks.

One limitation of out study is that we do not account for model uncertainty in portfolio selection. In our OOS exercises, we just assume that investors use optimization inputs obtained from the same estimators over the entire OOS period. Additionally, we restrict our investment universe to large-cap stocks and provide evidence unsupportive of complex methods for estimating expected returns and the covariance matrix. We do not rule out the possibility that the mean-variance optimization for small-cap stocks or in larger sets can benefit from advanced forecasting methods. However, estimating and managing transaction costs for small-cap stocks can be even more challenging than for large caps due to price impacts and other frictions not accounted for in our study.

Figure 1
Estimates of effective half-spreads for stocks used in OOS exercises – 1987:01-2020:12
The plot depicts half of effective spreads for the top-500 stocks by market capitalization over the OOS period 1987:01 to 2020:12. The purple (blue) line represents the maximum (minimum) spreads. The orange line represents the mean spreads.

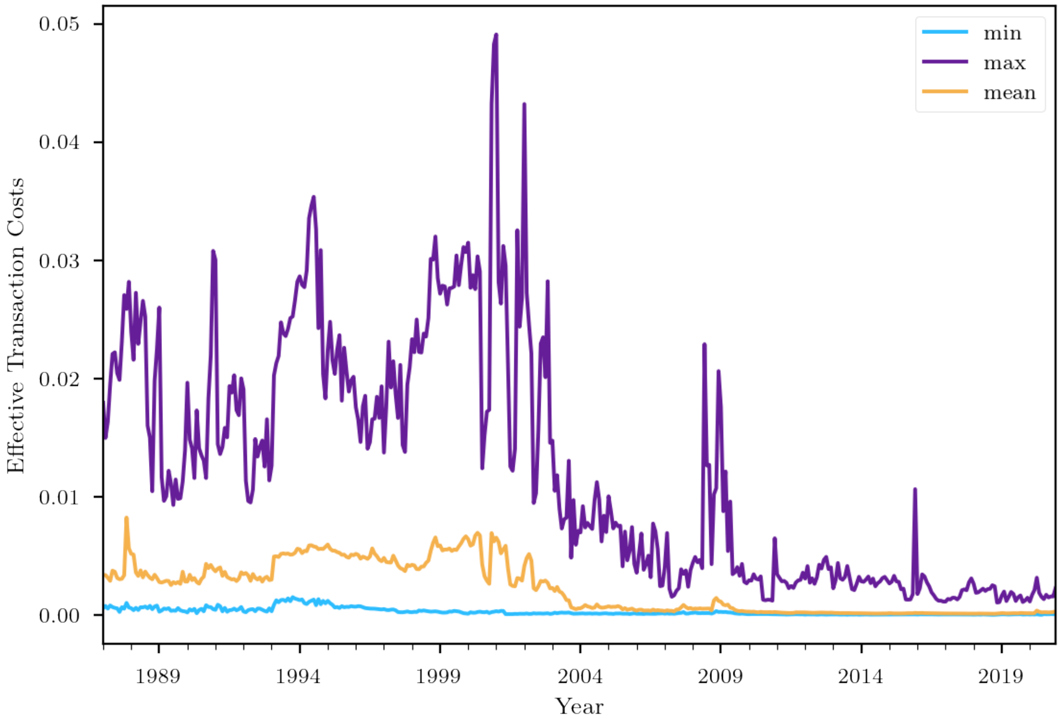

Figure 2
Comparison of transaction cost estimators for the estimation sample of PPP
The plot depicts monthly average transaction costs of stocks included in the estimation sample of PPP (i.e. the 500 largest stocks by market equity with non-missing 62 characteristics) over the 1964:07 to 2020:12 period. The blue line depicts dynamics of (ex ante) transaction cost estimates used in our main analysis, the purple line depicts dynamics of transaction cost estimates used in Brandt, Santa-Clara, and Valkanov (2009).

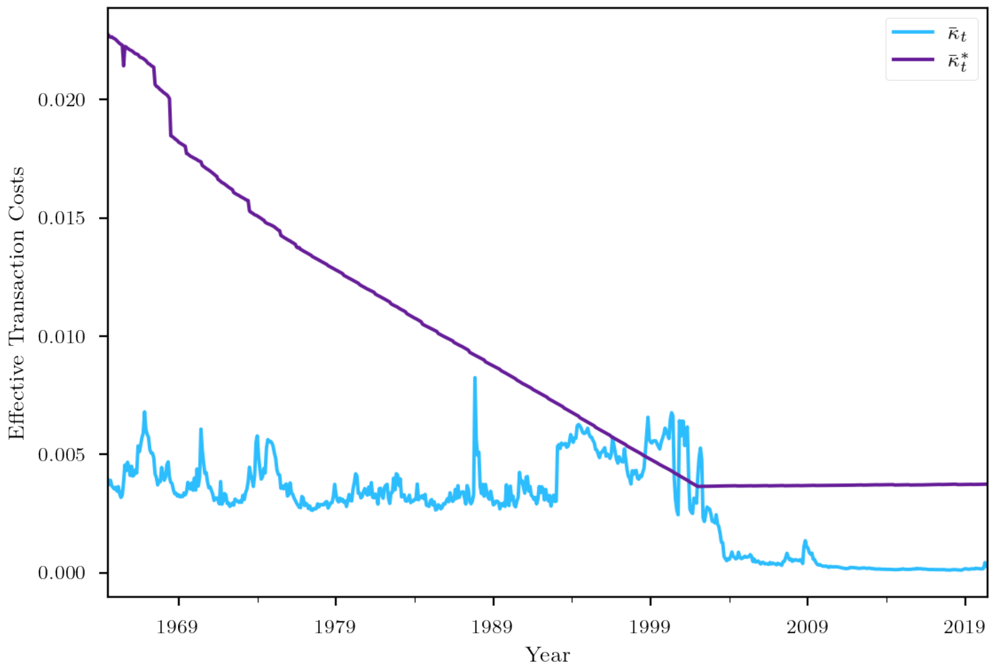

Table 1
OOS R^2(%) of expected return estimators

The table reports out-of-sample (OOS) R^2 of various expected return forecasting methods for the investment universe used in our OOS exercises over the 1987:01 to 2020:12 period. Besides OOS R^2, the table presents the corresponding paper of each method and the model type.

	Paper	Model Type	R^2_{oos}
	Return-based Estimators		
$\mathrm{SAM}_{\mathrm{mean}}$	Markowitz (1952)	Sample Mean	-1.75
$\mathrm{GAL}_{\mathrm{mean}}$	Barroso and Saxena (2022)	Sample Mean with Galton Correction	0.50
	Fama-MacBeth Regressions without/with Galton Correction		
$\mathrm{FM}_{\mathrm{large}}$	Fama and MacBeth (1973)	OLS for Large Stocks	0.48
$\mathrm{FM}_{\mathrm{wls}}$	Fama and MacBeth (1973)	WLS for All Stocks	0.50
$\mathrm{GFM}_{\mathrm{large}}$		OLS for Large Stocks with Galton Correction	0.61
$\mathrm{GFM}_{\mathrm{wls}}$		WLS for All Stocks with Galton Correction	0.64
	Machine Learning		
$\mathrm{AGLASSO}_{\mathrm{large}}$	Freyberger, Neuhierl, and Weber (2020)	Additive Model for Large Stocks with Adaptive Group Lasso	0.25
$\mathrm{IPCA}_{\mathrm{K=5,large}}$	Kelly, Pruitt, and Su (2019)	Linear Conditional Latent Factor Model for Large Stocks	0.58
$\mathrm{CA2}_{\mathrm{K=5}}$	Gu, Kelly, and Xiu (2021)	Conditional Autoencoder	0.61
NN3	Gu, Kelly, and Xiu (2020)	Neutral Network	0.66

Table 2

OOS performance of the GMV portfolio

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. We construct the global minimum variance (GMV) portfolio using various covariance matrix estimators: the daily sample covariance matrix (SAM), the covariance matrix implied by the single index model of Sharpe (1963, EFM_{capm}), the Galton covariance matrix of Barroso and Saxena (2022, GAL_{cov}), the covariance matrix estimated by the Principal Orthogonal ComplEment Thresholding method of Fan, Liao, and Mincheva (2013, POET), two linear shrinkage estimators proposed by Ledoit and Wolf (2003, 2004, LW_{capm} and LW_{cc}), the nonlinear shrinkage estimator of Ledoit and Wolf (2020, LW_{nl}), and the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019, DCC_{nl}). All portfolios are rebalanced monthly. VW and EW respectively denote the value-weighted and equal-weighted portfolios. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\overline{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal variances, equal Sharpe ratios, or equal net-of-costs Sharpe ratios between a GMV portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\sigma}}{\overline{\sigma}_{exp}}$	SNW	AS	BR	SR	TO	NSR
VW	8.59	14.99		0.00	0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41		0.00	45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
EFM_{capm}	6.84	12.27 (0.00)	5.60	-42.07	100.40	0.00	0.56 (0.94)	21.47	0.50 (0.75)
SAM	6.10	11.32 (0.00)	2.32	-280.41	334.05	0.00	0.54 (0.86)	158.20	0.24 (0.08)
GAL	6.82	10.41 (0.00)	1.38	-128.03	186.51	0.00	0.66 (0.65)	57.04	0.53 (0.81)
POET	6.61	10.53 (0.00)	2.11	-126.64	186.20	0.00	0.63 (0.76)	48.41	0.52 (0.78)
LW_{capm}	6.66	10.67 (0.00)	1.98	-176.88	233.82	0.00	0.62 (0.78)	84.31	0.46 (0.53)
LW_{cc}	6.65	11.14 (0.00)	1.94	-201.50	257.64	0.00	0.60 (0.90)	98.44	0.41 (0.42)
LW_{nl}	6.49	10.55 (0.00)	1.79	-155.71	210.97	0.00	0.61 (0.81)	74.17	0.46 (0.53)
DCC_{nl}	8.79	10.18 (0.00)	2.46	-131.89	192.37	0.00	0.86 (0.09)	240.58	0.38 (0.25)

Table 3

OOS performance of tangency portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of tangency portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a tangency portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	AS	BR	SR	TO	NSR
Panel A: Benchmark Portfolios										
VW	8.59	14.99			0.00	0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Covariance Matrix implied by the Single Index Model										
Return-based Estimators										
SAM_{mean}	-18.46	114.29	-0.20	12.41	-293.19	343.00	0.74	-0.16 (0.01)	315.18	-0.23 (0.00)
GAL_{mean}	6.80	11.87	0.56	5.34	-43.14	101.10	0.00	0.57 (1.00)	22.45	0.52 (0.79)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	10.07	182.17	0.12	17.13	-398.62	453.64	0.25	0.06 (0.35)	1071.02	-0.11 (0.05)
FM_{wls}	57.90	253.52	0.31	9.92	-1047.49	1099.09	0.49	0.23 (0.14)	2636.41	-0.27 (0.00)
GFM_{large}	14.11	13.48	0.71	4.56	-75.02	133.93	0.00	1.05 (0.03)	205.35	0.58 (0.97)
GFM_{wls}	14.20	13.38	0.85	4.42	-77.32	135.36	0.00	1.06 (0.03)	213.57	0.60 (0.88)
Machine Learning										
$AGLASSO_{large}$	9.84	98.16	0.14	11.38	-279.81	334.32	0.49	0.10 (0.17)	616.42	-0.10 (0.01)
$IPCA_{K=5,large}$	11.84	14.46	0.63	4.86	-77.96	135.13	0.00	0.82 (0.25)	131.44	0.48 (0.69)
$CA2_{K=5}$	33.93	52.00	0.49	4.85	-398.69	452.61	0.00	0.65 (0.72)	2792.56	-0.20 (0.00)
NN3	10.73	40.94	0.32	8.15	-157.52	214.15	0.25	0.26 (0.46)	432.62	-0.10 (0.01)
Panel C: Sample Covariance Matrix										
Return-based Estimators										
SAM_{mean}	2.96	186.66	0.01	4.78	-2746.89	2794.16	0.74	0.02 (0.04)	11560.70	-0.36 (0.00)
GAL_{mean}	5.08	12.02	0.39	2.34	-297.02	349.82	0.00	0.42 (0.46)	176.91	0.10 (0.02)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	112.10	205.16	0.30	4.55	-3290.18	3339.12	0.98	0.55 (0.91)	9176.28	-0.26 (0.00)
FM_{wls}	161.80	469.38	0.34	7.43	-4522.49	4569.11	1.23	0.34 (0.32)	11194.57	-0.14 (0.00)
GFM_{large}	40.23	24.55	0.64	1.93	-878.91	931.94	0.00	1.64 (0.00)	2120.49	0.16 (0.04)
GFM_{wls}	-503.77	3177.84	-1.24	30.21	-7996.05	8048.30	0.25	-0.16 (0.38)	16354.89	-0.17 (0.01)
Machine Learning										
$AGLASSO_{large}$	26.78	367.76	0.04	5.80	-4388.25	4434.84	1.23	0.07 (0.12)	10849.39	-0.46 (0.00)
$IPCA_{K=5,large}$	22.39	19.41	0.57	2.11	-627.25	678.56	0.00	1.15 (0.01)	929.17	0.04 (0.01)
$CA2_{K=5}$	88.16	180.77	0.27	3.76	-3504.69	3554.14	0.98	0.49 (0.79)	9771.79	-0.35 (0.00)
NN3	60.59	76.01	0.28	2.44	-2410.66	2461.17	0.74	0.80 (0.40)	6267.14	-0.36 (0.00)

