

Optimizing the Performance of Sample Mean-Variance Efficient Portfolios

Chris Kirby^a, Barbara Ostdiek^b

^a*Belk College of Business, University of North Carolina at Charlotte*

^b*Jones Graduate School of Business, Rice University*

Abstract

We propose a comprehensive empirical strategy for optimizing the out-of-sample performance of sample mean-variance efficient portfolios. After constructing a sample objective function that accounts for the impact of estimation risk, specification errors, and transaction costs on portfolio performance, we maximize the function with respect to a set of tuning parameters to obtain plug-in estimates of the optimal portfolio weights. The methodology offers considerable flexibility in specifying objectives, constraints, and modeling techniques. Moreover, the resulting portfolios have well-behaved weights, reasonable turnover, and substantially higher Sharpe ratios and certainty-equivalent returns than benchmarks such as the $1/N$ portfolio and S&P 500 index.

Key words: active management, conditioning information, estimation risk, mean-variance optimization, portfolio choice, turnover

JEL classification: G11; G12; C11

July 23, 2012

First draft: April 24, 2011

* Comments welcome. We thank Dejan Suskavcevic for excellent research assistance. A previous version of the paper was circulated under the title ‘Optimal Active Portfolio Management with Unconditional Mean-Variance Risk Preferences.’ Address correspondence to: Chris Kirby, Department of Finance, Belk College of Business Administration, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-0001.

1 Introduction

Empirical research on mean-variance portfolio optimization is typically conducted by substituting estimates of the mean vector and covariance matrix of asset returns into an expression for the optimal portfolio weights. The portfolios constructed using this “plug-in” approach are called sample mean-variance efficient portfolios. Although the plug-in approach is conceptually straightforward, a number of implementation issues arise that fall outside the scope of the optimization problem. These range from how to model changes in the investment opportunity set and estimate the model parameters, to how to account for transaction costs. In this paper, we propose a methodology that encompasses the plug-in approach within a broader empirical framework that fully accounts for the impact of such issues on the performance of the plug-in weights. This allows us to optimize the out-of-sample performance of sample mean-variance efficient portfolios with respect to specific investment objectives.

Our general empirical strategy is motivated by Skouras (2007). He suggests a decision-theoretic approach for estimating the parameters of any sufficiently regular rule that maps realizations of one or more random variables into decisions made by an economic agent. The parameter estimates are obtained by maximizing an objective function that measures the economic benefit to the agent of following the given rule. In other words, the approach is designed to optimize the performance of the rule from an economic perspective. Brandt et al. (2009) provide a nice example of this approach applied in a portfolio choice setting. They investigate a parametric portfolio choice rule that restricts the portfolio weights to be linear in a set of asset-specific variables, such as size and book-to-market measures. To implement the rule, they find the coefficients on the asset-specific variables that deliver the highest average utility for a specified historical sample period.

We use a similar methodology to optimize the performance of the plug-in approach. The basic strategy is as follows. First, we note that Ferson and Siegel (2001) derive an analytic expression for the weights that deliver an unconditionally mean-variance efficient (UMVE) portfolio in settings with time-varying investment opportunities. This establishes the portfolio rebalancing rule that is optimal in the absence of estimation risk and rebalancing costs. Next, we use historical asset returns to construct plug-in weights that depend on a small number of unknown parameters. This yields theoretically-motivated counterparts of the linear weight functions used by Brandt et al. (2009). Finally, we use the plug-in weights to generate a series of out-of-sample portfolio returns and find the values of the parameters that maximize the mean return subject to a constraint on the return variance. Thus we fully account for the impact of estimation risk when choosing the parameter values, and we can take turnover into account simply by using returns measured net of rebalancing costs.

Ferson and Siegel (2001) show that the weights that deliver an UMVE portfolio are a function of the conditional mean vector and conditional second-moment matrix of asset returns. Consistent with most studies in the portfolio choice literature, we focus on the case in which the conditioning information consists solely of past asset returns. For the empirical investigation, we estimate the conditional moments of returns via simple exponential-smoothing models that emphasize parsimony and impose minimal assumptions about the data generating process. Of course, using estimates in place of the true conditional moments introduces

estimation risk, i.e., uncertainty about portfolio returns that is incremental to the usual uncertainty about the individual asset returns. The presence of estimation risk complicates the portfolio choice problem because it affects the time-series properties of the plug-in weights and, hence, the expected portfolio performance (see, e.g., Kan and Zhou, 2007).

To illustrate the nature of the complications, consider a scenario in which all investors form conditional expectations via exponential smoothing and use the same smoothing constant. Even if we use an exponential-smoothing model to estimate the conditional moments of asset returns, the estimated optimal portfolio weights will generally deviate from the true weights because the unknown value of the smoothing constant must either be specified *a priori* or estimated from the data. In either case, the potential for choosing an incorrect smoothing constant causes the expected performance of the sample mean-variance efficient portfolio to fall short of the expected performance of the true optimal portfolio. Specification errors give rise to similar issues. If the models used to estimate the conditional moments of returns are misspecified, the plug-in weights are not consistent estimates of the true optimal weights. The result is a decline in expected portfolio performance.

From an empirical point of view, the challenge is to minimize the adverse impact of estimation risk and specification errors on portfolio performance. This argues for the use of specialized econometric methods. Under a conventional econometric approach, choosing the smoothing constants is a model-fitting problem. We might select the values that deliver the best forecasts of the returns and squared returns under mean-squared-error (MSE) loss. It is clear, however, that the model-fitting problem does not embody the same objective as the portfolio problem, which is to maximize expected utility under mean-variance risk preferences. Expected utility functions generally translate into asymmetric loss functions, and asymmetric loss favors estimates that are biased in an appropriate direction (Patton and Timmermann, 2007). The values of the smoothing constants that deliver the best-performing portfolio could be quite different from the values that minimize the MSE of the forecasts.

This insight lies at the core of our methodology. In effect, we treat any unknown parameter that influences the time-series properties of the plug-in weights as a tuning parameter, i.e., as a parameter that can be freely changed to tailor portfolio performance to specific constraints and objectives. In our framework, the goal is to maximize the unconditional expected return on the portfolio subject to a constraint on the unconditional return variance. We therefore select the model parameters based on the sample moments of the sequence of out-of-sample portfolio returns that result from using historical data to construct a time series of plug-in portfolio weights. Specifically, we find the parameter values that generate the highest average realized utility under mean-variance risk preferences. This optimizes the out-of-sample performance of the portfolio.

The proposed approach is not restrictive in terms of either modeling techniques or constraints on portfolio holdings. For example, we consider sample UMVE portfolios constructed using shrinkage estimators of the conditional moments of returns. The shrinkage factor is therefore included as an additional tuning parameter to be estimated in conjunction with other model parameters. We also address issues of portfolio turnover and rebalancing costs by explicitly accounting for the effect of turnover in the tuning-parameter optimization. This is accomplished by using returns measured net of rebalancing costs to construct the sam-

ple objective function. The approach can easily be extended to incorporate techniques for reducing turnover and the attendant rebalancing costs.

For example, Leland (1999) argues that partial-adjustment strategies are the appropriate way to deal with rebalancing costs. These strategies recognize that costly trading can make it inefficient to fully adjust to the estimated optimal weights each period. We develop a partial-adjustment strategy that defines a no-trade region around the estimated optimal weights each period using an estimate of the conditional expected utility loss from leaving the weights unchanged. If this loss is less than some cutoff, no adjustment is made; otherwise the existing weights are adjusted to the no-trade boundary. We estimate the optimal size of the no-trade region by including the cutoff in the set of tuning parameters.

We evaluate the effectiveness of the proposed methodology for three datasets that contain monthly returns on equally-weighted U.S. equity portfolios. Using portfolios as assets instead of individual stocks is common in research on mean-variance optimization. We do so because it allows us to assess the potential for exploiting well-known empirical regularities such as value, growth, and momentum effects. The first two datasets are the Fama-French 10 Industry portfolios and 25 Size/Book-to-Market portfolios. To construct the third dataset, we sort NYSE, AMEX, and NASDAQ firms into 30 Momentum/Volatility portfolios using their past returns and average absolute returns. The sample period is January 1946 to December 2009. We reserve the first 360 months of data to construct the plug-in weights for the initial investment period, leaving 408 months for performance evaluation.

To generate our empirical results, we maximize average realized utility under mean-variance risk preferences using a relative risk aversion of 15. This level of risk aversion imposes a substantial risk penalty, producing a relatively conservative investment style. The analysis reveals that our methodology performs well along a number of dimensions. For instance, if we estimate the optimal values of the tuning parameters and measure portfolio performance assuming proportional transaction costs of 50 basis points (bp), then the sample UMVE portfolio for the 10 Industry dataset has an estimated Sharpe ratio of 0.94. In comparison, the S&P 500 index and the $1/N$ portfolio have estimated Sharpe ratios of 0.41 and 0.54. The performance advantage of the sample UMVE portfolio is highly statistically significant and points to substantial benefits from employing mean-variance optimization.

This finding is tempered, however, by the high level of turnover required to realize these benefits. It averages over 380% per year. In an effort to reduce the turnover of the portfolio, we explore the use of partial adjustment techniques. This meets with only modest success for this dataset. Allowing for partial adjustment of the weights decreases the average turnover of the sample UMVE portfolio by only about 45 percentage points per year. The high turnover is probably related to the low cross-sectional dispersion in average returns for the 10 Industry dataset — a characteristic not shared with the other two datasets. Because there is little cross-sectional variation in the sample means, the performance gains may largely reflect the success of the optimizer in exploiting, by what turns out to be aggressive rebalancing, the time-series variation in conditional expected returns.

The results for the 25 Size/Book-to-Market and 30 Momentum/Volatility datasets are consistent with this hypothesis in the sense that turnover is of less concern. Under the same 50 bp

transaction costs assumption, the sample UMVE portfolios for these datasets have estimated Sharpe ratios of 0.95 and 1.46, respectively. The average turnover for the 25 Size/Book-to-Market dataset is 286% per year, which is still relatively high. But the average turnover is considerably lower for the 30 Momentum/Volatility dataset: 158% per year. With partial adjustment of the weights, these figures fall to 137% and 75%, respectively. This drop in average turnover is accompanied by an increase in the estimated Sharpe ratios of the sample UMVE portfolios to 1.03 and 1.54. In comparison, the $1/N$ portfolio generates an estimated Sharpe ratio of about 0.57 for both datasets.

Taking the effort to reduce turnover a step further, we combine partial adjustment of the weights with the use of shrinkage estimators while simultaneously imposing a long-only constraint. We find that prohibiting short sales leads to a considerable reduction in the estimated Sharpe ratios. For example, the long-only sample UMVE portfolio for the 30 Momentum/Volatility dataset has an estimated Sharpe ratio of 1.04. It is noteworthy, however, that this portfolio outperforms all of the benchmarks at the 1% significance level, and it does so despite having an average turnover of only 8% per year. Thus prohibiting short sales is an effective strategy for sharply reducing turnover, and it allows the sample UMVE portfolios to maintain a significant performance advantage over the benchmarks.

Overall the empirical evidence suggests that the proposed methodology leads to robust portfolio selection rules. It achieves robustness by expanding the scope of the optimization problem to encompass the effects of estimation risk, specification errors, and transaction costs on portfolio performance, and using an adaptive empirical strategy to select the values of the unknown parameters that appear in the expression for the plug-in weights. This is in contrast to the *ad hoc* strategies for selecting the values of these parameters considered elsewhere in the literature. Although researchers have long recognized that the performance of mean-variance optimization is sensitive to the choice of parameter values, there is a dearth of research on choosing these values in a robust fashion. The empirical strategy developed here represents a significant step forward in this regard.

2 Methodology for Optimizing the Performance of Sample UMVE Portfolios

The analysis is framed in terms of the portfolio problem of an investor who wants to rebalance his portfolio on a regular basis to take advantage of time-varying investment opportunities. We assume that the rebalancing process is accomplished using a formal rule that specifies how the portfolio weights respond to changes in the investment opportunity set. To identify the optimal rebalancing rule, we have to specify a well-defined investment objective. We analyze the case in which the objective is to maximize the unconditional expected return of the portfolio subject to a constraint on the unconditional portfolio variance. That is, we assume the goal is to construct a UMVE portfolio.

Although it is somewhat unusual to specify an unconditional investment objective in a setting with time-varying investment opportunities, this is consistent with the common practice of unconditional performance evaluation. Mutual fund ratings, for example, are largely determined by fund performance over an extended interval such as three or five years. Portfolios constructed using conditional objectives may fare poorly if their performance is evaluated

from an unconditional perspective (Dybvig and Ross, 1985). Of course, unconditional optimization is a special case of conditional optimization in our framework, because every UMVE portfolio is also conditionally mean-variance efficient (Hansen and Richard, 1987). We emphasize the unconditional representation of the portfolio problem because unconditional optimization with respect to a set of tuning parameters plays a key role in the analysis.

2.1 UMVE portfolios with time-varying investment opportunities

Ferson and Siegel (2001) provide a general characterization of the set of active portfolio strategies that deliver minimum-variance portfolios in the presence of time-varying investment opportunities. Suppose for illustration purposes that there are N risky assets. Let r_{t+1} denote the $N \times 1$ vector of asset returns for period $t + 1$ and let $r_{p,t+1} = w_t' r_{t+1}$ denote the portfolio return for period $t + 1$, where w_t is an $N \times 1$ vector of weights selected in period t that sum to 1. Ferson and Siegel (2001) show that the weights that produce the minimum value of $\sigma_p^2 = \text{Var}(r_{p,t+1})$ for a given value of $\mu_p = \text{E}[r_{p,t+1}]$ are of the form

$$w_t = \frac{\Omega_t^{-1} \iota}{\iota' \Omega_t^{-1} \iota} + \frac{\mu_p - \mu_{p_0}}{\mu_{p_1} - \mu_{p_0}} \left(\Omega_t^{-1} - \frac{\Omega_t^{-1} \iota \iota' \Omega_t^{-1}}{\iota' \Omega_t^{-1} \iota} \right) \mu_t \quad \forall t, \quad (1)$$

where $\mu_t = \text{E}_t[r_{t+1}]$ is the conditional mean vector of returns, $\Omega_t = \text{E}_t[r_{t+1} r_{t+1}']$ is the condition second moment matrix of returns,

$$\mu_{p_0} = \text{E} \left[\frac{\iota' \Omega_t^{-1} \mu_t}{\iota' \Omega_t^{-1} \iota} \right], \quad (2)$$

$$\mu_{p_1} = \text{E} \left[\mu_t' \Omega_t^{-1} \mu_t + (1 - \iota' \Omega_t^{-1} \mu_t) \frac{\iota' \Omega_t^{-1} \mu_t}{\iota' \Omega_t^{-1} \iota} \right], \quad (3)$$

and ι is an $N \times 1$ vector of ones. The scalars μ_{p_0} and μ_{p_1} denote the expected returns of two portfolios on the minimum-variance frontier: the portfolio with the minimum value of $\text{E}_t[r_{p,t+1}^2]$ and the portfolio with the maximum value of $\text{E}_t[r_{p,t+1}] - (1/2)\text{E}_t[r_{p,t+1}^2]$.

Equation (1) implies that we can construct the minimum-variance portfolio for a target expected return of μ_p by investing a fraction of wealth, $x_p = (\mu_p - \mu_{p_0}) / (\mu_{p_1} - \mu_{p_0})$, in the frontier portfolio with expected return μ_{p_1} and the remainder in the frontier portfolio with expected return μ_{p_0} . This construction is not unique because any two frontier portfolios span the entire minimum-variance frontier (Hansen and Richard, 1987). However, it is the only construction for which the weights of the two spanning portfolios can be expressed in terms of μ_t and Ω_t alone, i.e., without the use of any scaling constants.