Table 4

OOS performance of tangency portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of tangency portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	AS	BR	SR	TO	NSR
Panel A: Benchmark Portfolios										
VW	8.59	14.99			0.00	0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Galton Covariance Matrix										
Return-based Estimators										
SAM_{mean}	-348.10	1949.24	-0.61	20.25	-2687.18	2729.63	3.19	−0.18 (0.01)	3826.52	−0.20 (0.00)
GAL_{mean}	5.86	10.95	0.45	1.37	-139.33	196.18	0.00	0.54 (0.87)	72.55	0.38 (0.37)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	136.65	192.97	0.47	3.25	-1815.76	1867.26	0.25	0.71 (0.54)	5111.25	0.36 (0.28)
FM_{wls}	143.71	255.74	0.49	4.12	-1891.33	1940.40	0.49	0.56 (0.97)	4961.81	0.16 (0.06)
$\text{GFM}_{\text{large}}$	38.32	23.23	0.67	1.26	-515.33	571.36	0.00	1.65 (0.00)	1325.68	0.65 (0.66)
GFM_{wls}	38.39	30.91	0.66	1.42	-611.14	666.08	0.00	1.24 (0.01)	1598.93	0.49 (0.69)
Machine Learning										
$\text{AGLASSO}_{\text{large}}$	252.88	999.59	0.43	10.40	-2628.89	2679.28	0.49	0.25 (0.34)	6186.66	0.06 (0.14)
$\text{IPCA}_{\text{K=5,large}}$	20.86	17.73	0.57	1.32	-355.65	408.70	0.00	1.18 (0.01)	585.61	0.41 (0.45)
$\text{CA2}_{\text{K=5}}$	417.68	1469.25	0.60	8.92	-4685.24	4737.97	0.00	0.28 (0.28)	10723.97	0.20 (0.07)
NN3	56.50	62.78	0.38	1.72	-1097.31	1151.42	0.25	0.90 (0.24)	2777.50	0.11 (0.04)
Panel C: DCC-GARCH with Nonlinear Shrinkage										
Return-based Estimators										
SAM_{mean}	3.27	35.61	0.02	2.05	-718.72	767.31	0.00	0.09 (0.04)	858.12	−0.45 (0.00)
GAL_{mean}	8.39	10.05	0.66	2.35	-138.10	197.35	0.00	0.83 (0.13)	247.54	0.33 (0.16)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	78.84	160.34	0.53	7.14	-1090.74	1145.04	0.25	0.49 (0.72)	2857.79	0.24 (0.10)
FM_{wls}	354.52	1822.47	0.66	21.92	-3686.11	3737.50	0.49	0.19 (0.23)	8094.87	0.15 (0.14)
$\text{GFM}_{\text{large}}$	20.17	13.38	0.65	1.91	-280.56	339.59	0.00	1.51 (0.00)	685.26	0.46 (0.52)
GFM_{wls}	19.25	14.26	0.69	1.93	-301.41	359.56	0.00	1.35 (0.00)	747.40	0.38 (0.29)
Machine Learning										
$\text{AGLASSO}_{\text{large}}$	293.14	1076.32	0.67	22.96	-2292.71	2346.84	0.00	0.27 (0.19)	5141.74	0.22 (0.16)
$\text{IPCA}_{\text{K=5,large}}$	16.21	14.22	0.58	2.16	-251.63	308.58	0.00	1.14 (0.01)	441.79	0.33 (0.18)
$\text{CA2}_{\text{K=5}}$	81.50	104.78	0.47	3.26	-1495.14	1548.02	0.25	0.78 (0.34)	5581.28	−0.45 (0.00)
NN3	141.09	653.06	0.75	18.61	-1585.69	1640.39	0.25	0.22 (0.20)	3597.65	0.09 (0.10)

Table 5

OOS performance of RT portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of RT portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between an optimized portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	NSR
Panel A: Benchmark Portfolios									
VW	8.59	14.99			0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Covariance Matrix implied by the Single Index Model									
Return-based Estimators									
SAM_{mean}	24.47	61.38	0.16	4.21	-433.37	0.00	0.40 (0.46)	318.82	0.24 (0.16)
GAL_{mean}	50.01	76.76	0.62	5.26	-278.03	0.00	0.65 (0.71)	279.85	0.55 (0.91)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	65.56	54.62	0.51	3.74	-441.75	0.00	1.20 (0.01)	1300.62	0.56 (0.98)
FM_{wls}	62.88	52.68	0.53	3.61	-454.99	0.00	1.19 (0.01)	1320.37	0.53 (0.84)
$\text{GFM}_{\text{large}}$	70.81	67.24	0.72	4.61	-350.84	0.00	1.05 (0.03)	972.09	0.65 (0.73)
GFM_{wls}	69.37	66.16	0.84	4.53	-355.84	0.00	1.05 (0.03)	987.96	0.64 (0.74)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	61.91	55.36	0.56	3.79	-415.92	0.00	1.12 (0.01)	798.47	0.68 (0.63)
$\text{IPCA}_{\text{K=5,large}}$	59.50	74.06	0.67	5.08	-345.51	0.25	0.80 (0.31)	584.65	0.55 (0.93)
$\text{CA2}_{\text{K=5}}$	62.70	49.46	0.62	3.39	-436.18	0.00	1.27 (0.00)	1322.45	0.52 (0.81)
NN3	58.00	54.84	0.57	3.76	-394.44	0.00	1.06 (0.02)	1070.85	0.51 (0.78)
Panel C: Sample Covariance Matrix									
Return-based Estimators									
SAM_{mean}	-2.42	46.76	-0.02	2.69	-1173.18	0.00	-0.05 (0.02)	982.39	-0.46 (0.00)
GAL_{mean}	18.14	40.32	0.41	2.32	-1000.69	0.00	0.45 (0.56)	647.63	0.11 (0.03)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	66.05	31.62	0.46	1.82	-1259.88	0.00	2.09 (0.00)	3045.99	0.19 (0.07)
FM_{wls}	63.11	31.99	0.46	1.84	-1257.57	0.00	1.97 (0.00)	3043.73	0.10 (0.03)
$\text{GFM}_{\text{large}}$	64.58	32.62	0.73	1.87	-1198.18	0.00	1.98 (0.00)	2772.50	0.32 (0.21)
GFM_{wls}	61.31	33.07	0.80	1.90	-1198.37	0.00	1.85 (0.00)	2774.68	0.25 (0.11)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	49.01	33.18	0.39	1.91	-1260.29	0.00	1.48 (0.00)	2114.99	-0.04 (0.01)
$\text{IPCA}_{\text{K=5,large}}$	36.59	35.06	0.52	2.01	-1165.97	0.00	1.04 (0.04)	1664.49	0.05 (0.02)
$\text{CA2}_{\text{K=5}}$	52.52	30.27	0.43	1.74	-1264.54	0.00	1.74 (0.00)	3213.74	-0.39 (0.00)
NN3	50.48	30.89	0.43	1.77	-1293.82	0.00	1.63 (0.00)	3080.28	-0.28 (0.00)

Table 6

OOS performance of RT portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of RT portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between an optimized portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	NSR
Panel A: Benchmark Portfolios									
VW	8.59	14.99			0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	0.00	0.58	7.58	0.56
							(0.96)		(0.84)
Panel B: Galton Covariance Matrix									
Return-based Estimators									
SAM_{mean}	-0.58	30.44	-0.01	1.82	-456.90	0.00	-0.02	301.31	-0.25
							(0.03)		(0.00)
GAL_{mean}	12.38	23.12	0.44	1.39	-289.60	0.00	0.54	154.22	0.37
							(0.87)		(0.34)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	40.48	19.84	0.46	1.18	-513.92	0.00	2.04	1315.97	0.56
							(0.00)		(0.95)
FM_{wls}	39.09	19.59	0.46	1.17	-515.66	0.00	1.99	1312.77	0.50
							(0.00)		(0.72)
$\text{GFM}_{\text{large}}$	39.64	20.93	0.72	1.25	-465.56	0.00	1.89	1185.26	0.65
							(0.00)		(0.66)
GFM_{wls}	38.00	20.85	0.80	1.24	-467.45	0.00	1.82	1183.03	0.60
							(0.00)		(0.87)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	29.05	18.83	0.39	1.13	-479.94	0.00	1.54	872.38	0.33
							(0.00)		(0.30)
$\text{IPCA}_{\text{K=5,large}}$	23.50	21.45	0.54	1.28	-424.84	0.00	1.10	682.46	0.35
							(0.03)		(0.32)
$\text{CA2}_{\text{K=5}}$	33.72	18.55	0.46	1.11	-505.96	0.00	1.82	1343.93	0.14
							(0.00)		(0.04)
NN3	30.40	18.07	0.43	1.08	-497.61	0.00	1.68	1222.79	0.17
							(0.00)		(0.06)
Panel C: DCC-GARCH with Nonlinear Shrinkage									
Return-based Estimators									
SAM_{mean}	2.48	33.66	0.02	2.10	-650.69	0.00	0.07	733.53	-0.36
							(0.04)		(0.00)
GAL_{mean}	29.56	43.03	0.60	2.68	-511.10	0.25	0.69	938.60	0.26
							(0.54)		(0.07)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	42.20	30.41	0.40	1.89	-691.46	0.00	1.39	1678.90	0.25
							(0.00)		(0.10)
FM_{wls}	40.64	29.83	0.40	1.86	-693.46	0.00	1.36	1680.70	0.20
							(0.00)		(0.06)
$\text{GFM}_{\text{large}}$	45.08	37.23	0.63	2.32	-621.39	0.00	1.21	1533.13	0.36
							(0.01)		(0.27)
GFM_{wls}	43.42	36.39	0.70	2.27	-623.58	0.00	1.19	1524.30	0.35
							(0.01)		(0.23)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	38.63	27.01	0.38	1.68	-689.29	0.00	1.43	1289.31	0.31
							(0.00)		(0.20)
$\text{IPCA}_{\text{K=5,large}}$	33.84	37.68	0.53	2.35	-603.52	0.00	0.90	1079.73	0.28
							(0.12)		(0.12)
$\text{CA2}_{\text{K=5}}$	35.55	28.89	0.41	1.80	-707.16	0.00	1.23	1803.81	-0.08
							(0.01)		(0.00)
NN3	35.29	29.49	0.39	1.84	-700.84	0.00	1.20	1694.66	0.04
							(0.01)		(0.01)

Table 7
OOS performance of TCM-RT portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
					Panel A: Benchmark Portfolios						
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
					Panel B: Covariance Matrix implied by the Single Index Model						
					Return-based Estimators						
SAM_{mean}	26.53	59.38	0.18	4.07	-427.96	0.00	0.45 (0.59)	225.65	179.62	271.69	0.37 (0.40)
GAL_{mean}	48.06	73.22	0.62	5.02	-250.28	0.00	0.66 (0.69)	191.21	148.43	234.00	0.62 (0.81)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	61.14	57.04	0.51	3.91	-420.24	0.00	1.07 (0.03)	994.55	791.77	1197.33	0.79 (0.31)
FM_{wls}	58.70	54.33	0.52	3.72	-432.83	0.00	1.08 (0.02)	995.69	776.22	1215.17	0.79 (0.30)
$\text{GFM}_{\text{large}}$	62.07	69.42	0.67	4.76	-327.26	0.00	0.89 (0.13)	658.51	522.03	794.98	0.72 (0.50)
GFM_{wls}	59.41	68.17	0.78	4.67	-329.09	0.00	0.87 (0.16)	619.43	402.02	836.83	0.75 (0.39)
					Machine Learning						
$\text{AGLASSO}_{\text{large}}$	53.50	55.41	0.52	3.80	-401.81	0.00	0.97 (0.07)	502.86	449.99	555.74	0.81 (0.26)
$\text{IPCA}_{\text{K=5,large}}$	54.53	69.45	0.65	4.76	-333.47	0.00	0.79 (0.33)	321.38	280.45	362.30	0.71 (0.53)
$\text{CA2}_{\text{K=5}}$	57.54	51.27	0.62	3.51	-407.62	0.00	1.12 (0.01)	938.77	722.28	1155.25	0.84 (0.21)
NN3	52.42	56.83	0.56	3.89	-372.22	0.00	0.92 (0.10)	706.81	468.27	945.36	0.75 (0.39)
					Panel C: Sample Covariance Matrix						
					Return-based Estimators						
SAM_{mean}	2.42	42.51	0.02	2.44	-1087.55	0.00	0.06 (0.05)	577.18	199.64	954.72	−0.06 (0.02)
GAL_{mean}	18.88	33.35	0.50	1.92	-722.35	0.00	0.57 (0.98)	210.17	39.32	381.02	0.54 (0.88)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	56.18	30.53	0.43	1.75	-1149.48	0.00	1.84 (0.00)	2168.11	972.17	3364.04	1.17 (0.00)
FM_{wls}	53.30	30.77	0.43	1.77	-1142.28	0.00	1.73 (0.00)	2141.32	936.51	3346.13	1.10 (0.01)
$\text{GFM}_{\text{large}}$	49.03	31.46	0.64	1.81	-1044.25	0.00	1.56 (0.00)	1680.29	546.22	2814.37	1.20 (0.00)
GFM_{wls}	44.52	31.10	0.69	1.79	-1022.23	0.00	1.43 (0.00)	1618.38	431.63	2805.13	1.14 (0.01)
					Machine Learning						
$\text{AGLASSO}_{\text{large}}$	37.23	30.32	0.33	1.74	-1140.23	0.00	1.23 (0.01)	1167.74	618.46	1717.03	0.83 (0.25)
$\text{IPCA}_{\text{K=5,large}}$	28.51	31.96	0.47	1.84	-1013.37	0.00	0.89 (0.13)	738.25	233.89	1242.61	0.74 (0.42)
$\text{CA2}_{\text{K=5}}$	42.83	29.10	0.40	1.67	-1121.50	0.00	1.47 (0.00)	2176.12	874.43	3477.80	0.83 (0.22)
NN3	36.56	29.11	0.35	1.67	-1155.32	0.00	1.26 (0.00)	2020.86	616.60	3425.12	0.76 (0.37)