Setting $\mu_p \geq \mu_{p_0} / (1 - \mu_{p_1} + \mu_{p_0})$ delivers a UMVE portfolio (Ferson and Siegel, 2001). To see this, consider the problem of choosing w_t to maximize the quadratic objective function

$$Q_p(w_t) = \text{E}[w_t' r_{t+1}] - \frac{\gamma}{2} \text{Var}(w_t' r_{t+1}), \quad (4)$$

where $\gamma > 0$. The solution clearly delivers a UMVE portfolio because it maximizes $\text{E}[r_{p,t+1}]$ for some value of $\text{Var}(r_{p,t+1})$. Moreover, the solution must be of the form shown in equation

(1) because maximizing $Q_p(w_t)$ subject to $E[r_{p,t+1}] = \mu_p$ also minimizes $\text{Var}(r_{p,t+1})$ subject to $E[r_{p,t+1}] = \mu_p$.

Substituting equation (1) into equation (4) and applying the law of iterated expectations yields the concentrated objective function

$$Q_p(\mu_p) = \mu_p - \frac{\gamma}{2} \left(E[(\iota' \Omega_t^{-1} \iota)^{-1}] + \frac{(\mu_p - \mu_{p_0})^2}{\mu_{p_1} - \mu_{p_0}} - \mu_p^2 \right) \quad (5)$$

with μ_p as the choice variable. Using equation (5) we find that $Q_p(w_t)$ is maximized for

$$\mu_p = \frac{\mu_{p_0}}{1 - \mu_{p_1} + \mu_{p_0}} + \frac{1}{\gamma} \left(\frac{\mu_{p_1} - \mu_{p_0}}{1 - \mu_{p_1} + \mu_{p_0}} \right). \quad (6)$$

Hence, we have $\mu_p \geq \mu_{p_0}/(1 - \mu_{p_1} + \mu_{p_0})$ for a UMVE portfolio.

In the subsequent analysis we exploit the close connection between the problem of finding a UMVE portfolio and the problem of maximizing expected utility under quadratic risk preferences. To see the connection, suppose someone with utility of the form

$$U(w_t) = w_t' r_{t+1} - \frac{\psi}{2} (w_t' r_{t+1})^2 \quad (7)$$

wants to maximize $E[U(w_t)]$. Because maximizing $E[U(w_t)]$ subject to $E[r_{p,t+1}] = \mu_p$ is equivalent to minimizing $\text{Var}(r_{p,t+1})$ subject to $E[r_{p,t+1}] = \mu_p$, it again follows that the solution must be of the form shown in equation (1).

Substituting equation (1) into equation (7) and applying the law of iterated expectations yields the concentrated objective function

$$E[U(\mu_p)] = \mu_p - \frac{\psi}{2} \left(E[(\iota' \Omega_t^{-1} \iota)^{-1}] + \frac{(\mu_p - \mu_{p_0})^2}{\mu_{p_1} - \mu_{p_0}} \right), \quad (8)$$

which is maximized for $\mu_p = \mu_{p_0} + (\mu_{p_1} - \mu_{p_0})/\psi$. Hence, we obtain

$$w_t = \frac{\Omega_t^{-1} \iota}{\iota' \Omega_t^{-1} \iota} + \frac{1}{\psi} \left(\Omega_t^{-1} - \frac{\Omega_t^{-1} \iota \iota' \Omega_t^{-1}}{\iota' \Omega_t^{-1} \iota} \right) \mu_t \quad (9)$$

as the optimal vector of weights. This is the same vector of weights that maximizes

$$E_t[U(w_t)] = w_t' \mu_t - \frac{\psi}{2} w_t' \Omega_t w_t \quad (10)$$

subject to $w_t' \iota = 1$. Setting $\psi = \gamma(1 - \mu_{p_1} + \mu_{p_0})/(1 + \gamma\mu_{p_0})$ delivers the UMVE portfolio for a given value of γ . Note that we have $\psi < \gamma$ because $\mu_{p_1} > \mu_{p_0} > 0$.

2.2 Plug-in estimation of the optimal weights

Equation (1) is derived under the assumption that the values of μ_{p_0} , μ_{p_1} , μ_t , and Ω_t are known. Because this assumption is not satisfied in practice, we follow the related empirical literature

by using the plug-in approach to implement the Ferson and Siegel (2001) methodology. That is, we use historical data to estimate the unknown parameters, and substitute the parameter estimates into the formula for the optimal portfolio weights. Using estimates in place of the population parameters entails estimation risk: uncertainty about portfolio returns that is incremental to the uncertainty about individual asset returns.¹ It is important, therefore, to consider the impact of this risk on portfolio performance.

The empirical investigation focuses on the case in which the conditioning information consists solely of historical returns. We refer to the sample of returns ending at T_0 as the initial “holdout” sample. To use this sample to construct plug-in weights for the interval T_0 to $T_0 + 1$, we must first specify estimators for μ_{T_0} and Ω_{T_0} . Many methods could be used to model the conditional moments of returns. We employ a simple filtering technique that emphasizes parsimony and imposes minimal assumptions about the data generating process. In particular, we specify exponentially-weighted rolling estimators of the form

$$\hat{\mu}_{T_0}(\phi) = \left(\sum_{t=1}^{T_0} \phi^{T_0-t} \right)^{-1} \sum_{t=1}^{T_0} \phi^{T_0-t} r_t \quad (11)$$

$$\hat{\Omega}_{T_0}(\varphi) = \left(\sum_{t=1}^{T_0} \varphi^{T_0-t} \right)^{-1} \sum_{t=1}^{T_0} \varphi^{T_0-t} r_t r_t' \quad (12)$$

where the smoothing constants ϕ and φ satisfy $0 < \phi, \varphi \leq 1$.

The use of rolling estimators is common in research on mean-variance portfolio selection. A number of studies, for example, construct plug-in estimates of the portfolio weights by using a fixed-width rolling data window to estimate the mean vector and covariance matrix of asset returns.² This approach seeks to balance the benefits of increasing the sample size against the costs of including more distant observations that are less likely to reflect current market conditions. Although the use of a fixed-width window has some intuitive appeal, it is typically less efficient than methods that exploit the full historical sample of asset returns. The literature suggests that exponentially-weighted rolling estimators are preferred from an efficiency perspective (see, e.g., Foster and Nelson, 1996).

Developing a robust method for choosing the smoothing constants is central to our investigation. One possibility is to use a model-fitting approach. For instance, we might choose the

¹ This motivates Paye (2010) to propose a methodology in which multiple plug-in estimates of the weights are combined to obtain the final weights. He shows how to estimate the combination that minimizes the investor’s expected loss, but finds that it is more robust to use a simple average of the different plug-in estimates because this entails less estimation risk overall.

² Some recent examples include DeMiguel et al. (2009), Paye (2010), Tu and Zhou (2011), and Kirby and Ostdiek (2012).

smoothing constants by minimizing sample criteria of the form

$$\hat{G}(\phi) = \text{tr} \left\{ \frac{1}{T_0 - 1} \sum_{t=1}^{T_0-1} (r_{t+1} - \hat{\mu}_t(\phi))(r_{t+1} - \hat{\mu}_t(\phi))' \right\}, \quad (13)$$

$$\hat{H}(\varphi) = \text{tr} \left\{ \frac{1}{T_0 - 1} \sum_{t=1}^{T_0-1} (r_{t+1}r_{t+1}' - \hat{\Omega}_t(\varphi))(r_{t+1}r_{t+1}' - \hat{\Omega}_t(\varphi))' \right\}, \quad (14)$$

where $\text{tr}\{\cdot\}$ denotes the trace operator. This would deliver estimates of the values of ϕ and φ that produce the best forecasts of the returns and their squares and cross products under MSE loss. In general, however, we would not expect such an approach to be satisfactory, because the values that minimize $\hat{G}(\phi)$ and $\hat{H}(\varphi)$ are unlikely to coincide with the values that optimize the performance of the portfolio. The same is true for any approach that relies on an econometrically-motivated loss function.³

We can see this more clearly by considering the implications of specifying risk preferences of the form shown in equation (7). Under a model-fitting approach, we can express the vector of plug-in weights for period T_0 as

$$\hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi}) = \frac{\hat{\Omega}_{T_0}^{-1}(\hat{\varphi})\iota}{\iota'\hat{\Omega}_{T_0}^{-1}(\hat{\varphi})\iota} + \frac{1}{\psi} \left(\hat{\Omega}_{T_0}^{-1}(\hat{\varphi}) - \frac{\hat{\Omega}_{T_0}^{-1}(\hat{\varphi})\iota\iota'\hat{\Omega}_{T_0}^{-1}(\hat{\varphi})}{\iota'\hat{\Omega}_{T_0}^{-1}(\hat{\varphi})\iota} \right) \hat{\mu}_{T_0}(\hat{\phi}), \quad (15)$$

where $\hat{\phi} = g(r_1, r_2, \dots, r_{T_0})$ and $\hat{\varphi} = h(r_1, r_2, \dots, r_{T_0})$ for some model-specific functions $g(\cdot)$ and $h(\cdot)$. The optimal vector of weights, on the other hand, can be expressed as

$$w_{T_0}(\psi) = \frac{1}{\psi} \Omega_{T_0}^{-1}(\mu_{T_0} - \kappa_{T_0}(\psi)\iota), \quad (16)$$

where $\kappa_{T_0}(\psi) = (\psi - \iota'\Omega_{T_0}^{-1}\mu_{T_0})/\iota'\Omega_{T_0}^{-1}\iota$.

If we substitute equation (15) into the utility function in equation (7) and take conditional expectations, we obtain

$$\mathbb{E}_{T_0}[U(\hat{w}_{T_0}(\cdot))] = \kappa_{T_0}(\psi) + \psi \hat{w}_{T_0}(\psi)' \Omega_{T_0} \hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi}) - \frac{\psi}{2} \hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi})' \Omega_{T_0} \hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi}). \quad (17)$$

In comparison, substituting equation (16) into the utility function in equation (7) and taking conditional expectations yields

$$\mathbb{E}_{T_0}[U(w_{T_0}(\cdot))] = \kappa_{T_0}(\psi) + \frac{\psi}{2} w_{T_0}(\psi)' \Omega_{T_0} w_{T_0}(\psi). \quad (18)$$

Hence, under the utility-based loss function

$$L(w_{T_0}, \hat{w}_{T_0}) = U(w_{T_0}(\cdot)) - U(\hat{w}_{T_0}(\cdot)), \quad (19)$$

³ If we consider the case in which $r_t \sim i.i.d. \mathcal{N}(\mu, \Sigma)$, for example, using the maximum likelihood estimators of μ and Σ to construct the plug-in weights does not maximize the expected out-of-sample performance of the portfolio (Kan and Zhou, 2007).

the conditional expected loss is

$$E_{T_0}[L(w_{T_0}, \hat{w}_{T_0})] = \frac{\psi}{2}(w_{T_0}(\psi) - \hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi}))' \Omega_{T_0}(w_{T_0}(\psi) - \hat{w}_{T_0}(\psi, \hat{\phi}, \hat{\varphi})) \quad (20)$$

by using the plug-in weights in place of the optimal weights. The unconditional expected loss follows immediately by the law of iterated expectations.

It is apparent from equation (20) that the magnitude of the expected loss from using the plug-in weights depends on the nature of the functions $g(\cdot)$ and $h(\cdot)$. Under a model-fitting approach these functions are defined implicitly by solving for the values of ϕ and φ that maximize goodness-of-fit with respect to standard statistical criteria. In general, there is no guarantee that the resulting estimators deliver a small expected utility loss. To develop a decision-theoretic approach for choosing the smoothing constants, we treat them as tuning parameters, i.e., as parameters that we can freely change to tailor the performance of the plug-in weights to a particular investment objective.

2.3 Estimating the optimal values of the tuning parameters

Suppose we want to use the Ferson and Siegel (2001) framework to construct a sample UMVE portfolio. Under these circumstances, the problem is to select an active portfolio strategy from within the set of strategies that have weights of the form

$$\hat{w}_{T_0}(\vartheta) = \frac{\hat{\Omega}_{T_0}^{-1}(\varphi)\iota}{\iota'\hat{\Omega}_{T_0}^{-1}(\varphi)\iota} + \frac{1}{\psi} \left(\hat{\Omega}_{T_0}^{-1}(\varphi) - \frac{\hat{\Omega}_{T_0}^{-1}(\varphi)\iota\iota'\hat{\Omega}_{T_0}^{-1}(\varphi)}{\iota'\hat{\Omega}_{T_0}^{-1}(\varphi)\iota} \right) \hat{\mu}_{T_0}(\phi). \quad (21)$$

where $\vartheta = (\psi, \phi, \varphi)'$ denotes the vector of tuning parameters. To identify the optimal active strategy under unconditional mean-variance risk preferences, we have to find the value of ϑ that maximizes the quadratic objective function

$$Q_p(\vartheta) = E[\hat{w}_{T_0}(\vartheta)'r_{T_0+1}] - \frac{\gamma}{2}\text{Var}(\hat{w}_{T_0}(\vartheta)'r_{T_0+1}), \quad (22)$$

where γ measures relative risk aversion. If we assume that $\hat{\mu}_{T_0}(\phi)$ and $\hat{\Omega}_{T_0}(\varphi)$ are correctly-specified parametric models of the conditional mean vector and conditional second-moment matrix, then $Q_p(\vartheta)$ is maximized by setting $\psi = \gamma(1 - \mu_{p_1} + \mu_{p_0})/(1 + \gamma\mu_{p_0})$, and choosing ϕ and φ such that $\hat{\mu}_{T_0}(\phi) = \mu_{T_0}$ and $\hat{\Omega}_{T_0}(\varphi) = \Omega_{T_0}$. This yields the same vector of weights as maximizing $Q_p(w_t)$ in equation (4) for $t = T_0$.

The value of ϑ that maximizes $Q_p(\vartheta)$ is unknown in practice. However, we can use historical returns to construct a sample version of the objective function and estimate this value in a straightforward fashion. To implement our estimation methodology, we split the initial hold-out sample into an “initialization window” that contains the first $K_0 \geq N$ observations and an “estimation window” that contains the remaining $T_0 - K_0$ observations.⁴ The returns in the initialization window are used to initialize the rolling estimators of the conditional mean

⁴ The restriction $K_0 \geq N$ is imposed to ensure that the estimate of Ω_t is invertible for all $t \geq K_0$, i.e., for all dates contained in the estimation window.

vector and conditional second-moment matrix, and the returns in the estimation window are used to construct the sample objective function and estimate the optimal value of ϑ .

The proposed estimator is obtained by applying the weights $\{\hat{w}_t(\vartheta)\}_{t=K_0}^{T_0-1}$ to the returns in the estimation window. Note that $\hat{w}_t(\vartheta)$ depends only on the returns observed in periods 1 through t . It follows, therefore, that applying $\hat{w}_t(\vartheta)$ to r_{t+1} delivers an out-of-sample portfolio return for period $t + 1$. For any choice of ϑ , the sample mean and sample variance of the out-of-sample portfolio returns for the estimation window are given by

$$\hat{\mu}_p(\vartheta) = \frac{1}{T_0 - K_0} \sum_{t=K_0}^{T_0-1} \hat{w}_t(\vartheta)' r_{t+1}, \quad (23)$$

$$\hat{\sigma}_p^2(\vartheta) = \frac{1}{T_0 - K_0} \sum_{t=K_0}^{T_0-1} (\hat{w}_t(\vartheta)' r_{t+1} - \hat{\mu}_p(\vartheta))^2. \quad (24)$$

These sample moments are analogs of the population moments that appear on the right side of equation (22). Under suitable regularity conditions, therefore, the estimate of ϑ obtained by maximizing the sample objective function

$$\hat{Q}_p(\vartheta) = \hat{\mu}_p(\vartheta) - \frac{\gamma}{2} \hat{\sigma}_p^2(\vartheta) \quad (25)$$

converges as $T_0 - K_0 \rightarrow \infty$ to the value of this vector that maximizes $Q_p(\vartheta)$.⁵

In general, we expect the sample objective function to be constructed using misspecified econometric models. For example, the volatility modeling literature suggests that even for small values of N we need a heavily-parameterized model to fully capture the dynamics of Ω_t . Such a model may not be practical in portfolio-choice applications. If we instead use a more parsimonious specification, such as the rolling estimator considered here, the plug-in weights will differ from the true weights for all possible values of the tuning parameters. In this case, maximizing the sample objective function yields estimates of the values of tuning parameters that deliver a portfolio that is UMVE with respect to the choice set established by using the misspecified models to construct the plug-in weights. This yields the most efficient portfolio possible given the choice of modeling techniques.

We also expect the estimates of ϕ and φ obtained by maximizing $\hat{Q}_p(\vartheta)$ to serve the underlying investment objectives better than either *ad hoc* choices of the parameter values or the estimates obtained from a model-fitting approach. To understand why, note that our approach maximizes the average realized utility generated by the portfolio as opposed to a statistical goodness-of-fit criterion. Because utility-based objective functions generally translate into asymmetric loss functions, overestimating an expected return or variance will typically produce a different loss than underestimating this quantity by the same amount.

⁵ Note that global identification is not a concern in this setting, because our objective is limited to constructing an estimate of ϑ that converges to the value that maximizes $Q_p(\vartheta)$ as $T_0 - K_0 \rightarrow \infty$. There is no need to assume that the maximum of $Q_p(\vartheta)$ corresponds to a unique parameter configuration. For a detailed discussion of identification in the context of parameter estimation using economic loss functions, see Skouras (2007).

As a consequence, the tuning-parameter optimization favors estimates of ϕ and φ that are biased in an appropriate direction.⁶

By estimating ϕ and φ jointly with ψ , we allow the optimizer to fully evaluate all of the tradeoffs involved in choosing the tuning parameter values. For instance, reducing the value of ϕ might produce more accurate estimates of conditional expected returns, but it might also increase the time-series variation in the plug-in weights. In isolation this could be counterproductive. However, increasing the value of ψ might compensate for the increased variation in the plug-in weights and ultimately produce a higher value of the sample objective function. Assessing the potential for such tradeoffs is at the core of our strategy for identifying tuning parameter values that optimize the out-of-sample performance of the portfolio.⁷

Brandt et al. (2009) use a related estimation strategy to implement a parametric portfolio rule in large-scale applications. Under their approach, each asset weight is restricted to be linear in a set of asset-specific variables, such as market capitalization, book-to-market value, and lagged returns. This linear-weight-function approach is a middle ground between a fully-specified model of optimal portfolio choice and pure technical trading rules. Although it is an approximation, it has the advantage of drastically reducing computational demands when N is large by eliminating the need to estimate the optimal weights using the functional form implied by theory. Because the values of coefficients in the linear weight functions are unknown, Brandt et al. (2009) estimate them from the data. In particular, they find the coefficient values that maximize average utility over their sample period under a specified utility of wealth function. The proposed approach for optimizing the out-of-sample performance of the plug-in weights is a theory-based alternative to their methodology.

2.4 *Portfolio turnover and rebalancing costs*

Portfolio turnover is always a concern if transaction costs are greater than zero. In this situation, anything that increases turnover can decrease performance after accounting for rebalancing costs. Turnover is usually defined as the fraction of invested wealth traded in a given period to rebalance the portfolio. To see how to compute this measure, note that if one dollar is invested in the portfolio in period $t - 1$, there are $\hat{w}_{i,t-1}(\vartheta)(1 + r_{i,t})$ dollars invested in the i th asset of the portfolio in period t . Hence, the weight in asset i before the portfolio

⁶ Using asymmetric loss functions to evaluate forecast quality is an area of ongoing research. Under asymmetric loss, many of the properties traditionally associated with forecast optimality need not hold. Optimal forecasts can be biased, the forecast errors can display arbitrarily high orders of serial correlation, and the variance of the forecast errors can decline as the forecast horizon increases (Patton and Timmermann, 2007).

⁷ In our experience, the sample objective function in equation (25) tends to have a number of local optima, so some care is needed to avoid termination of the optimization algorithm at these points. We guard against this possibility by conducting multiple optimizations using a range of starting values. Despite this precaution, there may be some cases in which we fail to find the global optimum. Our analysis suggests, however, that the remaining improvement in the value of the objective function that could be achieved in such cases is very small.

is rebalanced is

$$\tilde{w}_{i,t}(\vartheta) = \frac{\hat{w}_{i,t-1}(\vartheta)(1 + r_{i,t})}{1 + \sum_{i=1}^N \hat{w}_{i,t-1}(\vartheta)r_{i,t}}, \quad (26)$$

and the turnover at time t is given by

$$\tau_{p,t}(\vartheta) = \frac{1}{2} \sum_{i=1}^N |\hat{w}_{i,t}(\vartheta) - \tilde{w}_{i,t}(\vartheta)|, \quad (27)$$

where $\hat{w}_{i,t}(\vartheta)$ is the desired weight in asset i at time t .⁸

One advantage of the proposed methodology is that we can take turnover and rebalancing costs directly into account. To illustrate, let $\tilde{r}_{p,t}$ denote the portfolio return net of rebalancing costs for period t . Now suppose that the cost of rebalancing the portfolio to the desired period t weights is subtracted from the return for period t , and that the level of transaction costs is constant both across assets and over time. Under these circumstances,

$$\tilde{r}_{p,t}(\vartheta) = (1 + \hat{w}'_{t-1}(\vartheta)r_t)(1 - 2\tau_{p,t}(\vartheta)c) - 1, \quad (28)$$

where c is the assumed level of proportional costs per transaction.⁹ We can therefore estimate the optimal values of the tuning parameters for a given c by using $\{\tilde{r}_{p,t}(\vartheta)\}_{t=1}^{K_0}$ to initialize the rolling estimators and $\{\tilde{r}_{p,t}(\vartheta)\}_{t=K_0+1}^{T_0}$ to construct the sample objective function. The assumption that c is constant could easily be relaxed. For instance, evidence suggests that the cost of trading U.S. equities has declined over time (Domowitz et al., 2001; Hasbrouck, 2009). This decline can be captured by specifying a linear time trend for trading costs of the form $c_t = c_0 + c_1t$ with appropriate values of c_0 and c_1 .

2.5 Shrinkage and partial-adjustment techniques

Shrinkage methods are a popular technique for improving the performance of the plug-in approach to constructing portfolio weights. The basic idea of shrinkage estimators, as first described by James and Stein (1961), is to reduce the extreme estimation errors that may occur when estimating the cross-section of means, variances, and covariances of asset returns. For example, we might shrink the sample mean for each asset towards the grand sample mean for all the assets. This mitigates the largest estimation errors and may reduce the variance of the estimators by enough to outweigh the biases introduced by the technique.

It is straightforward to apply shrinkage methods in the proposed framework. Consider esti-

⁸ Equation (27) is consistent with the measure of turnover used in the mutual fund industry, i.e., the lesser of the value of purchases and sales in the period divided by net asset value. Here the value of purchases equals the value of sales because there are no fund inflows or outflows.

⁹ Note that $\tau_{p,t}(\vartheta)$ is multiplied by 2 in equation (28) because turnover is the value of assets purchased or, equivalently in our framework, the value of assets sold as a fraction of total wealth. Both purchases and sales incur transaction costs, so rebalancing costs are given by $2\tau_{p,t}(\vartheta)c$.

mators for μ_t and Ω_t of the form

$$\hat{\mu}_t^*(\phi, \rho) = \rho \bar{\mu}_t + (1 - \rho) \hat{\mu}_t(\phi), \quad (29)$$

$$\hat{\Omega}_t^*(\varphi, \rho) = \rho \bar{\Omega}_t + (1 - \rho) \hat{\Omega}_t(\varphi), \quad (30)$$

where $\bar{\mu}_t$ and $\bar{\Omega}_t$ are the shrinkage targets and the shrinkage factor ρ satisfies $0 < \rho \leq 1$.¹⁰ Consistent with the general approach, we treat ρ as a tuning parameter. Thus the vector of plug-in weights for period t becomes

$$\hat{w}_t^*(\vartheta^*) = \frac{\hat{\Omega}_t^{*-1}(\varphi, \rho)\iota}{\iota' \hat{\Omega}_t^{*-1}(\varphi, \rho)\iota} + \frac{1}{\psi} \left(\hat{\Omega}_t^{*-1}(\varphi, \rho) - \frac{\hat{\Omega}_t^{*-1}(\varphi, \rho)\iota \iota' \hat{\Omega}_t^{*-1}(\varphi, \rho)}{\iota' \hat{\Omega}_t^{*-1}(\varphi, \rho)\iota} \right) \hat{\mu}_t^*(\phi, \rho), \quad (31)$$

where $\vartheta^* = (\vartheta', \rho)'$ contains the original tuning parameters plus the shrinkage factor. The shrinkage targets are obtained by averaging the sample means, sample second moments, and sample second moments of the returns that are in the investor's information set when the weights are selected. Specifically, we set $\bar{\mu}_{it} = \iota' \hat{\mu}_t(1)/N$ for all i , $\bar{\Omega}_{ii,t} = \text{tr}\{\hat{\Omega}_t(1)\}/N$ for all i , and $\bar{\Omega}_{ij,t} = (\iota' \hat{\Omega}_t(1)\iota - \text{tr}\{\hat{\Omega}_t(1)\})/(N^2 - N)$ for all $i \neq j$.

The empirical evidence suggests that shrinkage methods reduce the adverse impact of estimation risk, but they may not be the most effective way to address the issue of rebalancing costs. For this reason, we also consider partial-adjustment strategies. These strategies recognize that costly trading can make it inefficient to fully adjust to the estimated optimal position each period because there is an inherent tradeoff between the benefits of incorporating information about changes in the investment opportunity set and the attendant rebalancing costs. The idea behind partial-adjustment strategies is to strike an appropriate balance between these costs and benefits.

Brandt et al. (2009) propose one such strategy. They use a function of the form

$$d_t(\delta) = \frac{1}{N} \sum_{i=1}^N (\hat{w}_{i,t} - \tilde{w}_{i,t}^*(\delta))^2 \quad (32)$$

to measure the distance between the desired weights, \hat{w}_t , and weights before any rebalancing occurs, $\tilde{w}_t^*(\delta)$, and specify that no adjustment of the weights takes place if $d_t(\delta) \leq \delta$. Thus there is a no-trade region — a hypersphere of radius $\sqrt{\delta}$ — around \hat{w}_t . For cases in which $d_t(\delta) > \delta$, the weights are adjusted to the boundary of the no-trade region by setting

$$\hat{w}_t^*(\delta) = \varrho_t(\delta) \tilde{w}_t^*(\delta) + (1 - \varrho_t(\delta)) \hat{w}_t, \quad (33)$$

where $\varrho_t(\delta) = (\delta/d_t(\delta))^{1/2}$.

Although partial-adjustment strategies can be motivated by the theory of portfolio optimization in the presence of transaction costs (see, e.g., Leland, 1999), there is no claim that the optimal shape of the no-trade region is a hypersphere. Indeed, equation (20) suggests using a different shape for an investor with quadratic risk preferences. It shows that the conditional

¹⁰ We investigated using different shrinkage factors for the first and second moments, but found that this had little impact on our results.

expected loss generated by errors in estimating the optimal weights is a quadratic form in the conditional second moment matrix of returns. Accordingly, we propose a partial-adjustment strategy based on the distance measure

$$d_t(\vartheta^*) = (\hat{w}_t(\vartheta) - \tilde{w}_t^*(\vartheta^*))' \hat{\Omega}_t(\varphi) (\hat{w}_t(\vartheta) - \tilde{w}_t^*(\vartheta^*)), \quad (34)$$

where $\vartheta^* = (\vartheta', \delta)'$ contains the original tuning parameters plus the no-trade distance. The value of $d_t(\vartheta^*)$ approximates the conditional expected loss in utility from leaving the weights unchanged. If the anticipated loss is less than δ , no adjustment takes place. Otherwise the weights are adjusted to the no-trade boundary by setting

$$\hat{w}_t^*(\vartheta^*) = \varrho_t(\vartheta^*) \tilde{w}_t^*(\vartheta^*) + (1 - \varrho_t(\vartheta^*)) \hat{w}_t(\vartheta), \quad (35)$$

where $\varrho_t(\vartheta^*) = (\delta/d_t(\vartheta^*))^{1/2}$.

3 Empirical Application

To investigate the effectiveness of the proposed methodology, we consider an empirical application in which the goal is to create an optimal “fund-of-funds” strategy by investing in a defined set of characteristic-based portfolios that contain NYSE, AMEX, and NASDAQ firms. Using portfolios rather than individual stocks as assets has two advantages. First, it allows us to assess the extent to which the cross-sectional and time-series variation in the plug-in weights is related to well-known empirical regularities, such as value, growth, and momentum effects. This provides insights on the features of the research design that influence the performance of sample UMVE portfolios. Second, it allows us to directly relate our findings to the literature, because most of the empirical research on mean-variance optimization uses portfolios rather than individual stocks.

We use a number of equity benchmarks, such as the S&P 500 index, to draw inferences about the observed performance of the sample UMVE portfolios. Treasury bills are excluded from the set of assets to ensure that the observed differences in performance between the sample UMVE portfolios and the benchmarks are not driven by allocations to the conditionally risk-free security. As Brandt et al. (2009) point out, the first-order effect of including a risk-free security in the portfolio is simply to change the leverage, not the relative weightings of the risky assets. Thus, little is lost by excluding Treasury bills from consideration.

3.1 Datasets

The empirical investigation is conducted using monthly returns on three sets of equally-weighted stock portfolios for the period from January 1946 to December 2009 (768 observations).¹¹ We employ equally-weighted portfolios for the analysis so that naïve diversification, one of our benchmark strategies, corresponds to holding an equally-weighted portfolio of individual stocks. This allows us to attribute the observed differences in performance between

¹¹ We exclude data prior to 1945 from the analysis because of the atypical conditions that prevailed in U.S. equity markets during the Great Depression and World War II.

the $1/N$ portfolio and the sample UMVE portfolios to the impact of grouping individual firms on the basis of observed characteristics and using mean-variance optimization to take advantage of the resulting cross-sectional and time-series variation in the conditional moments of returns. In general, we expect the sorting rule to play an important role in the analysis because it affects key characteristics of the investment opportunity set.

Two of the three datasets are from a data library maintained by Ken French.¹² The first dataset is constructed by sorting individual firms into 10 Industry portfolios using standard industrial classification (SIC) codes. By using industry portfolios as assets, we potentially encompass the type of sector rotation strategies popular among professional money managers. The second dataset is constructed by sorting individual firms into 25 Size/Book-to-Market portfolios using market capitalization and book-to-market values. Using these portfolios as assets allows us to examine the interplay between value and growth effects. Both datasets are representative of those used in prior research (see, e.g., DeMiguel et al., 2009).