Table 8

OOS performance of TCM-RT portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between an optimized portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
Panel A: Benchmark Portfolios											
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
Panel B: Galton Covariance Matrix											
Return-based Estimators											
SAM_{mean}	2.53	28.17	0.03	1.68	-442.56	0.00	0.09 (0.07)	187.47	85.73	289.21	0.02 (0.03)
GAL_{mean}	13.48	21.28	0.53	1.28	-229.60	0.00	0.63 (0.76)	50.04	19.94	80.15	0.62 (0.82)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	35.71	20.22	0.44	1.21	-493.21	0.00	1.77 (0.00)	1005.83	607.61	1404.04	1.16 (0.01)
FM_{wls}	34.36	20.06	0.44	1.20	-494.32	0.00	1.71 (0.00)	991.67	588.11	1395.22	1.13 (0.01)
GFM_{large}	30.33	21.33	0.62	1.27	-429.22	0.00	1.42 (0.00)	780.89	341.81	1219.97	1.12 (0.01)
GFM_{wls}	27.54	21.08	0.67	1.26	-425.22	0.00	1.31 (0.00)	746.63	270.84	1222.43	1.06 (0.01)
Machine Learning											
$AGLASSO_{large}$	23.47	18.64	0.34	1.12	-454.99	0.00	1.26 (0.00)	534.25	356.67	711.83	0.91 (0.14)
$IPCA_{K=5,large}$	18.94	20.21	0.49	1.21	-401.36	0.00	0.94 (0.09)	337.88	141.41	534.34	0.81 (0.28)
$CA2_{K=5}$	28.68	19.36	0.43	1.16	-476.95	0.00	1.48 (0.00)	974.46	552.17	1396.75	0.90 (0.11)
NN3	23.35	18.32	0.37	1.09	-469.49	0.00	1.27 (0.00)	840.29	375.19	1305.38	0.83 (0.20)
Panel C: DCC-GARCH with Nonlinear Shrinkage											
Return-based Estimators											
SAM_{mean}	5.86	31.57	0.04	1.97	-632.64	0.00	0.19 (0.11)	526.42	255.18	797.67	0.01 (0.02)
GAL_{mean}	26.01	38.31	0.59	2.40	-452.62	0.25	0.68 (0.60)	533.87	133.69	934.05	0.61 (0.86)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	38.24	30.84	0.39	1.92	-663.77	0.00	1.24 (0.01)	1247.29	654.28	1840.30	0.81 (0.27)
FM_{wls}	36.79	30.26	0.40	1.88	-663.96	0.00	1.22 (0.01)	1235.23	638.72	1831.74	0.80 (0.30)
GFM_{large}	38.25	36.28	0.59	2.26	-582.05	0.00	1.05 (0.04)	1056.69	414.84	1698.53	0.76 (0.39)
GFM_{wls}	35.66	35.23	0.66	2.20	-575.95	0.00	1.01 (0.05)	989.00	350.23	1627.77	0.79 (0.31)
Machine Learning											
$AGLASSO_{large}$	33.66	27.12	0.35	1.69	-656.50	0.00	1.24 (0.00)	844.24	511.59	1176.90	0.88 (0.14)
$IPCA_{K=5,large}$	28.92	35.65	0.50	2.23	-570.90	0.00	0.81 (0.25)	638.18	247.36	1029.00	0.66 (0.68)
$CA2_{K=5}$	31.96	29.44	0.41	1.84	-663.13	0.00	1.09 (0.04)	1273.56	610.79	1936.34	0.67 (0.66)
NN3	30.23	28.82	0.37	1.80	-662.01	0.00	1.05 (0.05)	1184.67	483.64	1885.70	0.69 (0.58)

Table 9

OOS performance of TCM-RT long-only portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT long-only Markowitz portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\hat{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\hat{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$	BR	SR	TO	TO1	TO2	NSR
Panel A: Benchmark Portfolios										
VW	8.59	14.99			0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
Panel B: Covariance Matrix implied by the Single Index Model										
Return-based Estimators										
SAM_{mean}	17.65	29.77	0.30	2.04	0.00	0.59 (0.91)	41.63	26.50	56.77	0.58 (0.98)
GAL_{mean}	23.36	42.47	0.55	2.91	0.00	0.55 (0.91)	46.64	36.42	56.86	0.53 (0.85)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	27.76	37.50	0.62	2.57	0.00	0.74 (0.33)	245.33	191.10	299.55	0.64 (0.67)
FM_{wls}	27.06	35.24	0.67	2.42	0.00	0.77 (0.24)	245.77	192.53	299.00	0.67 (0.57)
GFM_{large}	28.03	41.89	0.66	2.87	0.00	0.67 (0.60)	157.65	100.38	214.92	0.63 (0.76)
GFM_{wls}	26.64	40.94	0.76	2.81	0.00	0.65 (0.67)	153.45	85.50	221.41	0.61 (0.81)
Machine Learning										
$AGLASSO_{large}$	25.19	33.72	0.63	2.31	0.00	0.75 (0.29)	134.17	122.05	146.29	0.68 (0.47)
$IPCA_{K=5,large}$	25.45	40.57	0.72	2.78	0.00	0.63 (0.77)	76.34	59.32	93.37	0.60 (0.87)
$CA2_{K=5}$	25.75	34.99	0.73	2.40	0.00	0.74 (0.31)	248.41	193.31	303.51	0.62 (0.72)
NN3	26.23	37.64	0.65	2.58	0.00	0.70 (0.48)	191.15	130.16	252.14	0.63 (0.74)
Panel C: Sample Covariance Matrix										
Return-based Estimators										
SAM_{mean}	10.46	19.37	0.24	1.11	0.00	0.54 (0.78)	23.85	15.46	32.24	0.52 (0.69)
GAL_{mean}	11.90	19.66	0.57	1.20	0.00	0.61 (0.80)	12.22	9.72	14.72	0.60 (0.83)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	15.89	20.69	0.53	1.22	0.00	0.77 (0.13)	191.02	164.65	217.39	0.61 (0.74)
FM_{wls}	14.97	20.51	0.55	1.23	0.00	0.73 (0.20)	184.42	160.80	208.03	0.58 (0.93)
GFM_{large}	14.23	21.11	0.58	1.29	0.00	0.67 (0.42)	148.24	99.30	197.19	0.59 (0.86)
GFM_{wls}	13.39	20.55	0.67	1.30	0.00	0.65 (0.53)	131.61	78.89	184.34	0.59 (0.89)
Machine Learning										
$AGLASSO_{large}$	13.42	19.01	0.53	1.20	0.00	0.71 (0.26)	104.75	99.31	110.19	0.61 (0.72)
$IPCA_{K=5,large}$	12.37	19.30	0.61	1.22	0.00	0.64 (0.60)	59.84	38.44	81.23	0.61 (0.78)
$CA2_{K=5}$	14.35	21.80	0.59	1.31	0.00	0.66 (0.50)	192.29	168.63	215.96	0.50 (0.56)
NN3	15.59	20.79	0.61	1.24	0.00	0.75 (0.14)	166.63	125.14	208.12	0.62 (0.66)

Table 10
OOS performance of reward-to-risk timing strategies
The table presents the OOS performance of the reward-to-risk timing strategy of Kirby and Ostdiek (2012). The portfolios are constructed using the same investment universe and optimization inputs as the Markowitz portfolios. The expected returns are estimated by 10 different methods. The variances of stocks are estimated by two methods. In Panel B, we use the variance implied by the single index model. In Panel C, we use the sample variance based on daily returns. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	AS	BR	SR	TO	NSR
Panel A: Benchmark Portfolios									
VW	8.59	14.99			0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Covariance Matrix implied by the Single Index Model									
Return-based Estimators									
SAM_{mean}	8.79	14.24	0.43	1.09	48.95	0.00	0.62 (0.35)	10.47	0.60 (0.57)
GAL_{mean}	8.86	13.76	0.72	1.15	45.76	0.00	0.64 (0.26)	6.57	0.63 (0.34)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	11.13	14.28	0.81	1.22	57.19	0.00	0.78 (0.01)	78.76	0.62 (0.49)
FM_{wls}	11.13	14.39	0.86	1.21	56.83	0.00	0.77 (0.01)	80.42	0.62 (0.51)
$\text{GFM}_{\text{large}}$	9.98	13.91	0.80	1.20	49.82	0.00	0.72 (0.05)	41.89	0.63 (0.39)
GFM_{wls}	9.94	13.97	0.94	1.19	49.38	0.00	0.71 (0.05)	42.19	0.63 (0.38)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	9.92	13.65	0.75	1.13	53.72	0.00	0.73 (0.01)	42.97	0.63 (0.28)
$\text{IPCA}_{\text{K=5,large}}$	9.19	13.54	0.87	1.19	50.06	0.00	0.68 (0.15)	23.99	0.62 (0.50)
$\text{CA2}_{\text{K=5}}$	11.19	14.41	0.97	1.19	55.55	0.00	0.78 (0.01)	80.15	0.60 (0.67)
NN3	10.30	14.07	0.85	1.18	51.52	0.00	0.73 (0.03)	52.95	0.62 (0.49)
Panel C: Sample Covariance Matrix									
Return-based Estimators									
SAM_{mean}	8.88	14.47	0.43	0.91	49.85	0.00	0.61 (0.40)	10.47	0.59 (0.62)
GAL_{mean}	8.98	13.94	0.73	0.94	46.58	0.00	0.64 (0.26)	6.22	0.63 (0.34)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	11.23	14.48	0.81	0.99	58.01	0.00	0.78 (0.01)	79.19	0.62 (0.52)
FM_{wls}	11.25	14.58	0.86	0.99	57.68	0.00	0.77 (0.01)	80.92	0.61 (0.53)
$\text{GFM}_{\text{large}}$	10.12	14.10	0.81	0.97	50.70	0.00	0.72 (0.05)	42.42	0.63 (0.40)
GFM_{wls}	10.10	14.16	0.95	0.97	50.30	0.00	0.71 (0.05)	42.76	0.63 (0.39)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	10.01	13.82	0.76	0.93	54.33	0.00	0.72 (0.02)	43.37	0.63 (0.31)
$\text{IPCA}_{\text{K=5,large}}$	9.34	13.69	0.89	0.96	50.78	0.00	0.68 (0.14)	24.20	0.62 (0.47)
$\text{CA2}_{\text{K=5}}$	11.30	14.63	0.97	0.98	56.35	0.00	0.77 (0.01)	80.69	0.59 (0.72)
NN3	10.43	14.26	0.85	0.96	52.44	0.00	0.73 (0.03)	53.42	0.62 (0.51)

Table 11
OOS performance of parametric portfolio policies
The table presents the OOS performance of parametric portfolio policies (PPP) proposed by Brandt, Santa-Clara, and Valkanov (2009). The portfolios are constructed using the same investment universe as the Markowitz portfolios. We use estimated coefficients along with monthly lagged characteristics to form OOS PPP. The coefficients are updated at the end of June. The benchmark portfolio is the value-weighted portfolio (VW) whose OOS performance is reported in Panel A. PPP3 denotes the portfolio policy based on the three characteristics used in Brandt, Santa-Clara, and Valkanov (2009) (i.e. value, momentum and size); PPP62 denotes the portfolio policy based on the 62 characteristics used in our study. $\text{PPP62}_{\text{lasso}}$ denotes the portfolio policy that incorporates a lasso penalty into PPP62. Panel B reports the results for portfolio policies without transaction cost management. Panels C and D report results for transaction-cost-managed portfolio policies. In Panel C, our ex ante effective spread estimator is used , while in Panel D, the ex ante effective spread estimator of Brandt, Santa-Clara, and Valkanov (2009) is used (κ^*). The columns show descriptive statistics of the OOS performance for each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), active share (AS), the sum of negative weights (SNW), bankruptcy rate (percentage of months with before-cost returns below -100%)), Sharpe ratio (SR), turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a portfolio policy and the benchmark portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	AS	SNW	BR	SR	TO	NSR
				Panel A: Benchmark Portfolios				
VW	8.59	14.99	0.00	0.00	0.00	0.57	1.47	0.57
				Panel B: No Trading Cost Management				
PPP3	10.00	19.99	146.27	-89.23	0.00	0.50 (0.69)	81.32	0.37 (0.26)
PPP62	149.01	261.47	3098.47	-3033.51	3.43	0.57 (0.99)	8120.60	−0.46 (0.00)
				Panel C: Trading Cost Management (κ)				
PPP3	8.63	16.45	109.02	-59.57	0.00	0.52 (0.74)	52.81	0.43 (0.30)
PPP62	58.77	58.89	972.47	-906.68	0.00	1.00 (0.08)	703.75	0.59 (0.92)
$\text{PPP62}_{\text{lasso}}$	17.59	37.59	273.13	-246.55	0.00	0.47 (0.59)	216.41	0.27 (0.12)
				Panel D: Trading Cost Management (κ^*)				
PPP3	7.78	14.62	52.26	-19.25	0.00	0.53 (0.51)	16.32	0.50 (0.29)
PPP62	24.19	24.17	415.14	-349.30	0.00	1.00 (0.05)	137.25	0.84 (0.21)
$\text{PPP62}_{\text{lasso}}$	10.89	33.46	241.43	-216.26	0.00	0.33 (0.17)	157.36	0.20 (0.04)