The third dataset is constructed using a sorting scheme that is motivated by the results of Kirby and Ostdiek (2012). They find that the cross-sectional dispersion in the sample means and sample variances of the asset returns influences the performance of mean-variance methods of portfolio selection. Although this is not surprising, it suggests that a more comprehensive approach to the mean-variance optimization problem might be beneficial. In particular, grouping firms in a manner specifically designed to create a large dispersion in both means and variances might lead to improved performance of the out-of-sample portfolio. To investigate this possibility, we employ a dataset that is constructed using past returns and average absolute returns to sort individual firms into 30 Momentum/Volatility portfolios.¹³

Figure 1 plots the annualized values of the sample mean returns and sample return volatilities for the three datasets. The patterns observed for the 10 Industry (panel A) and 25 Size/Book-to-Market (panel B) portfolios are familiar from other studies. First, sorting firms on SIC codes produces less dispersion in average returns than sorting firms on size and book-to-market values. The sample means range from 12.8% to 20.2% for the industry portfolios and from 10.6% to 23.4% for the size/book-to-market portfolios. Second, the choice of sorting scheme has less effect on the dispersion in sample volatilities. The range is 11.8% to 30.3% for the industries and 16.2% to 29.2% for size/book-to-market. In comparison, there is much more dispersion in sample volatilities for the 30 Momentum/Volatility portfolios (panel C), with a range of 9.5% to 47.4%. Moreover, the dispersion in sample means for these portfolios is comparable to that for the 25 Size/Book-to-Market portfolios: 11.1% to 25.6%. Thus the preliminary evidence suggests that including the choice of sorting scheme in the scope of the

¹² See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

¹³ The returns are drawn from the Center for Research in Security Prices monthly stock file. We form the portfolios for each month t as follows. First, we use the average absolute monthly return for months $t - 12$ to $t - 2$ to sort firms into volatility deciles. Next, we use the holding-period return for the interval $t - 12$ to $t - 2$ to sort the firms within each volatility decile into three momentum portfolios. Firms included in a portfolio for month t have a non-missing price for month $t - 13$, a non-missing return for month $t - 2$, a non-missing price and non-missing shares outstanding for month $t - 1$, and code -99.0 for any missing returns for months $t - 12$ to $t - 3$. These are the same filters used to construct the momentum portfolio dataset in the Ken French data library.

optimization problem is a promising strategy.

3.2 *Rolling-sample strategy for computing the plug-in weights*

We construct the sample UMVE portfolios using a rolling-sample approach in which the optimal values of the tuning parameters are reestimated with each new data point that becomes available. First, we split the dataset into the initial holdout sample, which contains the initial T_0 observations, and a performance evaluation sample, which contains the remaining T observations. The initial holdout sample is used to construct the initial estimates of the optimal tuning parameter values, which are then used to compute the plug-in weights for the interval T_0 to $T_0 + 1$. This delivers the portfolio return for period $T_0 + 1$. Next, we form an updated holdout sample that consists of the initial T_0 observations plus the observation for period $T_0 + 1$. The updated holdout sample is used to construct updated estimates of the optimal values of the tuning parameters, which are then used to compute the plug-in weights for the interval $T_0 + 1$ to $T_0 + 2$. This delivers the portfolio return for the period $T_0 + 2$. We continue updating the parameter estimates and computing the plug-in weights in this manner through the end of the performance evaluation sample.

To implement the rolling-sample approach, we have to choose a value for T_0 , the length of the initial holdout sample, and a value for K_0 , the length of the initialization window for the exponentially-weighted rolling estimators of the conditional moments of returns. These choices entail tradeoffs between opposing considerations. Increasing T_0 yields more precise estimates of the optimal values of the tuning parameters, but it also shortens the performance evaluation sample, making it more difficult to detect differences in performance across portfolios. Similarly, increasing K_0 makes the exponentially-weighted rolling estimators less noisy, especially in the early part of the holdout sample, but it also reduces the number of returns available to estimate the tuning parameters.

We look to prior research to guide the choice of K_0 , and choose T_0 based on our assessment of the amount of data needed for the initial tuning-parameter optimization. A number of studies, such as Chan et al. (1999) and DeMiguel et al. (2009), use rolling estimators with 5- to 10-year windows to construct sample mean-variance efficient portfolios. This suggests setting K_0 in the 60 to 120 range. We opt for a 10-year initialization window ($K_0 = 120$) rather than a shorter window length because of the difficulties in obtaining accurate estimates of expected returns (see Merton, 1980). To ensure that the initial estimates of the optimal tuning parameter values display reasonable precision we use a 20-year sample of monthly portfolio returns for the optimization, resulting in a 30-year initial holdout sample ($T_0 = 360$).

To construct the sample objective function for the optimization, we have to specify values for γ and c . Our choice of γ could have a significant impact on the findings, particularly if transaction costs are high. In effect, we are choosing the aggressiveness of the strategy. A low value of γ translates into an aggressive strategy, while a high value implies a more conservative investment style. We use $\gamma = 15$, which imposes a large risk penalty, to generate our results. This is consistent with an emphasis on low-turnover strategies. Plausible choices for c could range from as low as 5 bp for large institutional investors to as high as 50 bp for individual investors. To facilitate inference on the relationship between portfolio turnover

and the level of assumed transaction costs, we consider both of these values.

3.3 Performance benchmarks and statistical inference

We use three equity benchmarks to evaluate the performance of the sample UMVE portfolios. The first is the S&P 500 index. Monthly returns for this portfolio are drawn from the Center for Research in Security Prices monthly stock file. The second is the 1/ N portfolio, i.e., a portfolio with equal weight in each of the N assets under consideration. This is the benchmark advocated by DeMiguel et al. (2009). The third is a simple plug-in version of the global minimum variance (GMV) portfolio. This is a useful benchmark because researchers often report that it performs well in comparison to plug-in versions of other mean-variance efficient portfolios.¹⁴ Indeed, the GMV portfolio is the only portfolio that frequently outperforms naïve diversification in the DeMiguel et al. (2009) study.

We obtain the plug-in version of the GMV portfolio by mimicking the procedure used by DeMiguel et al. (2009). To illustrate, let $\Sigma_t = \Omega_t - \mu_t \mu_t'$ denote the conditional covariance matrix of returns. In a setting with no riskless asset, the GMV portfolio for period $t + 1$ is obtained by finding the vector of weights that minimizes $w_t' \Sigma_t w_t$ subject to $w_t' \iota = 1$. The solution is $w_t = \Sigma_t^{-1} \iota / \iota' \Sigma_t^{-1} \iota$.¹⁵ We estimate these weights for each month t in the performance evaluation window by replacing μ_t and Σ_t with rolling estimates of the form $\hat{\mu}_t = (1/L) \sum_{i=0}^{L-1} r_{t-i}$ and $\hat{\Sigma}_t = (1/L) \sum_{i=0}^{L-1} (r_{t-i} - \hat{\mu}_t)(r_{t-i} - \hat{\mu}_t)'$. Following DeMiguel et al. (2009), we set $L = 120$ for the empirical analysis.

The statistical significance of the observed differences in performance across portfolios is assessed using large-sample t -statistics. Let $\hat{\lambda}_{p_i}$ and $\hat{\lambda}_{p_j}$ denote the estimated Sharpe ratios of portfolios i and j for the evaluation period. If the two portfolios have the same population Sharpe ratio, then we have the large-sample approximation

$$\sqrt{T} \left(\frac{\hat{\lambda}_{p_i} - \hat{\lambda}_{p_j}}{\hat{V}_\lambda^{1/2}} \right) \overset{a}{\sim} \mathcal{N}(0, 1), \quad (36)$$

where \hat{V}_λ denotes a consistent estimator of the asymptotic variance of $\sqrt{T}(\hat{\lambda}_{p_i} - \hat{\lambda}_{p_j})$.¹⁶ To

¹⁴ This is often attributed to the fact that the weights of the GMV portfolio do not depend on expected returns, which reduces estimation risk (see, e.g., Jagannathan and Ma, 2003).

¹⁵ We call this the GMV portfolio to be consistent with the terminology used in previous studies. From an unconditional perspective, it is actually the global minimum second-moment portfolio. To see this, use equation (20) of Ferson and Siegel (2001) to establish that $w_t = \Omega_t^{-1} \iota / \iota' \Omega_t^{-1} \iota$ minimizes the unconditional second moment of the portfolio, and use the Sherman-Morrison matrix inversion formula to invert $\Omega_t = \Sigma_t + \mu_t \mu_t'$ and establish that $\Omega_t^{-1} \iota / \iota' \Omega_t^{-1} \iota = \Sigma_t^{-1} \iota / \iota' \Sigma_t^{-1} \iota$.

¹⁶ We use the generalized method of moments to construct this estimator. Let

$$e_t(\hat{\theta}) = \begin{pmatrix} \tilde{r}_{p_{it}} - r_{ft} - \hat{\sigma}_{p_i} \hat{\lambda}_{p_i} \\ \tilde{r}_{p_{jt}} - r_{ft} - \hat{\sigma}_{p_j} \hat{\lambda}_{p_j} \\ (\tilde{r}_{p_{it}} - r_{ft} - \hat{\sigma}_{p_i} \hat{\lambda}_{p_i})^2 - \hat{\sigma}_{p_i}^2 \\ (\tilde{r}_{p_{jt}} - r_{ft} - \hat{\sigma}_{p_j} \hat{\lambda}_{p_j})^2 - \hat{\sigma}_{p_j}^2 \end{pmatrix},$$

assess whether a given UMVE portfolio i outperforms a particular benchmark j , we find the p -value for the null hypothesis $\lambda_{p_i} - \lambda_{p_j} \leq 0$. If we reject the null, we conclude that the sample UMVE portfolio displays superior performance.

3.4 Results for the 10 Industry dataset

Table 1 documents the performance of the sample UMVE portfolios for the 10 Industry dataset. Panel A presents results for the baseline case and panel B for the case with shrinkage estimators. Each panel contains four sets of results. The initial two rows are for $c = 5$ bp with $\delta = 0$ (full adjustment to the estimated optimal weights) and $\delta \geq 0$ (partial adjustment to the no-trade boundary). The final two rows are for $c = 50$ bp with $\delta = 0$ and $\delta \geq 0$.

For each c and δ combination, we report the estimated expected portfolio return, the estimated portfolio standard deviation, the estimated portfolio Sharpe ratio, the estimated certainty-equivalent (CE) portfolio return, and the estimated expected portfolio turnover. The first four statistics are computed using transaction costs of both 5 and 50 basis points to assess the impact of using a value of c in the tuning-parameter optimization that differs from the level of transaction costs assumed for the performance evaluation. We also report p -values for three one-sided hypothesis tests: (i) the S&P 500 index performs at least as well as the UMVE portfolio (ii) the $1/N$ portfolio performs at least as well as the UMVE portfolio, and (iii) the plug-in GMV portfolio performs at least as well as the UMVE portfolio.

First consider the no-shrinkage case with $c = 5$ bp. If we set $\delta = 0$ and assume transaction costs of 5 bp for performance evaluation, the sample UMVE portfolio has an estimated expected return of 33.3%, an estimated standard deviation of 16.8%, an estimated Sharpe ratio of 1.65, and an estimated CE return of 12.1%. Setting $\delta \geq 0$ yields identical results because the optimizer chooses $\hat{\delta} = 0$ even when partial adjustment is not ruled out *a priori*. In comparison, the best-performing benchmark — the $1/N$ portfolio — has an estimated Sharpe ratio of 0.55. Its estimated expected return is less than half that of the sample UMVE portfolio, but its estimated standard deviation is almost three percentage points higher. The p -values indicate that the sample UMVE portfolio outperforms all the benchmarks at the 1% significance level. These findings highlight the promise of the methodology.

Note, however, that the sample UMVE portfolio achieves these results at the cost of very high portfolio turnover, averaging over 1400% per year. As a consequence, its performance deteriorates markedly if we impose transaction costs of 50 bp for performance evaluation. The estimated expected return falls to 20.2%, the estimated standard deviation increases to 17.3%, the estimated Sharpe ratio falls to 0.85, and the estimated CE return falls to -2.3% . Nonetheless, it still outperforms both the S&P 500 index and the plug-in GMV portfolio at the 5% significance level, and the $1/N$ portfolio at the 10% significance level.

where r_{ft} denotes the one-month Treasury bill rate and $\hat{\theta} = (\hat{\lambda}_{p_i}, \hat{\lambda}_{p_j}, \hat{\sigma}_{p_i}^2, \hat{\sigma}_{p_j}^2)'$ contains the sample Sharpe ratios and excess return variances for portfolios i and j . Under regularity conditions (see Hansen, 1982), $\sqrt{T}(\hat{\theta} - \theta) \overset{a}{\sim} \mathcal{N}(0, (\hat{D}'\hat{S}^{-1}\hat{D})^{-1})$ where $\hat{D} = (1/T) \sum_{t=T_0+1}^{T_0+T} \partial e_t(\hat{\theta})/\partial \hat{\theta}'$ and $\hat{S} = \hat{\Gamma}_0 + \sum_{l=1}^m (1 - l/(m+1))(\hat{\Gamma}_l + \hat{\Gamma}_l')$ with $\hat{\Gamma}_l = (1/T) \sum_{t=T_0+l+1}^{T_0+T} e_t(\hat{\theta})e_{t-l}(\hat{\theta})'$. For the empirical analysis, we set $m = 5$ and $\hat{V}_\lambda = \hat{V}_{11} - 2\hat{V}_{12} + \hat{V}_{22}$ where $\hat{V} \equiv (\hat{D}'\hat{S}^{-1}\hat{D})^{-1}$.

Figure 1 suggests a possible explanation for the high turnover for the 10 Industry dataset. Because the cross-sectional dispersion in sample means for the industry portfolios is relatively low, the performance advantage of the sample UMVE portfolio may largely reflect the success of the optimizer in exploiting time-series variation in the estimates of conditional expected returns. If this is the case, the high turnover is not surprising. Imposing a larger rebalancing penalty should reduce the aggressiveness of the optimizer and, consequently, improve performance under the 50 bp transaction costs assumption for performance evaluation.

Increasing c to 50 bp does produce a large decrease in the estimated expected turnover. It falls by over 1000 percentage points to about 380% per year. If we impose transaction costs of 5 bp for performance evaluation, the sample UMVE portfolio has an estimated expected return of 21.6%, an estimated standard deviation of 13.4%, an estimated Sharpe ratio of 1.21, and an estimated CE return of 8.2%. This represents a substantial reduction in performance, but the portfolio still outperforms all three benchmarks at the 1% significance level.

If we instead impose transaction costs of 50 bp for performance evaluation, the estimated expected return falls to 18.1%, the estimated standard deviation increases to 13.5%, the estimated Sharpe ratio falls to 0.94, and the estimated CE return falls to 4.5%. This is sufficient to outperform all three benchmarks at the 5% significance level. Setting $\delta \geq 0$ leads to a further reduction in turnover, but the performance of the sample UMVE is little changed under either transaction costs assumption. In all cases, the portfolio outperforms the three benchmarks at the 5% significance level.