Appendices

A Data transformation and missing Data handling

Following Freyberger, Neuhierl, and Weber (2020), Kelly, Pruitt, and Su (2019), Gu, Kelly, and Xiu (2020), Gu, Kelly, and Xiu (2021), and Chen, Pelger, and Zhu (2020), we use ranked-normalized characteristics. Specifically, we cross-sectionally rank all stocks by characteristic j in month t. Then, we normalized rank-transformed characteristic j to be within [-1, 1], that is,[42]

$$\tilde{c}_{i,j,t} = (\frac{\text{CSrank}(c_{i,j,t})}{N_{j,t} + 1} - 0.5) \times 2,$$

where $\text{CSrank}(c_{i,j,t})$ denotes cross-sectional rank of stock i by characteristic j in month t and $N_{j,t}$ denotes the number of stocks with non-missing characteristic j in month t. Rank transformation makes estimation insensitive to outliers and it is in line with the conventional portfolio sort approach.

When estimating parameters for methods based on large stocks, we drop small- and micro-cap stocks after data transformation. The exception is $\text{AGLASSO}_{\text{large}}$ since we need to use OLS estimates to compute BIC of a model; if we perform data transformation first and then keep large stocks, the design matrix of quadratic spline terms would be rank deficient in the sense that some columns associated with size-related characteristics such as firm size or bid-ask spread are vectors of zeroes. Hence, for $\text{AGLASSO}_{\text{large}}$, we first drop non-large stocks and then perform data transformation.

Our method for handling missing values is standard in the literature. When estimating parameters of NN3, $\text{CA2}_{\text{K=5}}$, and models based on Fama-MacBeth regressions, we follow Gu, Kelly, and Xiu (2020) and Gu, Kelly, and Xiu (2021) to fill in missing values with zeroes which are also the cross-sectional means and medians of ranked-normalized characteristics. When estimating parameters of $\text{AGLASSO}_{\text{large}}$ and $\text{IPCA}_{\text{K=5,large}}$, we follow respective studies to use stocks with non-missing 62 characteristics.

When forming return forecasts, we use the latest known ranked-normalized characteristics calculated from all stocks for all methods except $\text{AGLASSO}_{\text{large}}$. As what we do in parameter estimation, we fill in missing values with zeros. Given that in-

[42] The formula can be found in Equation A.1 of Gu, Kelly, and Xiu (2020) as well. Kelly, Pruitt, and Su (2019) do not multiply demeaned ranks by two so that their ranked-normalized characteristics are within [-0.5, 0.5]. Freyberger, Neuhierl, and Weber (2020) use $\tilde{c}_{i,j,t} = (\frac{\text{CSrank}(c_{i,j,t})}{N_{j,t}+1})$ so that their characteristics are within [0, 1].

vestment universe of interest consists of mega stocks with full return history over the past 60 months, there are few missing characteristics. Overall, for methods other than the adaptive group *lasso*, there are only 940 (0.46%) missing return forecasts, while there are 1263 (0.62%) missing return forecasts for the adaptive *lasso*. Extra 323 missing observations belong to stocks that are ranked in the top 500 by market equity at the beginning of a year but below NYSE median during the year. The 940 missing observations can be owing to changes in CRSP share code and/or CRSP exchange code during a holding period. We determine investment universe in December, we only consider U.S. common stocks listed on one of the three major exchanges; however, we apply the sample filters to the characteristic data *monthly*. A typical example is that few top-500 stocks that are defined as U.S. common stocks with a CRSP share code of 10 or 11 at the end of year $t-1$ become common stocks incorporated outside U.S (CRSP share code of 12) during year t and thus do no have valid stock characteristics after the change.

B Standard errors for hypothesis testing with Variance or Sharpe ratio

We follow Ledoit and Wolf (2008) and Ledoit and Wolf (2011) to compute standard errors used in our studentized i.i.d. bootstrap. Suppose there are two portfolios i and j whose excess returns in month t are r_{it} and r_{jt}, respectively. The returns can be either original returns or bootstrap returns. There are T return pairs. The two series are assumed to be strictly stationary and follow a bivariate return distribution with a mean vector $\begin{bmatrix} \mu_i \\ \mu_j \end{bmatrix}$ and a covariance matrix $\begin{bmatrix} \sigma_i^2 & \sigma_{ij} \\ \sigma_{ji} & \sigma_j^2 \end{bmatrix}$. Their sample estimates are denoted by $\begin{bmatrix} \hat{\mu}_i \\ \hat{\mu}_j \end{bmatrix}$ and $\begin{bmatrix} \hat{\sigma}_i^2 & \hat{\sigma}_{ij} \\ \hat{\sigma}_{ji} & \hat{\sigma}_j^2 \end{bmatrix}$. Also, the uncentered second moments of the two series are denoted by γ_i and γ_j, respectively. Let $\nu = (\mu_i, \mu_j, \gamma_i, \gamma_j)$ and $\hat{\nu} = (\hat{\mu}_i, \hat{\mu}_j, \hat{\gamma}_i, \hat{\gamma}_j)$. The difference of Sharpe ratios can be written as $\Delta_{sr} = f_{sr}(\nu)$ where $f_{sr}(a, b, c, d) = \frac{a}{\sqrt{c-a^2}} - \frac{b}{\sqrt{d-b^2}}$; the difference of (log) variance can be denoted by $\Delta_{var} = f_{var}(\nu)$ where $f_{var}(a, b, c, d) = log(c - a^2) - log(d - b^2)$. Ledoit and Wolf (2008) assume that

$$\sqrt{T}(\hat{\nu} - \nu) \to N(0; \Psi) \tag{B.1}$$

so that the distribution of the difference of Sharpe ratios or variances can be expresses as

$$\sqrt{T}(\hat{\Delta} - \Delta) \to N\big(0; \nabla' f(\nu) \Psi \nabla f(\nu)\big) \tag{B.2}$$

using the delta method where $\nabla' f_{sr}(\nu)$ is given by

$$\Big(\frac{c}{(c-a^2)^{1.5}}, \frac{-d}{(d-b^2)^{1.5}}, \frac{-0.5a}{(c-a^2)^{1.5}}, \frac{-0.5d}{(d-a^2)^{1.5}}\Big)$$

and $\nabla' f_{var}(\nu)$ is given by

$$\Big(\frac{2a}{c-a^2}, \frac{2b}{d-b^2}, \frac{1}{c-a^2}, \frac{-1}{(d-a^2)^{1.5}}\Big).$$

The standard error used in our i.i.d bootstrap[43] is calculated as

$$s(\hat{\Delta}) = \nabla' f(\hat{\nu}) \hat{\Psi} \nabla f(\hat{\nu}) \tag{B.3}$$

where $\hat{\Psi}$ for i.i.d data is the sample covariance matrix of $(r_{it}, r_{j,t}, r_{i,t}^2, r_{j,t}^2)$ (see Section 3.2.1 and Footnote 9 of Ledoit and Wolf (2008) for more details about the calculation of $\hat{\Psi}$).

C Estimation of effective spreads

We primarily use **one-half** of CRSP closing quoted spread (QS) to estimate effective one-way spreads ($\kappa_{i,t}$). QS at the end of month t for stock i is calculated as

$$\mathrm{QS_{i,t}} = \frac{1}{D_t} \sum_{d=1}^{D_t} \frac{Ask_{i,d} - Bid_{i,d}}{\big(\frac{Ask_{i,d}+Bid_{i,d}}{2}\big)},$$

where $Ask_{i,d}$ and $Bid_{i,d}$ denote daily closing bid and ask quotes, respectively. Daily price data are obtained from CRSP. D_t denotes the number of valid daily observations in month t. Zero daily bid-ask spreads or daily spreads higher than 50% are excluded from the calculation. Comparing several low-frequency effective spread estimators, Abdi and Ranaldo (2017) document that QS is *generally* the most accurate proxy from 1993 onwards in terms of cross-sectional and time-series correlations with effective spreads estimated from the Trade and Quote (TAQ) data as well as prediction errors for the TAQ spreads. A continuous series of daily quote data for NYSE and AMEX stocks becomes available on CRSP from December 28, 1992.[44] As for stocks listed on NASDAQ, daily quote data for National Market (NM) secu-

[43] Of course, one can conduct an inference based on asymptotic normality using Eq. (B.2). However, as Ledoit and Wolf (2008) argue, it is inferior to the studentized bootstrap method. In addition, the bootstrap method is used in the recent paper by DeMiguel et al. (2020).

[44] CRSP also provides daily quote data over the period December 31, 1925 to February 23 1942. For the period February 1942 to December 27, 1992, daily quote data are only available for stocks without closing prices. Bid and ask prices of NYSE and AMEX stocks are based on the last representative quote before market close of each trading day, while those of NASDAQ stocks are based on closing inside quotation.

rities begin on November 1, 1982 and become available for all securities from June 15, 1992. Since our OOS period begins in January 1987 and investment universe consists of the top-500 largest common stocks traded on NYSE, AMEX or NASDAQ by market equity, most of stocks in our investment universe have valid values of $\mathrm{QS_{i,t}}$. Specifically, over the period 1993-2020 (1987-1992), 98% (9.6%) of observations have non-missing $\mathrm{QS_{i,t}}$. We fill in missing values especially those over the 1977-1992 period with one half of Abdi and Ranaldo (2017)'s estimates for bid-ask spreads, which are computed using daily Close, High and Low (CHL) prices, all of which are available over the entire CRSP sample period. The two-day corrected version of CHL for stock i at the end of month t is given by

$$\mathrm{CHL_{i,t}} = \frac{1}{D_t}\sum_{d=1}^{D_t} \hat{s}_{i,t},\ \hat{s}_{i,d} = \sqrt{\max\{4(c_{i,d}-\eta_{i,d})(c_{i,d}-\eta_{i,d+1}), 0\}},$$

where D_t is the number of trading days in month t, $\hat{s}_{i,t}$ is a two-day estimate of the bid-ask spread, $c_{i,d}$ is the daily close log-price, and $\eta_{i,d}$ is the average of daily low and high log-prices.[45].

Abdi and Ranaldo (2017) show that in the absence of QS, CHL is the best performing estimator in terms of the evaluation criteria mentioned above. Notably, it even outperforms QS in terms of cross-sectional correlation and prediction errors during the 1993-2000 period, but the latter still has the highest time-series correlation coefficient with the TAQ effective spreads. Furthermore, DeMiguel et al. (2020) use CHL to check the robustness of their results regarding trading diversification benefits, that is, combining multiple characteristics can help reduce trading costs,

We fill in the remaining missing values with annual effective costs of Hasbrouck (2009, Gibbs) first. He estimates the Roll (1984) model by the Gibbs sampler using daily close prices only.[46] We then follow Novy-Marx and Velikov (2016) to fill in the rest of missing values with their closest peers in terms of the rank of market value and idiosyncratic volatility and their closest matches in terms of market value only.

D Additional tables

[45]We are grateful to the authors for making their code online. Note that we slightly modify their code by excluding the last daily observation ($\hat{s}_{i,d}$) from monthly average as it is computed with price information on the first trading day of next month ($\eta_{i,d+1}$), which is after portfolio formation. Again, we divide the measure by 2 as our proportional transaction cost. The adjusted measure is very close to the original one with a correlation of 99.8% for all stocks.

[46]Note that when implementing transaction-cost-managed strategies, we use previous year's Gibbs estimates as inputs to avoid look-ahead bias.

Table A1

62 firm-specific characteristics by category

The table lists the firm characteristics used in our empirical analysis by category.