Note also that there is a distinct pattern to the estimated CE returns in panel A. Specifically, for each level of transaction costs imposed for performance evaluation, using the matching level of c for tuning-parameter optimization leads to higher estimated CE returns than using the alternative level of c . This is as expected, because the chosen tuning parameter values maximize the historical performance of the portfolio under the selected value of c . This comparison confirms that the procedure is working as intended by appropriately adapting to changes in the level of assumed transaction costs.

Now consider the impact of using shrinkage estimators on the performance of the sample UMVE portfolio. In general, the results in panel B look a lot like those in panel A. We see the same patterns, but the performance of the sample UMVE portfolio is slightly improved for every c and δ combination. For example, with $c = 50$ bp and $\delta \geq 0$, the estimated expected portfolio turnover is about 285%. If we impose transaction costs of 50 bp for performance evaluation, the portfolio has an estimated expected return 20.1%, an estimated standard deviation of 14.1%, an estimated Sharpe ratio of 1.04, and an estimated CE return of 5.2%. This is sufficient to outperform all three benchmarks at the 1% significance level.

The small gain from employing shrinkage estimators is an interesting result. Other studies tend to find larger effects (see, e.g., the information ratios reported by Ledoit and Wolf, 2004). Perhaps this is because the benefit of using shrinkage techniques is tied to the tuning parameter values. Conventional approaches to specifying the tuning parameter values might yield values that are far from optimal, thereby leaving substantial room for improvement and making shrinkage estimators more valuable. The evidence thus far indicates that the additional flexibility offered by these estimators has little value within a framework that

includes tuning-parameter optimization.

3.4.1 Properties of the estimated tuning parameters and plug-in weights

To aid in understanding how the results in Table 1 are achieved, we document key properties of the estimated tuning parameters and plug-in weights. In the interest of space, we limit the analysis to one particular case: $c = 50$ bp and $\delta \geq 0$ with the use of shrinkage estimators. This is the most relevant case in terms of practical application. It delivers the lowest estimated turnover of the 12 scenarios considered, yet produces a sample UMVE portfolio that outperforms all of the benchmarks at the 1% significance level.

Figure 2, panel A plots the estimated optimal value of each tuning parameter through time. The left-hand scale is for the estimated smoothing constants, $\hat{\phi}$ and $\hat{\varphi}$; the right-hand scale is for the estimated risk penalty, $\hat{\psi}$, no-trade distance, $\hat{\delta}$, and shrinkage factor, $\hat{\rho}$.¹⁷ The shaded rectangles identify bear market periods (10% or greater market declines). In general, the estimates of the smoothing constants are quite stable. The value of $\hat{\phi}$ ranges from 0.68 to 0.84 with some upward trend through time, indicating that the estimates of conditional expected returns are only moderately persistent. In comparison, the estimates of the conditional second-moments of returns are very persistent: the value of $\hat{\varphi}$ ranges from 0.93 to 0.99. It is likely, therefore, that most of the time-series variation in the plug-in weights for this dataset is driven by variation in $\hat{\mu}_t(\hat{\phi})$.

The value of $\hat{\psi}$ is somewhat more variable through time, ranging from 27 to 57. Even at the low end, the risk-penalty used to construct the estimated weights is substantially higher than the value of γ used to construct the sample objective function. Recall that in Section 2.1 we found that setting $\psi < \gamma$ is necessary to obtain the UMVE portfolio for a given value of γ . However, this was in the absence of both rebalancing costs and estimation risk. The restriction on the value of ψ need not hold more generally, because rebalancing costs reduce the expected portfolio return and estimation risk inflates the variance of the portfolio return. These effects are likely to favor a more conservative investment strategy than that obtained by setting $\psi < \gamma$.

Notice that the value of $\hat{\psi}$ jumps upward about a third of the way through the performance evaluation period. Specifically, it increases from about 42 to 55. This is in response to the 1987 stock market crash. Once the October 1987 data enter the holdout sample, the crash is reflected in returns used to construct the sample objective function. The presence of large negative returns causes the optimizer to select a more conservative value of $\hat{\psi}$. Over time the influence of the crash dissipates as the length of the holdout sample increases, leading to a decline in the value of $\hat{\psi}$. It falls to below 35 before jumping upward again at the end of the performance evaluation period.

The value of $\hat{\delta}$, which defines the size of the no-trade region, ranges from 0.6 to 11.9. If the annualized value of the estimated conditional expected utility loss from leaving weights unchanged exceeds $\hat{\delta}$ bp, we rebalance the portfolio. Even at the high end of this range,

¹⁷ We report $\hat{\delta}$ as an annualized bp measure and $\hat{\rho}$ in percent, i.e., we multiply the estimates by 120,000 and 100, respectively, to obtain the values shown in the figure.

the no-trade region is quite small. This explains why we see only a modest impact from allowing partial adjustment of the estimated weights. The small no-trade region suggests that reducing the time-series variation in the weights entails a heavy performance penalty for this dataset.

The value of $\hat{\rho}$ ranges from about 24% to 42%. This is higher than might be anticipated based on the results in Table 1. Even though the methodology incorporates a substantial degree of shrinkage, the performance results are little changed from the no-shrinkage case. It appears, therefore, that the performance of the sample UMVE portfolio is not very sensitive to the choice of shrinkage factor over a large range of values. This again suggests that the additional flexibility offered by shrinkage estimators has little value when optimizing with respect to the tuning parameters.

In Figure 3, panel A we plot the minimum weight, the maximum weight, and the sum of the negative weights for each month in the performance evaluation period. The minimum weight is typically between -10% and -50% , while the maximum weight is typically between 50% and 100% . The sum of the negative weights, a measure of portfolio leverage, is typically above -100% . Although this is high by mutual fund standards, it is much more reasonable than the levels reported in some previous studies of mean-variance optimization. For example, our findings contrast sharply with those of DeMiguel et al. (2009). Extreme weights are common under their approach to constructing sample mean-variance efficient portfolios. In one case they report estimated weights ranging from $-148,195\%$ to $116,828\%$. Our analysis suggests that extreme weights can be a consequence implementation choices rather than a fundamental shortcoming of mean-variance optimization itself.

3.5 Results for the 25 Size/Book-to-Market dataset

The analysis for the 10 Industry dataset suggests that the proposed methodology for constructing plug-in weights improves the out-of-sample performance of mean-variance optimization by a substantial margin. The one aspect of the results that stands out as a possible concern is turnover, which is considerably higher than that of a typical actively-managed institutional portfolio. Even though the performance of the sample UMVE portfolios still looks quite good after accounting for the impact of plausible transaction costs, an estimated expected turnover of more than 250% per year could raise doubts about the economic relevance of the portfolio choice problem addressed by the research design.¹⁸ Thus, it is useful to understand more about the drivers of turnover in our framework.

To develop additional insights on turnover, we turn to the 25 Size/Book-to-Market dataset. Earlier it was noted that using size and book-to-market characteristics to sort firms into portfolios generates a wider dispersion in the sample means than sorting firms on SIC codes.

¹⁸ Griffin and Xu (2009) provide a useful perspective on this issue. They report that an annualized turnover in the neighborhood of 100% is not uncommon for actively-managed mutual funds and that the turnover for a meaningful portion of the hedge funds examined is between 100% and 200% per year. Moreover, they find that hedge funds often have turnover approaching 200% per year and that the turnover for a small share of the funds studied is over 200% per year.

This may indicate that the firms within each size/book-to-market portfolio are more homogeneous in terms of asset pricing dynamics. To the extent that the sorting scheme generates large cross-sectional differences in conditional expected returns that persist over time, we anticipate that it will generate strong persistence in the plug-in weights. This should translate into lower portfolio turnover, all else being equal.

Table 2 documents the performance of the sample UMVE portfolios for the 25 Size/Book-to-Market dataset. In the no-shrinkage case, setting $c = 5$ bp, $\delta = 0$, and imposing transaction costs of 5 bp for performance evaluation yields an estimated expected return of 21.9%, an estimated standard deviation of 12.6%, an estimated Sharpe ratio of 1.30, and an estimated CE return of 10.0%. This is sufficient to outperform the S&P 500 index and *the*1/ N portfolio at the 1% significance level. The estimated expected turnover — 464% per year — is high, but not nearly as high as the corresponding value in Table 1. If we impose transaction costs of 50 bp for performance evaluation, the estimated expected return falls to 17.7%, the estimated standard deviation increases to 12.8%, the estimated Sharpe ratio falls to 0.95, and the estimated CE return falls to 5.4%. This is sufficient to outperform the S&P 500 index and the 1/ N portfolio at the 5% significance level. Allowing for partial adjustment of the plug-in weights has little impact on portfolio performance.

In addition to the reduction in turnover, the other notable departure from the 10 Industry results is that the sample UMVE portfolio fails to significantly outperform the plug-in GMV portfolio. This is because of the strong performance of the plug-in GMV portfolio for this dataset. If we impose transaction costs of 5 bp for performance evaluation, the plug-in GMV portfolio has an estimated Sharpe ratio of 1.22 and a certainty equivalent return of 8.8%. Thus it outperforms both the S&P 500 index and 1/ N portfolio at the 1% significance level. In addition, its estimated expected turnover of 495% per year is only slightly higher than that of the sample UMVE portfolio. It appears that the *ad hoc* choice of a 120-month window length for the rolling estimators is well-suited to these data.

The disparity in the performance of the plug-in GMV portfolio across the 10 Industry and 25 Size/Book-to-Market datasets is noteworthy. It might be due solely to differences in the shape of the efficient frontier and its location in expected return/standard deviation space. But it could also be a direct reflection of the choice of L . The fundamental problem with using *ad hoc* choices of the tuning parameters is that values that perform well in one setting may perform poorly in another. This is why it is important to use the data to assess the effect of changing the values of the tuning parameters and to focus on economic objectives when developing a methodology for choosing these values. Doing so brings a level of robustness to parameter choices that is not possible otherwise.

The results for the case with $c = 50$ bp reinforce this point. With $\delta = 0$ the estimated expected turnover of the sample UMVE portfolio falls to about 285% per year. Allowing for partial adjustment of the weights reduces this further to about 135% per year. This reduction in turnover gives the sample UMVE portfolio a substantial advantage over the plug-in GMV portfolio in settings with high transaction costs. After imposing transaction costs of 50 bp for performance evaluation, the plug-in GMV portfolio has an estimated Sharpe ratio of 0.86 and a certainty equivalent return of 4.2%. The values for the sample UMVE portfolio are 1.03 and 6.1%. We find, therefore, that the sample UMVE portfolio outperforms the plug-in

GMV portfolio at the 5% significance level.

Using shrinkage estimators has a greater impact for the 25 Size/Book-to-Market dataset than for the 10 Industry dataset, but the effects are still small. If we set $c = 50$ bp, $\delta \geq 0$ and impose transaction costs of 50 bp for performance evaluation, the sample UMVE portfolio has an estimated Sharpe ratio of 1.11 and a certainty equivalent return of 7.4%, a small improvement on the 1.03 and 6.1% values obtained in the no-shrinkage case. In addition, the estimated expected turnover falls from about 135% per year to about 100% per year. This reduction could be meaningful if we strongly favor low-turnover strategies.

3.5.1 Properties of the estimated tuning parameters and estimated weights

Figure 2, panel B plots the estimated optimal value of each tuning parameter through time for the $c = 50$ bp and $\delta \geq 0$ case with the use of shrinkage estimators. There are two notable differences from the analogous graph for the 10 Industry dataset. Although the values of $\hat{\delta}$ and $\hat{\rho}$ are similar to those in panel A, the values of $\hat{\phi}$ and $\hat{\psi}$ are substantially higher. The former ranges from 0.990 to 0.997, while the latter ranges from 0.97 to 0.99. This indicates that the estimates of the conditional first and second moments of returns are very persistent. In addition, the value of $\hat{\psi}$ is lower than in panel A, ranging from about 12 to 26.

The high values of $\hat{\phi}$ and $\hat{\psi}$ undoubtedly contribute to the reduced turnover of the sample UMVE portfolios. In general, highly persistent estimates of the means, variance, and covariances can be expected to translate into highly persistent plug-in weights. The low end of the range of estimated expected turnover values for this dataset — about 100% — would fall in the upper tail of the distribution of annual turnover levels observed for active managers.

The high values of $\hat{\phi}$ and $\hat{\psi}$ can also help explain why the plug-in GMV portfolio obtained by setting $L = 120$ performs well for this dataset. With this choice of window length, the average age of the returns used to estimate the weights of the GMV portfolio is 60 months. To obtain the same average age with exponential smoothing requires a smoothing constant of 0.983. This is close to the estimated optimal values of the smoothing constants. It appears that, for this dataset, estimating the optimal choice of window length from the data would produce something close to our *ad hoc* choice of L . We could not have anticipated this, however, without first assessing the performance of the sample UMVE portfolios using historical data. There is no reliable *a priori* basis for predicting how a given choice of L will perform for a particular combination of dataset and investment objective.

Figure 3, panel B plots the minimum weight, the maximum weight, and the sum of the negative weights for each month in the performance evaluation period. The minimum weight is typically between -30% and -70% , while the maximum weight is typically between 40% and 70% . In these respects, the properties of the estimated weights are similar to those for the 10 Industry dataset. The sum of the negative weights, on the other hand, is typically in the -200% to -300% range, so the sample UMVE portfolio displays considerably higher leverage for this dataset than for the 10 Industry dataset.

The reason for the higher leverage is not immediately clear, but aggregating the estimated weights within in each book-to-market quintile reveals relatively large short positions in both

the lowest and highest book-to-market quintiles. Thus the sample UMVE portfolio shorts stocks at the both ends of the value/growth spectrum to invest in stocks in the middle of the spectrum. The extent to which the performance of the portfolio is driven by high leverage is not yet clear. There is, however, no evidence of the extreme short positions reported elsewhere in the literature on mean-variance portfolio optimization.

3.6 Results for the 30 Momentum/Volatility dataset

The difference in the results for the 10 Industry and 25 Size/Book-to-Market datasets suggests that the sorting rule used to create the dataset is an important consideration. This is not surprising in view of the potential for the sorting rule to affect properties such as the cross-sectional dispersion in conditional means and variances. Intuitively, we would expect to find that increasing the cross-sectional dispersion in the conditional moments enhances the performance of mean-variance optimization. To see if the evidence supports this hypothesis, we turn to the 30 Momentum/Volatility dataset, a dataset constructed using a sorting rule specifically designed to accomplish this objective.

Table 3 documents the performance of the sample UMVE portfolios for the 30 Momentum/Volatility dataset. We find clear support for our hypothesis. In the no-shrinkage case, for example, setting $c = 5$ bp, $\delta = 0$, and imposing transaction costs of 5 bp for performance evaluation yields an estimated expected return of 27.6%, an estimated standard deviation of 12.5%, an estimated Sharpe ratio of 1.76, an estimated CE return of 15.9%, and an estimated expected turnover of around 345% per year. This is the best performance for this combination of settings across the three datasets.

The sample UMVE portfolio also performs well in a relative sense. The plug-in GMV portfolio is again the best-performing benchmark. However, its performance for this dataset is not particularly impressive. It has an estimated Sharpe ratio of 0.68, an estimated certainty-equivalent return of 5.0%, and an estimated expected turnover of about 475% per year. We find, therefore, that the sample UMVE portfolio outperforms the plug-in GMV portfolio and the other two benchmarks at the 1% significance level.

The most compelling results are obtained with $c = 50$ bp and $\delta \geq 0$. First, this combination of settings delivers an estimated expected turnover of only 75% per year. With turnover of this magnitude, there is little difference in the performance of the portfolio across levels of transaction costs. Using transaction costs of 5 bp for performance evaluation, it has an estimated expected return of 23.7%, an estimated standard deviation of 11.2%, an estimated Sharpe ratio of 1.62, and an estimated CE return of 14.2%. The corresponding values with transaction costs of 50 bp are 23.0%, 11.3%, 1.54, and 13.4%. Regardless of the level of transaction costs, the sample UMVE portfolio outperforms all of the benchmarks at the 1% significance level. Employing shrinkage methods produces similar results.

These findings are indicative of the potential for mean-variance optimization to deliver on the promise of improved portfolio performance in out-of-sample applications. They show what can be achieved by employing a comprehensive approach to the portfolio choice problem. First, we lay the groundwork for successful implementation of the optimization methodology by constructing the dataset in a manner designed to increase the cross-sectional dispersion

in conditional means and variances. Second, we make the optimization methodology more robust by using the data to estimate the values of the tuning parameters that deliver the best-performing portfolio with respect to the stated investment objective. The effect of these two factors working in concert is apparent in these results.

3.6.1 *Properties of the estimated tuning parameters and plug-in weights*

In Figure 2, panel C, we plot the estimated optimal value of each tuning parameter through time for the $c = 50$ bp and $\delta \geq 0$ case with the use of shrinkage estimators. The value of $\hat{\phi}$ ranges from about 0.99 to 1, and the value of $\hat{\varphi}$ ranges from about 0.98 to 1. Thus the estimates of the conditional first and second moments of returns are again very persistent. The value of $\hat{\psi}$ remains within a relatively narrow band from around 12 to 17. This is also the true for the value of $\hat{\delta}$, which remains within a band of 0.2 to 2.5. In contrast, the value of $\hat{\rho}$ spans a relatively wide range: 23% to 61%.

If we compare the plot in panel C to the other two plots in Figure 2, we see that the values for the 30 Momentum/Volatility dataset look much more like those for the 25 Size/Book-to-Market dataset than those for the 10 Industry dataset. Perhaps this reflects the presence of common risk factors. Figure 1 suggests that both the size/book-to-market and momentum/volatility sorting rules do a reasonable job of grouping firms with similar conditional expected returns together. If the dynamics of conditional expected returns are driven by a set of common risk factors, it would not be surprising to find that the optimal values of the tuning parameters for the two datasets are similar.

In Figure 3, panel C, we plot the minimum weight, the maximum weight, and the sum of the negative weights for each month in the performance evaluation period. The minimum weight is typically between -10% and -40% , while the maximum weight is typically between 20% and 40% . The sum of the negative weights stays around -100% for the majority of the performance evaluation period. However, it is in the -100% to -200% range for most of the 1990's. The increase in leverage during this decade appears to be a response to the strong bull-market conditions that preceded the collapse of the tech bubble.

To more clearly illustrate the role of the sorting rule in generating our results, Figure 4 provides additional information about the properties of the estimated weights for this dataset. In panel A we plot the time series obtained by aggregating the estimated weights within each of the three momentum categories. This shows how funds are allocated across low-, intermediate- and high-momentum stocks without regard to volatility. We find that there is a monotonic relation between the aggregate weights and momentum characteristics. The aggregated weight for low-momentum stocks is always negative, the aggregate weight for high-momentum stocks is always positive, and the aggregate weight for intermediate-momentum stocks always lies between the other two.

In panel B we plot the time series obtained by aggregating the estimated weights within each of the ten volatility categories. This shows how funds are allocated across the volatility spectrum without regard to momentum. We find that the lower volatility stocks generally receive higher aggregate weights. The lowest volatility decile stands out in this regard, receiving an aggregate weight in the 60% to 100% range. In contrast, the highest volatility

deciles typically receive a weight close to zero. The exception to this is during the 1990s when an increasing long position in the most volatile decile is balanced against an increasing short position in decile nine. These results suggest that the sorting rule plays a significant role in determining the performance of the sample UMVE portfolios for this dataset.

3.7 Incorporating short-sales constraints

A potential criticism of the results presented in Tables 1–3 is that all investors face some limits on selling securities short. We can address this criticism by modifying the framework developed in Section 2 to include weight constraints. To illustrate, suppose that short sales are prohibited. If we take the values of ψ , ϕ , and φ as given, we can use numerical methods to solve the quadratic program

$$\begin{aligned} \max_{w_t} \quad & w_t' \hat{\mu}_t(\phi) - \frac{\psi}{2} w_t' \hat{\Omega}_t(\varphi) w_t, \\ \text{s.t.} \quad & w_t' \iota = 1, \\ & w_{it} \geq 0, \quad i = 1, 2, \dots, N, \end{aligned} \tag{37}$$

for any period t . The solution to this program is the vector of plug-in weights for period t under the long-only constraint. Accordingly, we construct the long-only version of the sample objective function for period T_0 by using these weights in place of the Ferson and Siegel (2001) weights for $t = K_0, K_0 + 1, \dots, T_0$, and proceed in an analogous fashion for each subsequent period. In all other respects, the analysis goes through unchanged. A similar approach can be used to accommodate more general weight constraints.

Table 4 summarizes the results obtained for each dataset under a long-only constraint for the shrinkage-estimators case. We begin with the results in panel A for the 10 Industry dataset. Prohibiting short sales clearly has a meaningful impact on the performance of the sample UMVE portfolios. For example, setting $c = 50$ bp, $\delta \geq 0$, and assuming transaction costs of 50 bp for performance evaluation yields an estimated Sharpe ratio of 0.79 and an estimated CE return of 2.7%, substantially lower than the corresponding values of 1.04 and 5.2% reported in Table 1. Nonetheless, the long-only sample UMVE portfolio still outperforms the S&P 500 index and naïve diversification at the 1% significance level and the long-only GMV portfolio at the 10% significance level. Thus the leverage generated by short sales does not appear to be an overriding factor in determining the performance of mean-variance optimization for this dataset.

The results in panel B show that prohibiting short sales has a larger effect on portfolio performance for the 25 Size/Book-to-Market dataset. For example, setting $c = 50$ bp, $\delta \geq 0$, and assuming transaction costs of 50 bp for performance evaluation yields an estimated Sharpe ratio of 0.64 and an estimated CE return of -3.5% , as opposed to the values of 1.11 and 7.4% reported in Table 2. Although the estimated Sharpe ratio of the long-only sample UMVE portfolio is higher than that of any of the benchmarks, it is not high enough to conclude that the performance advantage is statistically significant. We noted earlier that the unconstrained portfolios display considerably higher leverage for the 25 Size/Book-to-Market dataset than for the 10 Industry dataset. The evidence in Table 4 suggests that high

leverage is a key contributor to the performance of the portfolio in the unconstrained case for this dataset.

In terms of methodology, the results in panel C for the 30 Momentum/Volatility dataset are the most interesting. Setting $c = 50$ bp, $\delta \geq 0$, and assuming transaction costs of 50 bp for performance evaluation yields an estimated Sharpe ratio of 1.04 and an estimated CE return of 8.7%. The values of 1.64 and 14.3% reported in Table 3 are of course more impressive, but it is noteworthy that the long-only sample UMVE portfolio outperforms all of the benchmarks at the 1% significance level, and it does so with an estimated expected turnover of less than 10% per year. This supports our contention that expanding the scope of the optimization problem to include the sorting scheme enhances portfolio performance.

We should also point out that in no instance does imposing the long-only constraint lead to an increase in the estimated Sharpe ratio of the portfolio. It seems obvious that we should expect this to be the case. However, some studies in the literature report the opposite finding (e.g., see Table 3 of DeMiguel et al., 2009). This again raises questions about the efficacy of the research designed used in these studies. It would be surprising to find that imposing additional constraints on the portfolio improves its performance when using a robust approach for exploiting sample information.

3.8 Reward and risk analysis using factor models

The analysis thus far has examined the performance of the sample UMVE portfolios relative to a number of common benchmarks. To develop a more complete picture of the performance landscape, we use a linear factor model to analyze the reward and risk characteristics of the portfolios. Specifically, we fit the Carhart (1997) extension of the Fama and French (1993) three-factor model to the excess returns on the sample UMVE portfolios. The Carhart (1997) model includes market, size, book-to-market, and momentum factors.¹⁹ In the interest of space, we focus on the case using shrinkage estimators, with $c = 50$ bp, $\delta \geq 0$, and assuming transaction costs of 50 bp for performance evaluation.

Table 5 summarizes the results of the factor model regressions. We begin with panel A, which presents results for the scenario with unconstrained weights. The estimated annualized alphas of the sample UMVE portfolios range from 4.6% for the 10 Industry dataset to 11.2% for the 30 Volatility/Momentum dataset; all of the estimates are statistically significant. Hence, the evidence indicates that the average returns for the sample UMVE portfolios contain a component that is not captured by this model. This finding is not surprising. In general, we would not expect an unconditional factor model to fully explain the average returns for the sample UMVE portfolios because these portfolios have time-varying weights and hence time-varying expected returns.

The estimated factor loadings and R^2 values for the three datasets are more interesting. All of the estimated loadings for the 10 Industry dataset are positive and statistically significant. However, the largest is only 0.6, indicating that the sample UMVE portfolio returns

¹⁹ The monthly factor returns and one-month T-bill rate are obtained from the Ken French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

for this dataset are relatively insensitive to these sources of systematic risk. In addition, the R^2 for the regression is only 62%, so a substantial portion of the risk of the portfolio is idiosyncratic. Like the estimated alphas, this is consistent with the presence of time-varying factor exposures that cannot be captured using an unconditional factor model. The regressions for the 25 Size/Book-to-Market and 30 Volatility/Momentum datasets also produce relatively low R^2 values: 65% and 59%. But not all of the estimated factor loadings are statistically significant. The sample UMVE portfolio for the 25 Size/Book-to-Market has no statistically significant exposure to the size and momentum factors, and the portfolio for the 30 Volatility/Momentum dataset has no statistically significant exposure to the size factor.

Panel B shows how the results change under the long-only constraint. First, the estimated alphas decrease to 1.7% for the 10 Industry and the 25 Size/Book-to-Market datasets and to 3.7% for the 30 Volatility/Momentum dataset. Imposing the constraint reduces the magnitude of the unexplained component of the average portfolio returns. Second, the regression R^2 increases to 78% for the 10 Industry and 30 Volatility/Momentum datasets, and to 92% for the 25 Size/Book-to-Market. Imposing the constraint reduces the idiosyncratic component of risk. Third, all of the estimated factor loadings are statistically significant.

3.9 Additional robustness checks

Before concluding we discuss two additional robustness checks to address potential concerns about our research design. First, we determine if our findings are overly sensitive to the value of γ used to construct the sample objective function. Although our choice of γ will clearly affect the aggressiveness of the sample UMVE portfolios, we want to exclude the possibility that this alters the overall conclusions about the performance of our methodology. To do so, we repeat the empirical analysis using $\gamma = 5$ instead of $\gamma = 15$. As expected, the results are indicative of a much more aggressive investment style. For example, if we set $c = 50$ bp, $\delta \geq 0$, assume transaction costs of 50 bp for performance evaluation, and use shrinkage estimators of the conditional moments of returns, the sample UMVE portfolio for the 10 Industry dataset has an estimated expected return of 38.8%, an estimated standard deviation of 27.1%, an estimated Sharpe ratio of 1.23, an estimated CE return of 20.4%, and an estimated expected turnover of 818% per year. But changing the value of γ does not affect our general conclusions regarding the performance of the sample UMVE portfolio. It still outperforms all three of the benchmarks at the 1% significance level.

Second, we determine whether using equally-weighted portfolios as assets has a undue influence on the results. Perhaps this explains why our methodology performs so well in comparison to the methodologies employed in prior studies that use value-weighted portfolios as assets. To exclude this possibility, we repeat the empirical analysis for versions of the 10 Industry and 25 Size/Book-to-Market datasets that contain value-weighted portfolios. This does produce some changes in the results. For instance, both the estimated Sharpe ratios and estimated expected turnover figures for the 10 Industry dataset are lower than those reported in Table 1. But the performance comparisons remain favorable to the sample UMVE portfolios. For example, if we use shrinkage estimators, set $c = 50$ bp, $\delta \geq 0$, and assume transaction costs of 50 bp for performance evaluation, the sample UMVE portfolio for the 25 Size/BTM dataset has an estimated expected return of 17%, an estimated standard de-

viation of 13%, an estimated Sharpe ratio of 0.89, an estimated CE return of 4.6%, and an estimated expected turnover of 84% per year. This is sufficient to outperform all three benchmarks at the 5% level. Thus the choice of weighting scheme does not seem to be a critical issue.

4 Closing Remarks

Mean-variance portfolio optimization has recently come under fire for its ostensibly poor performance in out-of-sample applications. Much of this recent criticism is motivated by DeMiguel et al. (2009). The authors of this study construct sample mean-variance efficient portfolios using several variants of the plug-in approach and find that naïve diversification outperforms these portfolios for most of the datasets considered. This poor showing by the sample mean-variance efficient portfolios, which is interpreted as evidence that “the errors in estimating means and covariances erode all the gains from optimal, relative to naïve, diversification,” leads the authors to conclude that “there are still many ‘miles to go’ before the gains promised by optimal portfolio choice can actually be realized out of sample.”

Our investigation supports a sharply contrasting view of the out-of-sample performance of mean-variance optimization. We begin by noting that many aspects of the plug-in approach, such as how to model changes in the investment opportunity set and how to estimate the model parameters, fall outside the scope of the traditional optimization framework. This gives rise to significant robustness issues because *ad hoc* choices with respect to the details of the research design can lead to unforeseen consequences. For example, DeMiguel et al. (2009) focus on the performance of the sample version of the tangency portfolio, which is obtained by implementing the plug-in approach using a time-varying value of relative risk aversion. Although in principle this is fine, it leads to extreme turnover and poor performance because the sample tangency portfolio frequently targets a conditional expected return in excess of 100% per year (Kirby and Ostdiek, 2012).

To address the optimization problem in a more comprehensive fashion, we expand the scope of the analysis to encompass the effects of estimation risk, specification errors, and transaction costs on portfolio performance, and we use an adaptive empirical procedure to select the values of the unknown parameters that appear in the expression for the plug-in weights. The empirical analysis demonstrates that our methodology leads to robust portfolio selection rules, regardless of whether transaction costs are high or low. The resulting sample UMVE portfolios have well-behaved weights, reasonable turnover, and substantially higher estimated Sharpe ratios and certainty-equivalent returns than common performance benchmarks such as the $1/N$ portfolio and S&P 500 index. This is true regardless of whether we permit short sales or impose a long-only constraint.

As further evidence on the value of taking a comprehensive approach to the optimization problem, we show that constructing the dataset by sorting firms into portfolios on the basis of historical measures of momentum and volatility leads to sample UMVE portfolios that perform particularly well. Indeed, the sample UMVE portfolios for the 30 Momentum/Volatility dataset outperform the S&P 500 index, $1/N$ portfolio, and a simple plug-in version of the GMV portfolio at the 1% significance level. This finding is noteworthy because prior studies

report that the GMV portfolio performs well relative to other sample efficient portfolios. In view of this finding and the evidence in general, we believe the proposed methodology represents a significant step forward in addressing the challenges of mean-variance portfolio choice in the presence of estimation risk, model misspecification, and transaction costs.

With respect to future research, there are a number of interesting ways to extend the methodology. One possibility is to expand the information set by allowing the conditional moments of returns to depend on both past returns and other financial variables, such as interest rates, dividend yields and credit spreads. Another is to further expand the scope of the optimization problem by allowing the number of portfolios N under a given sorting rule to be an additional tuning parameter. This would allow the data to inform our choice of the scale of the optimization problem rather than imposing the scale *a priori*. We are currently pursuing ideas along these lines.