	Past Returns	
(1)	$\text{ret}_{1,2}$	Short-term reversal
(2)	$\text{ret}_{2,6}$	Recent momentum
(3)	$\text{ret}_{2,12}$	Momentum
(4)	$\text{ret}_{7,12}$	Intermediate momentum
(5)	$\text{ret}_{12,36}$	Long-term reversal
	Investment	
(6)	Investment	Log growth in total assets
(7)	ΔCEQ	percentage change in book equity
(8)	ΔPI2A	Change in PPE and Inventory scaled by lagged total assets
(9)	ΔShrout	Log growth in split-adjusted shares outstanding from t-36 to t-1
(10)	IVC	Change in inventory scaled by average total assets
(11)	NOA	Net operating assets over lagged total assets
	Profitability	
(12)	ATO	Sales to lagged net operating assets
(13)	CTO	Capital turnover
(14)	$\Delta(\Delta GM - \Delta Sales)$	Δ(% change in gross margin and % change in sales)
(15)	EPS	Earnings per share
(16)	IPM	Before-tax profit margin (pretax income over sales)
(17)	IPM_{adj}	Fama-French 48 industry adjusted IPM
(18)	PCM	Price-to-cost margin (sales minus COGS to sales)
(19)	PM	Profit margin (OI after depreciation over sales)
(20)	Profitability	Gross profitability scaled by book value of equity
(21)	RNA	Return on net operating assets
(22)	ROA	Return on assets
(23)	ROC	Productivity of cash
(24)	ROE	Return on equity
(25)	ROIC	Return on invested capital
(26)	S2C	Sales to cash
(27)	SAT	Sales to total assets
(28)	SAT_{adj}	Fama-French 48 industry adjusted SAT
	Intangibles	
(29)	AOA	Absolute value of operating accruals
(30)	OL	Operating leverage
(31)	Tan	Tangibility
(32)	OA	Operating accruals
	Value	
(33)	A2ME	Assets to market value of equity
(34)	BEME	Log book value minus log market value of equity in prior fiscal year
(35)	BEME_{adj}	Fama-French 48 industry adjusted book to market ratio
(36)	C	Ratio of cash and short-term investments to total assets
(37)	C2D	Cash flow to total debt
(38)	ΔSO	Log change in split-adjusted shares outstanding in prior fiscal year
(39)	Debt2P	Total debt to market value of equity
(40)	E2P	Earnings to market value of equity
(41)	Free CF	Free cash flow to Book Equity
(42)	LDP	Dividend yield over past 12 months
(43)	NOP	Net payouts to market value of equity
(44)	O2P	Operating payouts to market value of equity
(45)	Q	Tobin's Q
(46)	S2P	Sales to market value of equity
(47)	ΔSales	Sales Growth
Trading Frictions		
(48)	AT	Log Total assets
(49)	Beta	Frazzini-Pederson beta
(50)	Beta daily	Dimson beta using daily returns
(51)	DTO	De-trended turnover - market turnover
(52)	Idio vol	Idio vol of Fama-French 3-factor model
(53)	LME	Log market value of equity
(54)	LME_{adj}	Fama-French 48 industry adjusted market value of equity
(55)	LTO	Monthly turnover in month t-1
(56)	PTH	Price in month t-2 to 52-week high
(57)	Ret_{max}	Maximum daily return in month t-1
(58)	Spread	Average daily bid-ask spread in month t-1
(59)	σ_{to}	Standard deviation of daily turnover
(60)	σ_{vol}	Standard deviation of daily volume
(61)	SUV	Standard unexplained volume
(62)	σ_{ret}	Standard deviation of daily returns

Table A2

OOS performance of complete portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). complete portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a complete portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	AS	BR	SR	TO	NSR
					Panel A: Benchmark Portfolios					
VW	8.59	14.99			0.00	0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
					Panel B: Covariance Matrix implied by the Single Index Model					
					Return-based Estimators					
SAM_{mean}	448.40	1027.50	0.19	4.81	-6485.28	8708.69	25.49	0.44 (0.56)	19137.93	-0.05 (0.01)
GAL_{mean}	399.44	621.83	0.58	5.46	-2191.06	4882.47	19.85	0.64 (0.73)	32395.00	-0.03 (0.01)
					Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	763.60	740.95	0.47	4.20	-5442.00	7594.55	14.71	1.03 (0.04)	21017.11	0.12 (0.07)
FM_{wls}	697.51	681.58	0.49	4.12	-5257.17	7138.40	13.73	1.02 (0.05)	19717.29	0.10 (0.04)
$\text{GFM}_{\text{large}}$	608.83	663.10	0.63	4.87	-3338.42	5852.47	15.69	0.92 (0.11)	11943.09	0.40 (0.42)
GFM_{wls}	501.74	560.62	0.73	4.90	-2827.18	4920.68	14.46	0.89 (0.14)	12525.90	0.34 (0.30)
					Machine Learning					
$\text{AGLASSO}_{\text{large}}$	580.53	559.01	0.47	3.69	-4432.18	6349.22	15.93	1.04 (0.03)	21463.59	0.29 (0.21)
$\text{IPCA}_{\text{K=5,large}}$	469.15	593.02	0.60	4.84	-3071.12	5370.19	16.42	0.79 (0.30)	11596.44	0.20 (0.08)
$\text{CA2}_{\text{K=5}}$	593.65	559.10	0.57	3.96	-4321.25	5921.99	11.52	1.06 (0.02)	16283.60	0.26 (0.19)
NN3	517.98	590.79	0.50	4.19	-3779.22	5819.76	13.73	0.88 (0.16)	14819.03	0.16 (0.07)
					Panel C: Sample Covariance Matrix					
					Return-based Estimators					
SAM_{mean}	19.24	459.43	0.01	2.69	-11700.72	11979.28	15.93	0.04 (0.04)	41224.45	-1.02 (0.00)
GAL_{mean}	59.66	124.42	0.44	2.41	-2977.50	3503.62	1.23	0.48 (0.62)	3524.28	-0.23 (0.00)
					Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	679.23	309.44	0.47	1.84	-12422.33	12837.88	3.92	2.20 (0.00)	47556.50	-0.13 (0.00)
FM_{wls}	628.81	295.88	0.48	1.84	-11837.13	12212.75	3.92	2.13 (0.00)	36904.86	0.00 (0.01)
$\text{GFM}_{\text{large}}$	421.82	206.50	0.75	1.98	-7113.44	7633.25	0.98	2.04 (0.00)	22817.47	0.07 (0.01)
GFM_{wls}	340.40	175.43	0.84	1.97	-6154.16	6590.37	0.98	1.94 (0.00)	14980.61	0.20 (0.06)
					Machine Learning					
$\text{AGLASSO}_{\text{large}}$	458.46	294.84	0.36	2.04	-10249.33	10668.91	4.17	1.55 (0.00)	21448.96	-0.26 (0.00)
$\text{IPCA}_{\text{K=5,large}}$	248.27	191.05	0.62	2.27	-5661.37	6134.49	1.47	1.30 (0.00)	10766.41	-0.18 (0.00)
$\text{CA2}_{\text{K=5}}$	439.42	245.08	0.43	1.74	-10652.95	10920.84	2.45	1.79 (0.00)	40374.68	-0.36 (0.00)
NN3	411.41	258.77	0.43	1.90	-10253.71	10624.02	2.45	1.59 (0.00)	30414.36	-0.45 (0.00)

Table A3

OOS performance of complete portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of complete portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a complete portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	AS	BR	SR	TO	NSR
Panel A: Benchmark Portfolios										
VW	8.59	14.99			0.00	0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			0.00	45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Galton Covariance Matrix										
Return-based Estimators										
SAM_{mean}	5.06	227.54	0.01	1.92	-3342.99	3483.87	5.15	0.02 (0.03)	6037.53	-0.48 (0.00)
GAL_{mean}	29.23	59.23	0.45	1.68	-618.23	874.19	0.00	0.49 (0.69)	509.57	0.09 (0.02)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	316.65	148.15	0.51	1.35	-3471.09	3688.51	0.25	2.14 (0.00)	9576.36	0.38 (0.35)
FM_{wls}	297.34	139.51	0.52	1.33	-3323.83	3518.21	0.00	2.13 (0.00)	9025.69	0.07 (0.09)
$\text{GFM}_{\text{large}}$	200.19	108.81	0.76	1.57	-1962.98	2235.99	0.49	1.84 (0.00)	5032.31	0.44 (0.49)
GFM_{wls}	161.90	90.00	0.88	1.52	-1666.43	1893.47	0.25	1.80 (0.00)	4414.12	0.22 (0.18)
Machine Learning										
$\text{AGLASSO}_{\text{large}}$	213.08	138.07	0.40	1.48	-2846.22	3059.31	0.25	1.54 (0.00)	8056.87	0.04 (0.03)
$\text{IPCA}_{\text{K=5,large}}$	127.70	98.72	0.64	1.74	-1559.84	1796.75	0.25	1.29 (0.00)	2754.52	0.32 (0.22)
$\text{CA2}_{\text{K=5}}$	217.65	109.82	0.52	1.21	-2852.66	2991.63	0.00	1.98 (0.00)	7531.97	0.07 (0.01)
NN3	182.23	105.50	0.48	1.24	-2587.80	2781.81	0.00	1.73 (0.00)	7055.96	-0.16 (0.00)
Panel C: DCC-GARCH with Nonlinear Shrinkage										
Return-based Estimators										
SAM_{mean}	94.24	329.47	0.06	1.89	-7549.94	8323.61	12.50	0.29 (0.21)	28014.78	-0.58 (0.00)
GAL_{mean}	133.38	141.77	0.57	2.22	-2178.16	3070.09	1.47	0.94 (0.04)	5349.67	0.35 (0.21)
Fama-MacBeth Regressions without/with the Galton Correction										
FM_{large}	399.36	211.24	0.41	1.54	-6270.98	7013.88	2.94	1.89 (0.00)	181139.51	-0.10 (0.14)
FM_{wls}	369.74	198.62	0.42	1.53	-5988.61	6650.79	2.45	1.86 (0.00)	16035.07	0.41 (0.45)
$\text{GFM}_{\text{large}}$	292.17	170.76	0.62	1.82	-3767.32	4618.96	1.72	1.71 (0.00)	11147.19	0.53 (0.81)
GFM_{wls}	239.81	144.78	0.71	1.81	-3258.70	3970.61	0.98	1.66 (0.00)	9288.15	0.53 (0.82)
Machine Learning										
$\text{AGLASSO}_{\text{large}}$	304.62	212.70	0.28	1.62	-6074.38	6700.47	2.70	1.43 (0.00)	17325.47	-0.05 (0.01)
$\text{IPCA}_{\text{K=5,large}}$	219.21	160.13	0.57	1.92	-3308.20	4081.82	0.98	1.37 (0.00)	8206.46	0.24 (0.14)
$\text{CA2}_{\text{K=5}}$	269.16	158.38	0.40	1.41	-5377.29	5854.51	1.47	1.70 (0.00)	14559.44	0.16 (0.06)
NN3	271.16	171.84	0.38	1.49	-5438.21	6034.90	1.72	1.58 (0.00)	14540.97	0.11 (0.04)

Table A4

OOS performance of TCM portfolios

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). transaction-costs-managed (TCM) portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\hat{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below −100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a complete portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\hat{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
Panel A: Benchmark Portfolios											
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58	7.58	8.31	6.84	0.56
							(0.96)				(0.84)
Panel B: Covariance Matrix implied by the Single Index Model											
Return-based Estimators											
SAM_{mean}	446.46	943.90	0.21	4.79	-5398.64	21.32	0.47	35706.85	51811.01	19602.70	−0.16
							(0.68)				(0.01)
GAL_{mean}	339.75	540.78	0.59	5.62	-1575.28	13.97	0.63	8094.87	5124.46	11065.28	0.38
							(0.78)				(0.36)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	599.38	667.31	0.45	4.43	-4257.80	13.97	0.90	13256.00	9994.09	16517.92	0.55
							(0.15)				(0.92)
FM_{wls}	542.99	616.46	0.47	4.35	-4126.43	12.50	0.88	14802.26	9213.12	20391.41	0.50
							(0.17)				(0.75)
GFM_{large}	458.56	576.72	0.60	5.05	-2438.85	14.46	0.80	7788.91	5741.93	9835.88	0.56
							(0.28)				(0.98)
GFM_{wls}	356.88	489.48	0.68	5.23	-1993.72	12.25	0.73	9638.36	5341.53	13935.18	0.42
							(0.47)				(0.45)
Machine Learning											
$AGLASSO_{large}$	441.40	505.35	0.43	3.87	-3476.71	13.48	0.87	11872.95	9399.96	14345.93	0.30
							(0.17)				(0.36)
$IPCA_{K=5,large}$	375.41	520.18	0.60	5.03	-2294.75	13.24	0.72	8788.66	8974.05	8603.26	0.14
							(0.48)				(0.13)
$CA2_{K=5}$	469.06	518.73	0.56	4.31	-3340.02	9.56	0.90	11044.45	6440.02	15648.88	0.62
							(0.13)				(0.83)
NN3	409.60	534.67	0.49	4.46	-2870.98	12.25	0.77	19943.61	4008.97	35878.24	0.55
							(0.37)				(0.93)
Panel C: Sample Covariance Matrix											
Return-based Estimators											
SAM_{mean}	75.04	368.06	0.06	2.47	-8770.41	12.50	0.20	42806.70	10133.29	75480.10	−0.29
							(0.15)				(0.00)
GAL_{mean}	49.27	92.37	0.50	2.06	-1983.83	0.49	0.53	989.23	443.57	1534.89	0.42
							(0.84)				(0.45)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	431.96	247.12	0.40	1.82	-9263.71	3.68	1.75	36275.58	7148.42	65402.75	0.88
							(0.00)				(0.18)
FM_{wls}	392.41	236.34	0.41	1.84	-8733.38	3.92	1.66	201596.13	6076.66	397115.60	−0.02
							(0.00)				(0.35)
GFM_{large}	211.36	159.07	0.56	1.91	-5059.88	1.47	1.33	11719.34	6634.84	16803.85	0.03
							(0.00)				(0.18)
GFM_{wls}	176.33	130.21	0.65	1.84	-4329.30	0.74	1.35	8771.73	1726.76	15816.71	0.74
							(0.00)				(0.53)
Machine Learning											
$AGLASSO_{large}$	245.25	212.65	0.26	1.83	-7337.09	3.43	1.15	11830.79	5873.24	17788.34	0.36
							(0.01)				(0.50)
$IPCA_{K=5,large}$	139.90	141.70	0.54	2.10	-3912.16	0.74	0.99	4181.65	1921.15	6442.16	0.66
							(0.04)				(0.68)
$CA2_{K=5}$	251.05	197.19	0.33	1.75	-7946.57	2.45	1.27	19844.33	4652.39	35036.27	0.61
							(0.00)				(0.84)
NN3	225.92	208.89	0.32	1.91	-7559.10	2.21	1.08	18468.71	3435.92	33501.50	0.58
							(0.02)				(0.95)