References

- Brandt, M., P. Santa-Clara, and R. Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447
- Carhart, M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.
- Chan, L., J. Karceski, and J. Lakonishok, 1999, On portfolio optimization: Forecasting covariances and choosing the risk model, *Review of Financial Studies* 12, 937–974.
- DeMiguel, V., L. Garlappi, and R. Uppal, 2009, Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy, *Review of Financial Studies* 22, 1915–1953.
- Domowitz, I., J. Glen, and A. Madhavan, 2001, Liquidity, volatility, and equity trading costs across countries and over time, *International Finance* 4, 221–255.
- Dybvig, P., and S. Ross, 1985, Differential information and performance measurement using a security market line, *Journal of Finance* 40, 383–399.
- Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Ferson, W., and A. Siegel, 2001, The efficient use of conditioning information in portfolios, *Journal of Finance* 56, 967–982.
- Foster, D., and D. Nelson, 1996, Continuous record asymptotics for rolling sample variance estimators, *Econometrica* 64, 139–174.
- Griffin, J., and J. Xu, 2007, How Smart are the Smart Guys? A Unique View from Hedge Fund Stock Holdings, *Review of Financial Studies* 22, 2531–2570.
- Hansen, L., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.
- Hansen, L., and S. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.
- Hasbrouck, J., 2009, Trading costs and returns for US equities: Estimating effective costs from daily data, *Journal of Finance* 64, 1445–1477.
- Jagannathan, R., and T. Ma., 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1684.
- James, W., and C. Stein, 1961, Estimation with quadratic loss, *Proceedings of the Fourth Berkeley Symposium on Mathematics and Statistics*, 361–379.
- Kan, R., and G. Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621–656.
- Kirby, C., and B. Ostdiek, 2012, It’s all in the timing: Simple active portfolio strategies that outperform naïve diversification, *Journal of Financial and Quantitative Analysis* 47, 437–467
- Ledoit, O., and M. Wolf, 2004, Honey, I shrunk the sample covariance matrix: Problems in mean-variance optimization, *Journal of Portfolio Management* 30, 110–119.

- Leland, H., 1999, Optimal portfolio management with transactions costs and capital gains taxes, working paper, University of California Berkeley.
- Merton, R., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.
- Patton, A., and A. Timmermann, 2007, Properties of optimal forecasts under asymmetric loss and nonlinearity, *Journal of Econometrics* 140, 884–918.
- Paye, B., 2010, Portfolios of estimated portfolios: Combination approaches to estimating optimal portfolio weights, working paper, University of Georgia.
- Skouras, S., 2007. Decisionmetrics: A decision-based approach to econometric modelling, *Journal of Econometrics* 137, 414-440.
- Tu, J., and G. Zhou, 2011, Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies, *Journal of Financial Economics* 99, 204–215.

Table 1

Performance of Sample UMVE Portfolios for the 10 Industry Dataset

The table documents the performance of the sample UMVE portfolios for the 10 Industry dataset. Results for the three benchmarks (S&P 500, 1/N, and plug-in GMV portfolios) are presented above panel A. For each portfolio, we report annualized estimates of the expected turnover ($\hat{\tau}_p$), the expected return ($\hat{\mu}_p$), return volatility ($\hat{\sigma}_p$), Sharpe ratio ($\hat{\lambda}_p$), and certainty-equivalent return (\hat{Q}_p). The p -values indicate the probability of obtaining a higher $\hat{\lambda}_p$ under the null hypothesis that the corresponding benchmark portfolio performs at least as well as the UMVE portfolio. The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma = 15$ with exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. The analysis is conducted with both full- and partial-adjustment of the portfolio weights ($\delta = 0$ and $\delta \geq 0$) using two levels of proportional transaction costs to construct the sample objective function ($c = 5$ and $c = 50$ basis points). For each value of c , we report $\hat{\mu}_p$, $\hat{\sigma}_p$, $\hat{\lambda}_p$, and \hat{Q}_p using returns measured net of proportional one-way transaction costs of both five and 50 basis points. The performance evaluation period is January 1976 to December 2009 (408 months).

	$\hat{\tau}_p$	Imposing Transaction Costs of 5 bp					Imposing Transaction Costs of 50 bp								
		$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV
S&P 500	0	11.84	15.27	0.41	-5.65	—	0.879	0.543	11.84	15.27	0.41	-5.65	—	0.868	0.329
1/N	14	16.26	19.46	0.55	-12.15	0.121	—	0.299	16.14	19.46	0.54	-12.26	0.132	—	0.162
GMV	138	10.63	11.83	0.43	0.13	0.457	0.701	—	9.39	11.87	0.33	-1.18	0.671	0.838	—
Panel A: No Shrinkage															
$c = 5$ bp, $\delta = 0$	1424	33.27	16.80	1.65	12.09	0.000	0.000	0.000	20.19	17.33	0.85	-2.32	0.014	0.085	0.002
$c = 5$ bp, $\delta \geq 0$	1424	33.27	16.80	1.65	12.09	0.000	0.000	0.000	20.19	17.33	0.85	-2.32	0.014	0.085	0.002
$c = 50$ bp, $\delta = 0$	381	21.61	13.36	1.21	8.23	0.000	0.000	0.000	18.14	13.47	0.94	4.53	0.001	0.022	0.000
$c = 50$ bp, $\delta \geq 0$	336	21.96	13.84	1.19	7.60	0.000	0.000	0.000	18.90	13.95	0.96	4.32	0.001	0.016	0.000
Panel B: Using Shrinkage Estimators															
$c = 5$ bp, $\delta = 0$	1381	36.50	17.91	1.73	12.43	0.000	0.000	0.000	23.82	18.61	0.98	-2.15	0.003	0.021	0.000
$c = 5$ bp, $\delta \geq 0$	1381	36.50	17.91	1.73	12.43	0.000	0.000	0.000	23.82	18.61	0.98	-2.15	0.003	0.021	0.000
$c = 50$ bp, $\delta = 0$	326	22.97	13.70	1.28	8.90	0.000	0.000	0.000	20.02	13.87	1.05	5.59	0.000	0.002	0.000
$c = 50$ bp, $\delta \geq 0$	286	22.73	13.93	1.24	8.18	0.000	0.000	0.000	20.14	14.13	1.04	5.18	0.000	0.003	0.000

Table 2

Performance of Sample UMVE Portfolios for the 25 Size/Book-to-Market Dataset

The table documents the performance of the sample UMVE portfolios for the 25 Size/Book-to-Market dataset. Results for the three benchmarks (S&P 500, 1/N, and plug-in GMV portfolios) are presented above panel A. For each portfolio, we report annualized estimates of the expected turnover ($\hat{\tau}_p$), the expected return ($\hat{\lambda}_p$), return volatility ($\hat{\sigma}_p$), Sharpe ratio ($\hat{\lambda}_p$), and certainty-equivalent return (\hat{Q}_p). The p -values indicate the probability of obtaining a higher $\hat{\lambda}_p$ under the null hypothesis that the corresponding benchmark portfolio performs at least as well as the UMVE portfolio. The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma = 15$ with exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. The analysis is conducted with both full- and partial-adjustment of the portfolio weights ($\delta = 0$ and $\delta \geq 0$) using two levels of proportional transaction costs to construct the sample objective function ($c = 5$ and $c = 50$ basis points). For each value of c , we report $\hat{\mu}_p$, $\hat{\sigma}_p$, $\hat{\lambda}_p$, and \hat{Q}_p using returns measured net of proportional one-way transaction costs of both five and 50 basis points. The performance evaluation period is January 1976 to December 2009 (408 months).

	$\hat{\tau}_p$	Imposing Transaction Costs of 5 bp					Imposing Transaction Costs of 50 bp								
		$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV
S&P 500	0	11.84	15.27	0.41	-5.65	—	0.956	1.000	11.84	15.27	0.41	-5.65	—	0.950	0.996
1/N	11	15.83	18.00	0.57	-8.47	0.044	—	1.000	15.73	18.00	0.57	-8.57	0.050	—	0.956
GMV	495	21.22	12.85	1.22	8.84	0.000	0.000	—	16.70	12.90	0.86	4.21	0.004	0.044	—
Panel A: No Shrinkage															
$c = 5$ bp, $\delta = 0$	464	21.93	12.61	1.30	10.01	0.000	0.000	0.150	17.72	12.84	0.95	5.35	0.001	0.017	0.131
$c = 5$ bp, $\delta \geq 0$	459	21.92	12.62	1.29	9.98	0.000	0.000	0.157	17.75	12.85	0.95	5.36	0.001	0.017	0.127
$c = 50$ bp, $\delta = 0$	286	19.90	12.36	1.16	8.44	0.000	0.000	0.794	17.30	12.43	0.95	5.71	0.000	0.008	0.134
$c = 50$ bp, $\delta \geq 0$	137	20.47	13.08	1.14	7.64	0.000	0.000	0.833	19.24	13.26	1.03	6.05	0.000	0.003	0.027
Panel B: Using Shrinkage Estimators															
$c = 5$ bp, $\delta = 0$	372	22.68	12.69	1.34	10.61	0.000	0.000	0.101	19.33	13.10	1.05	6.45	0.000	0.003	0.034
$c = 5$ bp, $\delta \geq 0$	370	22.66	12.69	1.34	10.59	0.000	0.000	0.104	19.33	13.11	1.05	6.44	0.000	0.003	0.033
$c = 50$ bp, $\delta = 0$	176	21.07	12.60	1.23	9.15	0.000	0.000	0.474	19.47	12.73	1.09	7.32	0.000	0.001	0.018
$c = 50$ bp, $\delta \geq 0$	99	20.82	12.77	1.19	8.59	0.000	0.000	0.608	19.93	12.91	1.11	7.43	0.000	0.001	0.013

Table 3

Performance of Sample UMVE Portfolios for the 30 Momentum/Volatility Dataset

The table documents the performance of the sample UMVE portfolios for the 30 Momentum/Volatility dataset. Results for the three benchmarks (S&P 500, 1/N, and plug-in GMV portfolios) are presented above panel A. For each portfolio, we report annualized estimates of the expected turnover ($\hat{\tau}_p$), the expected return ($\hat{\mu}_p$), return volatility ($\hat{\sigma}_p$), Sharpe ratio ($\hat{\lambda}_p$), and certainty-equivalent return (\hat{Q}_p). The p -values indicate the probability of obtaining a higher $\hat{\lambda}_p$ under the null hypothesis that the corresponding benchmark portfolio performs at least as well as the UMVE portfolio. The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma = 15$ with exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. The analysis is conducted with both full- and partial-adjustment of the portfolio weights ($\delta = 0$ and $\delta \geq 0$) using two levels of proportional transaction costs to construct the sample objective function ($c = 5$ and $c = 50$ basis points). For each value of c , we report $\hat{\mu}_p$, $\hat{\sigma}_p$, $\hat{\lambda}_p$, and \hat{Q}_p using returns measured net of proportional one-way transaction costs of both five and 50 basis points. The performance evaluation period is January 1976 to December 2009 (408 months).

	Imposing Transaction Costs of 5 bp					Imposing Transaction Costs of 50 bp									
	$\hat{\tau}_p$	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. 1/N	p vs. GMV
S&P 500	0	11.84	15.27	0.41	-5.65	—	0.905	0.840	11.84	15.27	0.41	-5.65	—	0.896	0.232
1/N	14	16.85	19.88	0.57	-12.79	0.095	—	0.658	16.73	19.88	0.56	-12.90	0.104	—	0.096
GMV	476	12.05	9.67	0.68	5.04	0.160	0.342	—	7.75	9.85	0.23	0.48	0.768	0.904	—
Panel A: No Shrinkage															
$c = 5$ bp, $\delta = 0$	344	27.57	12.49	1.76	15.87	0.000	0.000	0.000	24.45	12.78	1.48	12.21	0.000	0.001	0.000
$c = 5$ bp, $\delta \geq 0$	344	27.57	12.49	1.76	15.87	0.000	0.000	0.000	24.45	12.78	1.48	12.21	0.000	0.001	0.000
$c = 50$ bp, $\delta = 0$	158	22.06	10.20	1.62	14.25	0.000	0.000	0.000	20.64	10.38	1.46	12.56	0.000	0.001	0.000
$c = 50$ bp, $\delta \geq 0$	75	23.68	11.22	1.62	14.24	0.000	0.000	0.000	23.01	11.31	1.54	13.42	0.000	0.000	0.000
Panel B: Using Shrinkage Estimators															
$c = 5$ bp, $\delta = 0$	230	28.58	13.54	1.71	14.84	0.000	0.000	0.000	26.51	13.78	1.53	12.26	0.000	0.000	0.000
$c = 5$ bp, $\delta \geq 0$	230	28.58	13.54	1.70	14.84	0.000	0.000	0.000	26.51	13.78	1.53	12.26	0.000	0.000	0.000
$c = 50$ bp, $\delta = 0$	110	24.81	11.90	1.63	14.18	0.000	0.000	0.000	23.81	12.02	1.53	12.98	0.000	0.000	0.000
$c = 50$ bp, $\delta \geq 0$	69	25.94	12.44	1.64	14.33	0.000	0.000	0.000	25.33	12.55	1.58	13.52	0.000	0.000	0.000

Table 4

Performance of Sample UMVE Portfolios under a Long-Only Constraint

The table documents the performance of sample long-only UMVE portfolios using a relative risk aversion of $\gamma = 15$ and shrinkage estimators. For each portfolio, we report annualized estimates of the expected turnover ($\hat{\tau}_p$), the expected return ($\hat{\mu}_p$), return volatility ($\hat{\sigma}_p$), Sharpe ratio ($\hat{\lambda}_p$), and certainty-equivalent return (\hat{Q}_p). The p -values indicate the probability of obtaining a higher $\hat{\lambda}_p$ under the null hypothesis that the corresponding benchmark portfolio (S&P 500, $1/N$, or plug-in long-only GMV) performs at least as well as the UMVE portfolio. The sample UMVE portfolios are constructed using exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. The analysis is conducted with both full- and partial-adjustment of the portfolio weights ($\delta = 0$ and $\delta \geq 0$) using two levels of proportional transaction costs to construct the sample objective function ($c = 5$ and $c = 50$ basis points). For each value of c , we report $\hat{\mu}_p$, $\hat{\sigma}_p$, $\hat{\lambda}_p$, and \hat{Q}_p using returns measured net of proportional one-way transaction costs of both five and 50 basis points. The performance evaluation period is January 1976 to December 2009 (408 months).