Table A5

OOS performance of TCM portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of transaction-cost-managed (TCM) portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
Panel A: Benchmark Portfolios											
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
Panel B: Galton Covariance Matrix											
Return-based Estimators											
SAM_{mean}	18.51	191.77	0.03	1.72	-2756.75	4.17	0.10 (0.06)	3728.10	3356.76	4099.44	-0.28 (0.00)
GAL_{mean}	23.65	49.25	0.46	1.53	-424.38	0.00	0.48 (0.63)	111.19	90.92	131.45	0.44 (0.52)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	211.26	119.63	0.45	1.29	-2734.55	0.25	1.77 (0.00)	8588.80	5266.40	11911.19	0.27 (0.46)
FM_{wls}	195.20	113.32	0.46	1.29	-2593.57	0.25	1.72 (0.00)	5438.04	3672.33	7203.74	1.01 (0.03)
GFM_{large}	108.59	85.01	0.59	1.48	-1423.39	0.49	1.28 (0.00)	2520.61	1609.48	3431.74	0.93 (0.06)
GFM_{wls}	81.79	68.94	0.65	1.40	-1221.38	0.25	1.19 (0.00)	2049.09	992.24	3105.94	0.91 (0.07)
Machine Learning											
$AGLASSO_{large}$	134.61	115.03	0.31	1.41	-2293.52	0.00	1.17 (0.01)	5462.52	3220.77	7704.28	0.28 (0.30)
$IPCA_{K=5,large}$	68.71	82.72	0.49	1.63	-1246.48	0.25	0.83 (0.22)	1171.46	1273.96	1068.95	0.36 (0.39)
$CA2_{K=5}$	132.79	88.05	0.43	1.18	-2145.75	0.00	1.51 (0.00)	4535.05	2530.17	6539.92	0.90 (0.10)
NN3	110.13	87.97	0.39	1.23	-1987.51	0.25	1.25 (0.00)	3608.32	1806.57	5410.07	0.80 (0.24)
Panel C: DCC-GARCH with Nonlinear Shrinkage											
Return-based Estimators											
SAM_{mean}	101.82	295.33	0.07	1.83	-6579.21	9.80	0.34 (0.33)	32010.15	33654.84	30365.46	-0.20 (0.00)
GAL_{mean}	110.56	118.56	0.57	2.15	-1727.95	0.74	0.93 (0.05)	4645.39	386.08	8904.69	0.76 (0.32)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	302.28	189.88	0.37	1.61	-5255.46	2.70	1.59 (0.00)	15992.22	4255.88	27728.57	1.07 (0.01)
FM_{wls}	277.00	177.77	0.38	1.58	-5038.79	1.72	1.56 (0.00)	13024.33	7694.69	18353.98	0.26 (0.39)
GFM_{large}	206.65	148.16	0.56	1.85	-3068.72	0.98	1.39 (0.00)	21751.56	9198.77	34304.34	-0.05 (0.13)
GFM_{wls}	162.47	123.15	0.62	1.83	-2607.24	0.74	1.32 (0.00)	5744.65	1332.34	10156.96	1.06 (0.01)
Machine Learning											
$AGLASSO_{large}$	224.57	188.60	0.24	1.64	-5144.98	2.45	1.19 (0.01)	9842.37	5469.91	14214.84	0.53 (0.87)
$IPCA_{K=5,large}$	154.73	139.51	0.51	1.89	-2744.68	0.49	1.11 (0.00)	5926.77	2547.70	9305.85	0.62 (0.83)
$CA2_{K=5}$	198.00	142.70	0.37	1.52	-4402.07	1.23	1.39 (0.00)	10245.18	2913.33	17577.03	0.85 (0.18)
NN3	206.04	154.26	0.36	1.57	-4481.86	1.47	1.34 (0.00)	9779.93	2475.35	17084.51	0.95 (0.05)

Table A6
OOS performance of TCM-BC portfolios
At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). TCM-BC portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). TCM-BC portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and a budget constraint. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a complete portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
						Panel A: Benchmark Portfolios					
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
						Panel B: Covariance Matrix implied by the Single Index Model					
						Return-based Estimators					
SAM_{mean}	144.07	719.23	0.09	4.20	-6077.02	21.57	0.20 (0.14)	29634.96	17451.05	41818.86	-0.22 (0.00)
GAL_{mean}	16.77	54.37	0.80	4.14	-367.60	0.00	0.31 (0.29)	201.23	15.27	387.19	0.28 (0.24)
						Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	359.72	307.78	0.43	2.56	-4910.27	5.15	1.17 (0.01)	16180.44	10492.74	21868.13	0.30 (0.31)
FM_{wls}	336.38	289.24	0.44	2.53	-4646.80	5.88	1.16 (0.01)	11865.43	5845.57	17885.30	0.69 (0.59)
GFM_{large}	161.36	155.87	0.69	2.49	-2463.68	1.47	1.04 (0.05)	9600.98	2381.81	16820.15	0.68 (0.63)
GFM_{wls}	124.22	132.68	0.74	2.51	-2065.57	0.98	0.94 (0.12)	4492.03	1348.60	7635.46	0.70 (0.56)
						Machine Learning					
$AGLASSO_{large}$	190.34	263.57	0.28	2.63	-3741.63	4.41	0.72 (0.54)	27404.04	33936.50	20871.59	-0.18 (0.01)
$IPCA_{K=5,large}$	97.31	157.92	0.46	2.73	-2240.95	1.23	0.62 (0.87)	3078.63	2397.93	3759.33	0.27 (0.24)
$CA2_{K=5}$	240.10	230.24	0.43	2.36	-3837.14	3.19	1.04 (0.05)	9768.34	4164.34	15372.33	0.62 (0.84)
NN3	200.61	205.64	0.44	2.29	-3434.12	2.70	0.98 (0.09)	7377.55	4455.00	10300.10	0.39 (0.51)
						Panel C: Sample Covariance Matrix					
						Return-based Estimators					
SAM_{mean}	54.19	364.66	0.04	2.40	-9341.19	12.25	0.15 (0.11)	29695.60	13406.49	45984.71	-0.38 (0.00)
GAL_{mean}	8.71	22.69	0.50	1.83	-485.58	0.00	0.38 (0.45)	105.77	13.65	197.89	0.37 (0.42)
						Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	393.82	228.50	0.39	1.75	-9357.20	2.94	1.72 (0.00)	27242.17	7269.98	47214.36	0.86 (0.22)
FM_{wls}	366.09	220.00	0.40	1.77	-8815.38	2.45	1.66 (0.00)	24423.83	5979.51	42868.15	0.93 (0.11)
GFM_{large}	176.33	112.52	0.64	1.67	-4581.01	0.00	1.57 (0.00)	8990.40	2340.72	15640.08	1.09 (0.02)
GFM_{wls}	136.40	98.85	0.67	1.73	-3925.70	0.00	1.38 (0.00)	7378.02	1384.48	13371.57	1.03 (0.04)
						Machine Learning					
$AGLASSO_{large}$	211.03	196.73	0.23	1.79	-7276.83	2.70	1.07 (0.04)	2924932.95	5499.41	5844366.48	-0.17 (0.13)
$IPCA_{K=5,large}$	85.49	105.10	0.47	1.94	-3480.29	0.49	0.81 (0.30)	3147.88	1248.02	5047.74	0.56 (0.98)
$CA2_{K=5}$	223.22	185.67	0.31	1.71	-7948.68	1.72	1.20 (0.01)	30738.41	4334.82	57142.01	0.41 (0.55)
NN3	199.59	191.60	0.31	1.84	-7635.23	1.72	1.04 (0.05)	21720.87	3287.43	40154.31	0.21 (0.20)

Table A7

OOS performance of TCM-BC portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-BC portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). TCM-BC portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and a budget constraint. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
					Panel A: Benchmark Portfolios						
VW	8.59	14.99			0.00	0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
					Panel B: Galton Covariance Matrix						
					Return-based Estimators						
SAM_{mean}	20.43	197.51	0.03	1.75	-2978.56	3.43	0.10 (0.07)	4837.90	5527.44	4148.36	-0.33 (0.00)
GAL_{mean}	8.06	14.44	0.54	1.22	-166.36	0.00	0.56 (0.95)	35.39	9.05	61.72	0.55 (0.91)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	194.11	103.93	0.45	1.19	-2764.59	0.00	1.87 (0.00)	9897.60	12458.20	7336.99	-0.05 (0.37)
FM_{wls}	180.83	98.35	0.46	1.19	-2612.91	0.25	1.84 (0.00)	5149.21	3376.00	6922.43	1.17 (0.01)
GFM_{large}	87.95	54.41	0.68	1.18	-1397.10	0.00	1.62 (0.00)	2398.01	1420.00	3376.02	1.16 (0.01)
GFM_{wls}	66.27	44.85	0.73	1.15	-1157.13	0.00	1.48 (0.00)	1955.69	845.57	3065.81	1.14 (0.01)
					Machine Learning						
$AGLASSO_{large}$	114.74	98.07	0.29	1.32	-2270.24	0.00	1.17 (0.01)	3938.64	2497.39	5379.89	0.57 (1.00)
$IPCA_{K=5,large}$	51.13	59.02	0.50	1.43	-1208.92	0.25	0.87 (0.20)	897.36	763.10	1031.63	0.65 (0.73)
$CA2_{K=5}$	123.41	79.39	0.42	1.11	-2178.66	0.00	1.55 (0.00)	4481.57	2479.92	6483.22	0.90 (0.12)
NN3	94.81	73.24	0.38	1.10	-2017.30	0.00	1.29 (0.00)	3525.59	1694.52	5356.65	0.79 (0.32)
					Panel C: DCC-GARCH with Nonlinear Shrinkage						
					Return-based Estimators						
SAM_{mean}	18.80	282.02	0.01	1.82	-6857.05	10.29	0.07 (0.05)	18410.24	11314.40	25506.07	-0.42 (0.00)
GAL_{mean}	8.99	21.13	0.46	1.61	-434.19	0.00	0.43 (0.53)	228.03	34.05	422.01	0.39 (0.44)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	210.74	143.32	0.32	1.35	-5319.30	0.74	1.47 (0.00)	11188.03	4040.75	18335.31	0.85 (0.22)
FM_{wls}	194.33	137.15	0.33	1.35	-5056.59	0.25	1.42 (0.00)	11458.13	3770.96	19145.29	0.81 (0.30)
GFM_{large}	102.72	76.81	0.53	1.36	-2718.63	0.00	1.34 (0.00)	4634.57	1504.75	7764.40	0.94 (0.10)
GFM_{wls}	79.95	66.77	0.56	1.39	-2329.56	0.00	1.20 (0.01)	3920.59	965.16	6876.02	0.90 (0.14)
					Machine Learning						
$AGLASSO_{large}$	148.21	167.17	0.18	1.64	-4963.54	2.45	0.89 (0.22)	8045.55	4159.00	11932.10	0.39 (0.47)
$IPCA_{K=5,large}$	58.31	81.88	0.37	1.57	-2331.82	0.25	0.71 (0.54)	1942.60	1071.18	2814.02	0.46 (0.63)
$CA2_{K=5}$	136.66	118.44	0.29	1.35	-4492.41	0.74	1.15 (0.02)	9672.37	2672.27	16672.46	0.64 (0.78)
NN3	131.16	119.47	0.27	1.34	-4589.39	0.49	1.10 (0.03)	10279.48	2300.27	18258.68	0.55 (0.95)

Table A8

OOS performance of TCM-RT long-only portfolios (Galton and GARCH covariances)

At the end of each calendar year, we select the 500 largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT long-only portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\mu_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\sigma_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below −100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\mu_{exp}}$	$\frac{\hat{\sigma}}{\sigma_{exp}}$	BR	SR	TO	TO1	TO2	NSR
					Panel A: Benchmark Portfolios					
VW	8.59	14.99			0.00	0.57	1.47	1.77	1.16	0.57
EW	9.44	16.41			0.00	0.58 (0.96)	7.58	8.31	6.84	0.56 (0.84)
					Panel B: Galton Covariance Matrix					
					Return-based Estimators					
SAM_{mean}	9.94	18.02	0.24	1.08	0.00	0.55 (0.84)	19.40	13.09	25.71	0.54 (0.75)
GAL_{mean}	10.32	17.28	0.56	1.13	0.00	0.60 (0.84)	9.62	7.98	11.26	0.59 (0.87)
					Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	14.08	18.62	0.53	1.15	0.00	0.76 (0.13)	167.98	150.70	185.26	0.60 (0.82)
FM_{wls}	13.41	18.40	0.54	1.16	0.00	0.73 (0.17)	162.70	147.58	177.81	0.58 (0.95)
$\text{GFM}_{\text{large}}$	12.45	18.77	0.56	1.22	0.00	0.66 (0.45)	127.55	89.52	165.58	0.58 (0.94)
GFM_{wls}	11.51	18.32	0.64	1.23	0.00	0.63 (0.64)	113.00	71.53	154.47	0.56 (0.96)
					Machine Learning					
$\text{AGLASSO}_{\text{large}}$	12.09	16.93	0.53	1.14	0.00	0.71 (0.21)	91.60	89.77	93.43	0.62 (0.65)
$\text{IPCA}_{\text{K=5,large}}$	10.64	17.23	0.59	1.15	0.00	0.62 (0.73)	52.59	35.83	69.35	0.58 (0.92)
$\text{CA2}_{\text{K=5}}$	12.51	19.55	0.57	1.24	0.00	0.64 (0.57)	169.35	154.24	184.45	0.48 (0.43)
NN3	13.10	18.61	0.57	1.17	0.00	0.70 (0.25)	144.74	112.89	176.60	0.58 (0.95)
					Panel C: DCC-GARCH with Nonlinear Shrinkage					
					Return-based Estimators					
SAM_{mean}	8.77	18.52	0.19	1.16	0.00	0.47 (0.41)	77.39	52.64	102.15	0.42 (0.20)
GAL_{mean}	12.58	18.82	0.51	1.23	0.00	0.67 (0.51)	104.01	47.35	160.67	0.63 (0.69)
					Fama-MacBeth Regressions without/with the Galton Correction					
FM_{large}	14.91	20.17	0.48	1.30	0.00	0.74 (0.22)	202.93	172.00	233.85	0.58 (0.92)
FM_{wls}	14.80	19.89	0.52	1.30	0.00	0.74 (0.18)	199.48	168.31	230.66	0.60 (0.84)
$\text{GFM}_{\text{large}}$	14.45	20.38	0.53	1.34	0.00	0.71 (0.32)	173.16	121.18	225.13	0.61 (0.77)
GFM_{wls}	14.22	19.73	0.63	1.32	0.00	0.72 (0.27)	161.73	102.96	220.50	0.64 (0.61)
					Machine Learning					
$\text{AGLASSO}_{\text{large}}$	14.37	19.30	0.53	1.26	0.00	0.74 (0.22)	147.55	123.70	171.40	0.64 (0.62)
$\text{IPCA}_{\text{K=5,large}}$	13.52	18.80	0.59	1.26	0.00	0.72 (0.31)	115.31	70.29	160.34	0.66 (0.55)
$\text{CA2}_{\text{K=5}}$	14.02	19.29	0.59	1.28	0.00	0.73 (0.22)	198.70	166.11	231.29	0.57 (0.97)
NN3	15.12	18.89	0.58	1.22	0.00	0.80 (0.10)	176.85	129.70	224.01	0.67 (0.44)

Table A9

OOS performance of reward-to-risk timing strategies with Galton or GARCH variances

The table presents the OOS performance of the reward-to-risk timing strategy of Kirby and Ostdiek (2012). The portfolios are constructed using the same investment universe and optimization inputs as the Markowitz portfolios. The expected returns are estimated by 10 different methods. The variances of stocks are estimated by two methods. In Panel B, we use the Galton variance. In Panel C, we use the conditional variance implied by the GARCH model. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\bar{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). The OOS period is from 1987:01 to 2020:12. TO_{first} ($\text{TO}_{\text{second}}$) denotes average portfolio turnover over the first (second) half of the OOS period. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between a strategy and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\bar{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\bar{\sigma}_{exp}}$	AS	BR	SR	TO	NSR
Panel A: Benchmark Portfolios									
VW	8.59	14.99			0.00	0.00	0.57	1.47	0.57
EW	9.44	16.41			45.74	0.00	0.58 (0.96)	7.58	0.56 (0.84)
Panel B: Galton Covariance Matrix									
Return-based Estimators									
SAM_{mean}	8.90	14.75	0.42	0.94	49.93	0.00	0.60 (0.54)	10.51	0.58 (0.79)
GAL_{mean}	8.96	14.24	0.73	0.97	46.44	0.00	0.63 (0.37)	6.26	0.62 (0.46)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	11.18	14.76	0.81	1.01	58.00	0.00	0.76 (0.02)	79.63	0.60 (0.66)
FM_{wls}	11.20	14.86	0.86	1.01	57.68	0.00	0.75 (0.02)	81.27	0.60 (0.68)
$\text{GFM}_{\text{large}}$	10.10	14.40	0.81	0.99	50.57	0.00	0.70 (0.08)	42.82	0.62 (0.53)
GFM_{wls}	10.08	14.46	0.95	0.99	50.19	0.00	0.70 (0.08)	43.14	0.62 (0.52)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	9.98	14.07	0.75	0.95	54.28	0.00	0.71 (0.03)	43.54	0.62 (0.43)
$\text{IPCA}_{\text{K=5,large}}$	9.33	13.95	0.89	0.97	50.53	0.00	0.67 (0.18)	24.47	0.61 (0.59)
$\text{CA2}_{\text{K=5}}$	11.27	14.91	0.96	1.00	56.36	0.00	0.76 (0.01)	80.99	0.58 (0.86)
NN3	10.40	14.53	0.85	0.99	52.39	0.00	0.72 (0.05)	53.65	0.60 (0.63)
Panel C: DCC-GARCH with Nonlinear Shrinkage									
Return-based Estimators									
SAM_{mean}	9.04	14.18	0.43	0.98	50.35	0.00	0.64 (0.20)	23.34	0.60 (0.61)
GAL_{mean}	9.05	13.58	0.73	1.01	46.92	0.00	0.67 (0.15)	21.07	0.63 (0.38)
Fama-MacBeth Regressions without/with the Galton Correction									
FM_{large}	11.03	14.13	0.80	1.06	58.10	0.00	0.78 (0.01)	80.88	0.62 (0.51)
FM_{wls}	11.06	14.25	0.86	1.06	57.79	0.00	0.78 (0.01)	82.56	0.62 (0.52)
$\text{GFM}_{\text{large}}$	10.05	13.76	0.81	1.04	50.96	0.00	0.73 (0.04)	46.99	0.64 (0.37)
GFM_{wls}	10.03	13.83	0.95	1.04	50.58	0.00	0.73 (0.04)	47.26	0.63 (0.36)
Machine Learning									
$\text{AGLASSO}_{\text{large}}$	10.11	13.56	0.77	1.00	54.51	0.00	0.75 (0.01)	48.84	0.64 (0.23)
$\text{IPCA}_{\text{K=5,large}}$	9.33	13.41	0.88	1.03	51.15	0.00	0.70 (0.10)	33.93	0.62 (0.50)
$\text{CA2}_{\text{K=5}}$	11.24	14.26	0.99	1.04	56.21	0.00	0.79 (0.01)	81.09	0.61 (0.57)
NN3	10.40	13.93	0.86	1.03	52.48	0.00	0.75 (0.02)	55.50	0.63 (0.40)

Table A10

OOS performance of TCM-RT portfolios (**N = 50**)

At the end of each calendar year, we select the **50** largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT portfolios constructed using combinations of 10 return forecasting methods with the covariance matrix implied by the single index model (Panel B) or the daily sample covariance matrix (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly. The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\hat{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below -100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between an optimized portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\hat{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
Panel A: Benchmark Portfolios											
VW	8.42	14.79			0.00	0.00	0.57	1.86	2.21	1.51	0.57
EW	8.72	14.79			0.00	0.00	0.59 (0.66)	6.55	7.33	5.77	0.58 (0.84)
Panel B: Covariance Matrix implied by the Single Index Model											
Return-based Estimators											
SAM_{mean}	11.85	24.00	0.26	1.70	-138.68	0.00	0.49 (0.72)	56.09	39.16	73.01	0.46 (0.62)
GAL_{mean}	16.87	26.65	0.64	1.91	-76.10	0.00	0.63 (0.74)	27.81	16.57	39.05	0.62 (0.77)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	17.09	22.48	0.45	1.60	-156.50	0.00	0.76 (0.38)	366.62	302.45	430.78	0.53 (0.84)
FM_{wls}	16.37	21.61	0.46	1.53	-157.80	0.00	0.76 (0.37)	366.70	296.03	437.37	0.53 (0.82)
GFM_{large}	19.92	25.60	0.68	1.82	-110.72	0.00	0.78 (0.28)	206.90	145.37	268.43	0.69 (0.53)
GFM_{wls}	19.75	25.42	0.81	1.81	-109.30	0.00	0.78 (0.29)	198.60	111.69	285.51	0.71 (0.46)
Machine Learning											
$AGLASSO_{large}$	18.49	24.04	0.53	1.73	-142.07	0.00	0.77 (0.37)	167.14	157.15	177.12	0.66 (0.68)
$IPCA_{K=5,large}$	17.93	25.52	0.67	1.82	-113.44	0.00	0.70 (0.50)	84.04	74.28	93.80	0.66 (0.66)
$CA2_{K=5}$	18.64	21.94	0.63	1.56	-148.10	0.00	0.85 (0.18)	339.66	267.03	412.29	0.65 (0.70)
NN3	19.48	22.89	0.65	1.62	-130.78	0.00	0.85 (0.14)	239.70	154.32	325.09	0.74 (0.38)
Panel C: Sample Covariance Matrix											
Return-based Estimators											
SAM_{mean}	6.57	20.05	0.14	1.12	-214.73	0.00	0.33 (0.30)	79.86	37.40	122.31	0.29 (0.23)
GAL_{mean}	10.90	16.39	0.67	1.06	-84.97	0.00	0.67 (0.54)	17.02	6.57	27.47	0.66 (0.56)
Fama-MacBeth Regressions without/with the Galton Correction											
FM_{large}	15.96	18.91	0.40	1.06	-268.82	0.00	0.84 (0.21)	563.26	348.32	778.19	0.53 (0.85)
FM_{wls}	15.50	18.79	0.42	1.05	-263.84	0.00	0.82 (0.24)	562.02	337.86	786.18	0.52 (0.81)
GFM_{large}	16.34	18.31	0.66	1.05	-200.38	0.00	0.89 (0.09)	390.72	188.67	592.76	0.73 (0.41)
GFM_{wls}	15.38	17.53	0.76	1.04	-192.16	0.00	0.88 (0.10)	361.85	138.12	585.58	0.75 (0.34)
Machine Learning											
$AGLASSO_{large}$	12.35	18.42	0.36	1.10	-220.68	0.00	0.67 (0.67)	249.51	194.43	304.59	0.50 (0.77)
$IPCA_{K=5,large}$	13.96	17.20	0.68	1.05	-185.10	0.00	0.81 (0.24)	133.81	77.24	190.38	0.74 (0.40)
$CA2_{K=5}$	15.28	18.91	0.48	1.07	-272.61	0.00	0.81 (0.29)	536.28	307.70	764.86	0.54 (0.86)
NN3	15.91	19.36	0.54	1.09	-248.75	0.00	0.82 (0.21)	448.04	197.19	698.90	0.64 (0.72)

Table A11

OOS performance of TCM-RT portfolios (Galton and GARCH covariances, $\mathbf{N = 50}$)

At the end of each calendar year, we select the **50** largest common stocks with a full return history over the past 60 months and a full return history over the subsequent 12 months. The sample is kept fixed for the following 12 months. Panel A presents the OOS performance of two benchmark portfolios: the value-weighted portfolio (VW) and the equal-weighted portfolio (EW). Panels B and C present the OOS performance of TCM-RT portfolios constructed using combinations of 10 return forecasting methods with the Galton covariance matrix of Barroso and Saxena (2022) (Panel B) or the DCC-GARCH estimator of Engle, Ledoit, and Wolf (2019) (Panel C). TCM-RT portfolios refer to portfolios constructed using mean-variance optimization with transaction cost management and risk targeting. All portfolios are rebalanced monthly.The columns present descriptive statistics of the OOS performance of each strategy: average realized excess return ($\hat{\mu}$), realized standard deviation ($\hat{\sigma}$), ratio of average realized returns to average expected returns ($\frac{\hat{\mu}}{\hat{\mu}_{exp}}$), ratio of realized volatility to average exepcted volatility ($\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$), the sum of negative weights (SNW), active share (AS), bankruptcy rate (i.e., percentage of months with before-cost returns below −100%, BR), Sharpe ratio (SR), portfolio turnover (TO) and net-of-costs Sharpe ratio (NSR). TO1 (TO2) denotes average portfolio turnover over the 1st (2nd) half of the OOS period. The OOS period is from 1987:01 to 2020:12. A two-sided p-value for the null hypothesis of equal Sharpe ratios or equal net-of-costs Sharpe ratios between an optimized portfolio and the value-weighted portfolio is reported in parentheses.

	$\hat{\mu}$	$\hat{\sigma}$	$\frac{\hat{\mu}}{\hat{\mu}_{exp}}$	$\frac{\hat{\sigma}}{\hat{\sigma}_{exp}}$	SNW	BR	SR	TO	TO1	TO2	NSR
					Panel A: Benchmark Portfolios						
VW	8.42	14.79			0.00	0.00	0.57	1.86	2.21	1.51	0.57
EW	8.72	14.79			0.00	0.00	0.59 (0.66)	6.55	7.33	5.77	0.58 (0.84)
					Panel B: Galton Covariance Matrix						
					Return-based Estimators						
SAM_{mean}	6.46	15.76	0.18	0.90	-119.50	0.00	0.41 (0.43)	46.79	22.74	70.84	0.38 (0.36)
GAL_{mean}	8.74	12.98	0.64	0.87	-24.08	0.00	0.67 (0.41)	7.86	4.63	11.09	0.67 (0.42)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	12.21	14.41	0.43	0.83	-147.93	0.00	0.85 (0.18)	343.01	236.52	449.49	0.57 (1.00)
FM_{wls}	11.83	14.20	0.44	0.82	-147.68	0.00	0.83 (0.19)	338.51	229.57	447.45	0.56 (0.97)
GFM_{large}	11.97	13.78	0.66	0.86	-90.72	0.00	0.87 (0.07)	211.48	120.69	302.27	0.73 (0.33)
GFM_{wls}	11.32	13.12	0.77	0.87	-84.63	0.00	0.86 (0.08)	190.45	88.05	292.86	0.76 (0.24)
					Machine Learning						
$AGLASSO_{large}$	9.72	13.68	0.39	0.89	-106.90	0.00	0.71 (0.53)	147.90	126.80	168.99	0.57 (0.99)
$IPCA_{K=5,large}$	10.29	12.73	0.69	0.85	-81.31	0.00	0.81 (0.18)	76.12	50.65	101.60	0.75 (0.32)
$CA2_{K=5}$	12.57	14.62	0.56	0.86	-139.27	0.00	0.86 (0.14)	319.18	209.51	428.86	0.62 (0.78)
NN3	13.41	14.62	0.63	0.86	-120.33	0.00	0.92 (0.04)	245.45	129.03	361.86	0.77 (0.25)
					Panel C: DCC-GARCH with Nonlinear Shrinkage						
					Return-based Estimators						
SAM_{mean}	5.62	19.18	0.12	1.17	-205.38	0.00	0.29 (0.20)	156.62	77.43	235.81	0.22 (0.11)
GAL_{mean}	12.30	18.97	0.68	1.28	-99.15	0.00	0.65 (0.63)	138.70	28.09	249.32	0.62 (0.77)
					Fama-MacBeth Regressions without/with the Galton Correction						
FM_{large}	13.23	19.11	0.37	1.19	-238.96	0.00	0.69 (0.57)	494.18	309.65	678.71	0.42 (0.47)
FM_{wls}	13.30	18.54	0.40	1.16	-232.01	0.00	0.72 (0.48)	493.91	301.86	685.96	0.45 (0.53)
GFM_{large}	15.38	19.02	0.65	1.23	-173.77	0.00	0.81 (0.19)	352.28	178.10	526.46	0.66 (0.61)
GFM_{wls}	14.75	18.12	0.75	1.20	-166.79	0.00	0.81 (0.16)	334.33	138.22	530.43	0.69 (0.47)
					Machine Learning						
$AGLASSO_{large}$	12.82	18.65	0.38	1.17	-202.29	0.00	0.69 (0.59)	280.00	204.79	355.20	0.51 (0.79)
$IPCA_{K=5,large}$	13.76	18.51	0.65	1.25	-173.87	0.00	0.74 (0.38)	192.48	89.79	295.17	0.66 (0.63)
$CA2_{K=5}$	12.41	18.41	0.44	1.18	-242.61	0.00	0.67 (0.62)	467.35	272.56	662.13	0.43 (0.50)
NN3	12.86	18.31	0.47	1.16	-221.03	0.00	0.70 (0.50)	406.82	183.40	630.24	0.53 (0.82)

References

Abdi, F., and A. Ranaldo. 2017. A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies* 30:4437–80.

Agrawal, A., R. Verschueren, S. Diamond, and S. Boyd. 2018. A rewriting system for convex optimization problems. *Journal of Control and Decision* 5:42–60.

Avramov, D., S. Cheng, and L. Metzker. 2022. Machine learning vs. economic restrictions: Evidence from stock return predictability. *Management Science* .

Bai, J., and S. Shi. 2011. Estimating high dimensional covariance matrices and its applications .

Barroso, P., and A. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140:744–67.

Barroso, P., and K. Saxena. 2022. Lest we forget: Learn from out-of-sample forecast errors when optimizing portfolios. *The Review of Financial Studies* 35:1222–78.

Bessembinder, H., M. J. Cooper, and F. Zhang. 2018. Characteristic-based benchmark returns and corporate events. *The Review of Financial Studies* 32:75–125. ISSN 0893-9454. doi:10.1093/rfs/hhy037.

Bessler, W., and D. Wolff. 2015. Do commodities add value in multi-asset portfolios? an out-of-sample analysis for different investment strategies. *Journal of Banking & Finance* 60:1–20.

Best, M. J., and R. R. Grauer. 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The Review of Financial Studies* 4:315–42.

Bianchi, D., M. Büchner, T. Hoogteijling, and A. Tamoni. 2021. Corrigendum: Bond risk premiums with machine learning. *The Review of Financial Studies* 34:1090–103.

Black, F. 1972. Capital market equilibrium with restricted borrowing. *The Journal of Business* 45:444–55.

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies* 22:3411–47.

Chan, L. K., J. Karceski, and J. Lakonishok. 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *The Review of Financial Studies* 12:937–74.

Chen, L., M. Pelger, and J. Zhu. 2020. Deep learning in asset pricing. *Available at SSRN 3350138* .

Chen, P., E. Lezmi, T. Roncalli, and J. Xu. 2019. A note on portfolio optimization with quadratic transaction costs. *Available at SSRN 3683466* .

Chernov, M., M. Dahlquist, and L. A. Lochstoer. 2020. Pricing currency risks. *Available at SSRN 3753971* .

Chopra, V. K., and W. T. Ziemba. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management* 19:6–11.

Chung, K. H., and H. Zhang. 2014. A simple approximation of intraday spreads using daily data. *Journal of Financial Markets* 17:94–120.

Cochrane, J. H., and J. Saa-Requejo. 2000. Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of political economy* 108:79–119.

Daniel, K., D. Hirshleifer, and L. Sun. 2020. Short-and long-horizon behavioral factors. *The review of financial studies* 33:1673–736.

Davis, M. H., and A. R. Norman. 1990. Portfolio selection with transaction costs. *Mathematics of operations research* 15:676–713.

De Nard, G., O. Ledoit, and M. Wolf. 2021. Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics* 19:236–57.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies* 22:1915–53.

DeMiguel, V., A. Martin-Utrera, and F. J. Nogales. 2013. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking & Finance* 37:3018–34.

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal. 2020. A transaction-cost perspective on the multitude of firm characteristics. *The Review of Financial Studies* 33:2180–222.

DeMiguel, V., and A. Olivares-Nadal. 2018. Technical note—a robust perspective on transaction costs in portfolio optimization. *Operations Research* .

Diamond, S., and S. Boyd. 2016. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research* 17:1–5.

Drechsler, I., and Q. F. S. Drechsler. 2018. The shorting premium and asset pricing anomalies. *The Shorting Premium and Asset Pricing Anomalies (May 4, 2018)* .

Engle, R. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20:339–50.

Engle, R. F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of The Econometric Society* 987–1007.

Engle, R. F., O. Ledoit, and M. Wolf. 2019. Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37:363–75.

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

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

———. 2018. Choosing factors. *Journal of Financial Economics* 128:234–52.

———. 2020. Comparing cross-section and time-series factor models. *The Review of Financial Studies* 33:1891–926.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

Fan, J., Y. Fan, and J. Lv. 2008. High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics* 147:186–97.

Fan, J., Y. Liao, and M. Mincheva. 2013. Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society. Series B, Statistical methodology* 75.

Filippou, I., D. Rapach, M. P. Taylor, and G. Zhou. 2022. Out-of-sample exchange rate prediction: A machine learning perspective. *Available at SSRN 3455713* .

Freyberger, J., A. Neuhierl, and M. Weber. 2020. Dissecting characteristics nonparametrically. *The Review of Financial Studies* 33:2326–77.

Green, J., J. R. M. Hand, and X. F. Zhang. 2017. The characteristics that provide independent information about average u.s. monthly stock returns. *The Review of Financial Studies* 30:4389–436. ISSN 0893-9454. doi:10.1093/rfs/hhx019.

Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33:2223–73.

———. 2021. Autoencoder asset pricing models. *Journal of Econometrics* 222:429–50.

Hafner, C. M., and O. Reznikova. 2012. On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis* 56:3533–45.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. … and the cross-section of expected returns. *The Review of Financial Studies* 29:5–68.

Hasbrouck, J. 2009. Trading costs and returns for us equities: Estimating effective costs from daily data. *The Journal of Finance* 64:1445–77.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction.* Springer Science & Business Media.

Haugen, R. A., and N. L. Baker. 1996. Commonality in the determinants of expected stock returns. *Journal of Financial Economics* 41:401–39.

Hautsch, N., and S. Voigt. 2019. Large-scale portfolio allocation under transaction costs and model uncertainty. *Journal of Econometrics* 212:221–40.

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

Huang, J., J. L. Horowitz, and F. Wei. 2010. Variable selection in nonparametric additive models. *The Annals of Statistics* 38:2282 – 2313. doi:10.1214/09-AOS781.

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58:1651–83.

James, W., C. Stein, et al. 1961. Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California.

Jensen, T. I., B. T. Kelly, S. Malamud, and L. H. Pedersen. 2022. Machine learning and the implementable efficient frontier. *Available at SSRN 4187217* .

Jobson, J. D., and B. Korkie. 1980. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association* 75:544–54.

Jorion, P. 1986. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative analysis* 21:279–92.

Kan, R., and D. R. Smith. 2008. The distribution of the sample minimum-variance frontier. *Management Science* 54:1364–80.

Kelly, B. T., S. Pruitt, and Y. Su. 2019. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134:501–24.

Kirby, C., and B. Ostdiek. 2012. It's all in the timing: simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* 47:437–67.

Kozak, S., S. Nagel, and S. Santosh. 2018. Interpreting factor models. *The Journal of Finance* 73:1183–223.

Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10:603–21.

———. 2004. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management* 30:110–9.

———. 2008. Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance* 15:850–9.

———. 2011. Robust performances hypothesis testing with the variance. *Wilmott* 2011:86–9.

———. 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics* 40:1024–60.

———. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies* 30:4349–88.

———. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics* 48:3043–65.

Lewellen, J. 2015. The cross-section of expected stock returns. *Critical Finance Review* 4:1–44.

Lintner, J. 1965. Security prices, risk, and maximal gains from diversification. *The Journal of Finance* 20:587–615.

Magill, M. J., G. M. Constantinides, et al. 1976. Portfolio selection with transaction costs. *Journal of Economic Theory* 13:245–63.

Markowitz, H. 1952. Portfolio selection. *The Journal of Finance* 7:77–91. ISSN 00221082, 15406261.

Michaud, R. O. 1989. The markowitz optimization enigma: Is 'optimized'optimal? *Financial Analysts Journal* 45:31–42.

MOSEK, A. 2022. *Optimizer api for python 9.3.13*.

Nesterov, Y. E., and M. J. Todd. 1997. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations research* 22:1–42.

Novy-Marx, R., and M. Velikov. 2016. A taxonomy of anomalies and their trading costs. *The Review of Financial Studies* 29:104–47.

Okhrin, Y., and W. Schmid. 2006. Distributional properties of portfolio weights. *Journal of Econometrics* 134:235–56.

———. 2007. Comparison of different estimation techniques for portfolio selection. *AStA Advances in Statistical Analysis* 91:109–27.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle. 2020. Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics* 1–17.

Roll, R. 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance* 39:1127–39.

Schwarz, G., et al. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–4.

Sharpe, W. F. 1963. A simplified model for portfolio analysis. *Management science* 9:277–93.

———. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19:425–42.

Stambaugh, R. F., and Y. Yuan. 2017. Mispricing factors. *The review of financial studies* 30:1270–315.

Stein, C., et al. 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California.

Taksar, M., M. J. Klass, and D. Assaf. 1988. A diffusion model for optimal portfolio selection in the presence of brokerage fees. *Mathematics of operations research* 13:277–94.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58:267–88.

Xiong, S., B. Dai, J. Huling, and P. Z. Qian. 2016. Orthogonalizing em: A design-based least squares algorithm. *Technometrics* 58:285–93.

Yuan, M., and Y. Lin. 2006. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68:49–67.