	Imposing Transaction Costs of 5 bp					Imposing Transaction Costs of 50 bp									
	$\hat{\tau}_p$	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. $1/N$	p vs. GMV	$\hat{\mu}_p$	$\hat{\sigma}_p$	$\hat{\lambda}_p$	\hat{Q}_p	p vs. Mkt	p vs. $1/N$	p vs. GMV
Panel A: 10 Industry Dataset															
$c = 5$ bp, $\delta = 0$	301	20.28	15.17	0.97	3.01	0.000	0.000	0.012	17.55	15.28	0.79	0.04	0.003	0.018	0.198
$c = 5$ bp, $\delta \geq 0$	303	20.24	15.23	0.97	2.84	0.000	0.000	0.013	17.49	15.34	0.78	-0.16	0.003	0.020	0.215
$c = 50$ bp, $\delta = 0$	105	17.11	13.33	0.87	3.79	0.000	0.002	0.019	16.17	13.38	0.80	2.75	0.002	0.012	0.094
$c = 50$ bp, $\delta \geq 0$	102	17.07	13.33	0.87	3.75	0.000	0.002	0.021	16.15	13.38	0.79	2.73	0.001	0.012	0.097
Panel B: 25 Size/Book-to-Market Dataset															
$c = 5$ bp, $\delta = 0$	339	15.71	16.32	0.62	-4.28	0.020	0.208	0.317	12.64	16.45	0.43	-7.65	0.430	0.980	0.972
$c = 5$ bp, $\delta \geq 0$	311	15.94	16.48	0.63	-4.43	0.016	0.162	0.271	13.13	16.59	0.46	-7.52	0.332	0.954	0.940
$c = 50$ bp, $\delta = 0$	74	15.55	15.90	0.63	-3.42	0.023	0.171	0.164	14.89	15.91	0.59	-4.10	0.054	0.366	0.273
$c = 50$ bp, $\delta \geq 0$	17	15.95	16.06	0.65	-3.40	0.018	0.123	0.111	15.80	16.05	0.64	-3.52	0.022	0.135	0.050
Panel C: 30 Momentum/Volatility Dataset															
$c = 5$ bp, $\delta = 0$	58	15.62	9.45	1.07	8.92	0.001	0.003	0.054	15.10	9.52	1.01	8.31	0.002	0.007	0.235
$c = 5$ bp, $\delta \geq 0$	28	15.86	9.63	1.07	8.90	0.001	0.002	0.045	15.61	9.66	1.04	8.61	0.001	0.004	0.049
$c = 50$ bp, $\delta = 0$	40	15.49	9.50	1.05	8.72	0.001	0.005	0.060	15.12	9.53	1.01	8.31	0.002	0.009	0.115
$c = 50$ bp, $\delta \geq 0$	8	15.56	9.53	1.05	8.74	0.001	0.004	0.038	15.49	9.55	1.04	8.65	0.001	0.004	0.006

Table 5**Factor Model Decomposition of Sample UMVE Portfolio Returns**

The table documents the results of fitting a linear four-factor risk model to the excess returns on the sample UMVE portfolios constructed using a relative risk aversion of $\gamma = 15$ and shrinkage estimators. The risk factors are the excess return on the S&P 500 index (MKT- r_f), the return on the small-minus-big (SMB) size portfolio, the return on high-minus-low (HML) book-to-market portfolio, and the return on the momentum (MOM) portfolio. We report the estimated intercept ($\hat{\alpha}$), the estimated factor loadings ($\hat{\beta}_i \forall i$), the corresponding standard errors (in parentheses), and the regression R^2 . The strategies are implemented using exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. The analysis is conducted with partial-adjustment of the portfolio weights ($\delta \geq 0$) using proportional transaction costs of $c = 50$ basis points to construct the sample objective function. The portfolio returns are measured net of proportional one-way transaction costs of 50 basis points. The performance evaluation period is January 1976 to December 2009 (408 months).

	$\gamma = 15, c = 50 \text{ bp}, \delta \geq 0$					
	$\hat{\alpha}$	$\hat{\beta}_1$ MKT- r_f	$\hat{\beta}_2$ SMB	$\hat{\beta}_3$ HML	$\hat{\beta}_4$ MOM	R^2
Panel A: Unconstrained Weights						
10 Industry Dataset	4.57 (1.74)	0.60 (0.04)	0.31 (0.08)	0.41 (0.08)	0.37 (0.05)	0.62
25 Size/Book-to-Market Dataset	7.18 (1.66)	0.66 (0.04)	0.02 (0.05)	0.61 (0.06)	0.00 (0.04)	0.65
30 Momentum/Volatility Dataset	11.18 (1.75)	0.53 (0.04)	0.02 (0.06)	0.43 (0.07)	0.36 (0.06)	0.59
Panel B: Long-Only Weights						
10 Industry Dataset	1.74 (1.34)	0.73 (0.03)	0.33 (0.05)	0.45 (0.05)	0.11 (0.03)	0.78
25 Size/Book-to-Market Dataset	1.68 (0.85)	0.95 (0.02)	0.33 (0.04)	0.49 (0.04)	-0.12 (0.02)	0.92
30 Momentum/Volatility Dataset	3.70 (0.96)	0.52 (0.03)	0.21 (0.04)	0.37 (0.03)	0.05 (0.03)	0.78

Figure 1

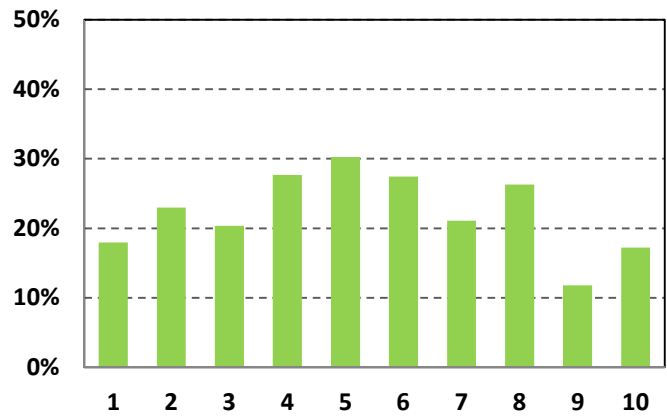
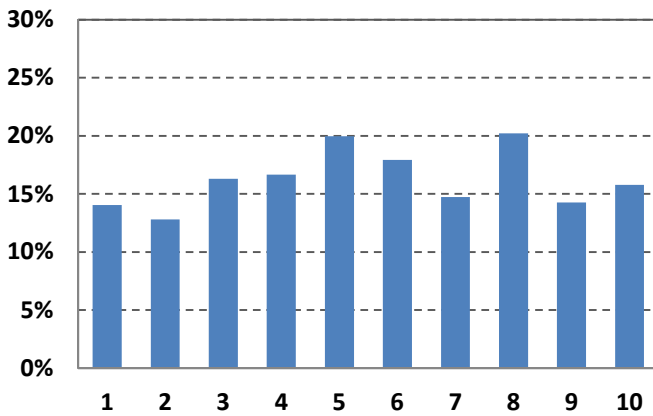
Reward and Risk Characteristics of the Datasets

The figure summarizes the sample reward and risk characteristics for the 10 Industry dataset (panel A), 25 Size/BTM dataset (panel B), and 30 Momentum/Volatility dataset (panel C). The first graph in each panel shows the cross-section of annualized mean returns and the second shows the cross-section of annualized return standard deviations. The full sample period is January 1946 to December 2009 (768 monthly observations). The reported statistics correspond to the performance evaluation period, i.e, observations 361 to 768.

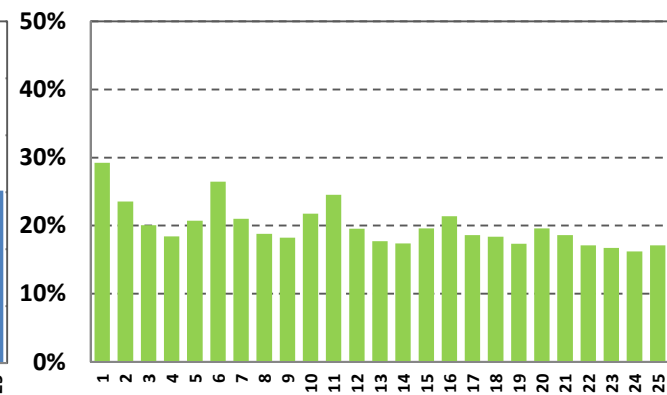
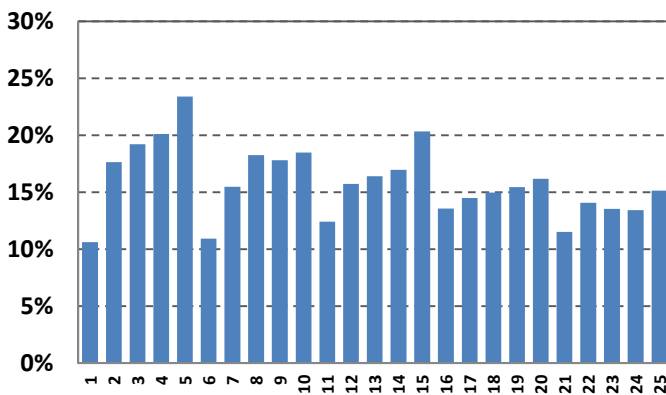
Mean Return

Volatility

Panel A. 10 Industry Portfolios



Panel B. 25 Size/Book-to-Market Portfolios



Panel C. 30 Momentum/Volatility Portfolios

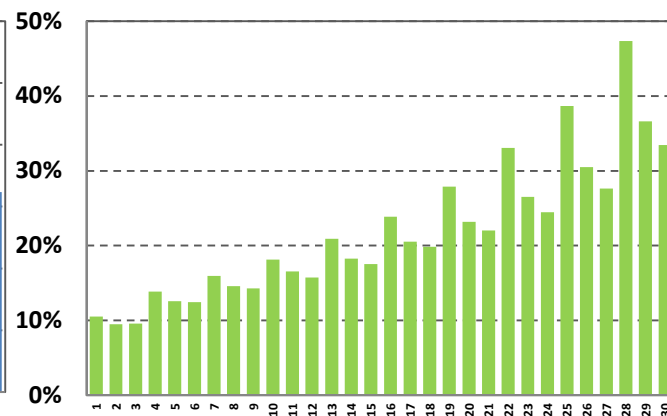
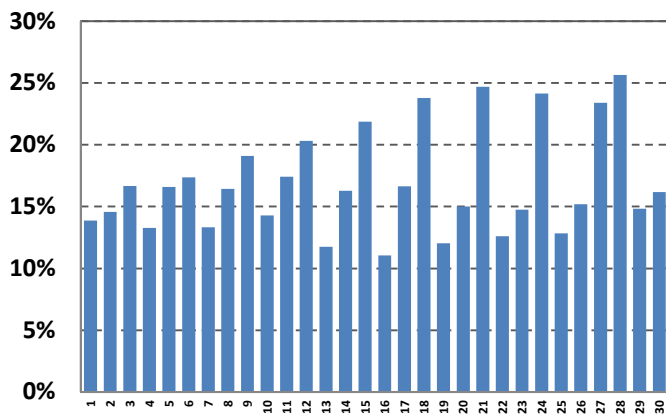
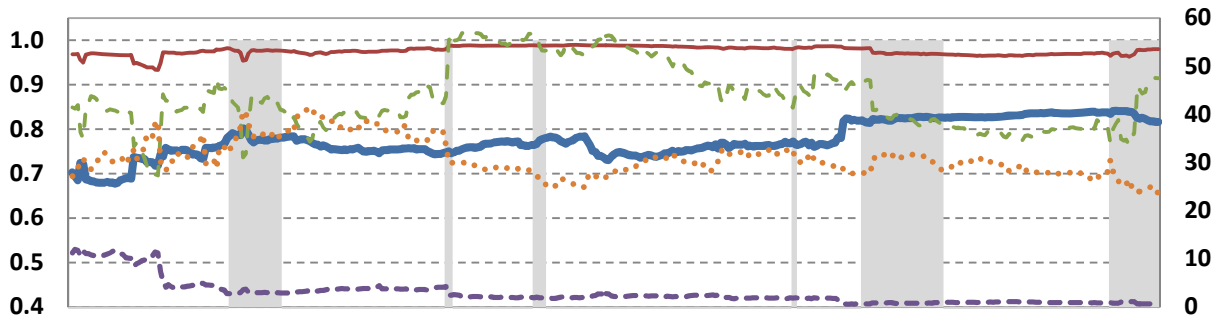


Figure 2

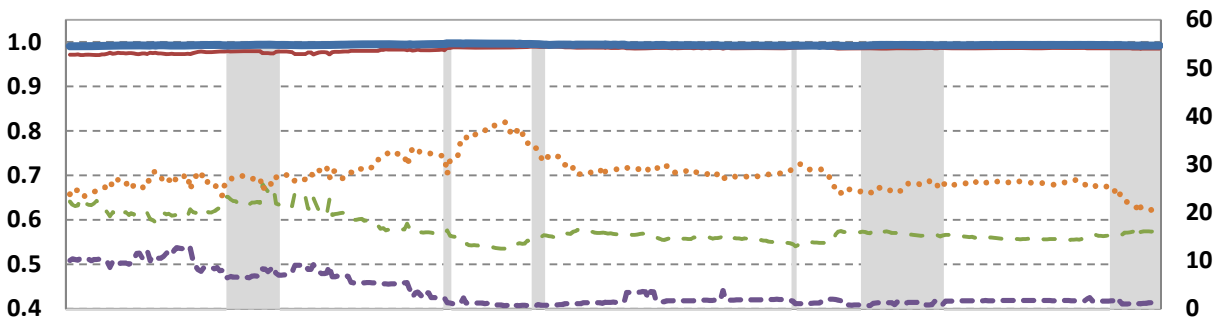
Estimated Optimal Tuning Parameter Values for the Sample UMVE Portfolios

The figure plots the times series of estimated optimal tuning parameter values for the 10 Industry dataset (panel A), 25 Size/BTM dataset (panel B) and 30 Momentum/Volatility dataset (panel C). The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma=15$, proportional one-way transactions costs of $c=50$ basis points, partial-adjustment of the portfolio weights ($\delta \geq 0$), and shrinkage versions of the exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. Each graph plots the time series of the estimated smoothing constants, ϕ and φ , (both left-hand scale), and the estimated risk penalty, ψ , no-trade distance, δ , and shrinkage factor, ρ , (all right-hand scale). We report the estimates of δ in annualized basis points and the estimates of ρ in percent. The shaded rectangles indicate bear markets (10% or greater declines in the S&P 500 index). The data begin in January 1946. The performance evaluation period is January 1976 to December 2009 (408 months).

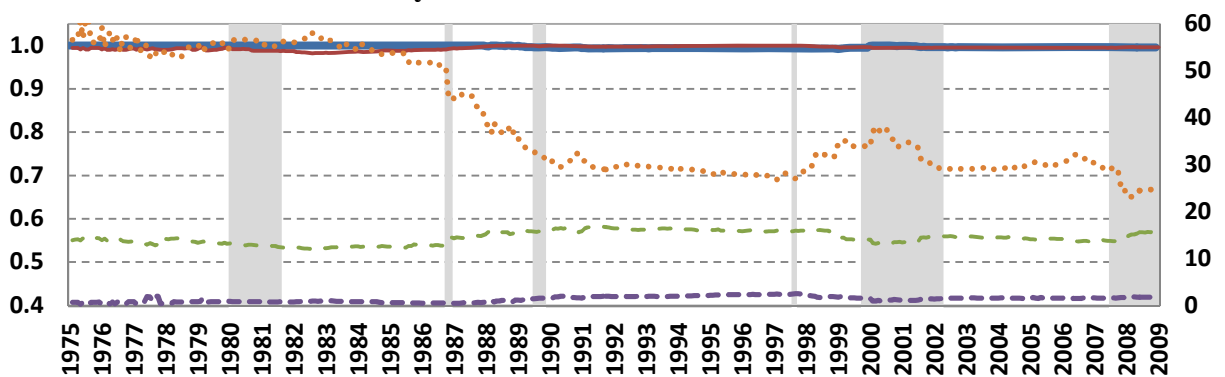
Panel A. 10 Industry Dataset



Panel B. 25 Size/Book-to-Market Dataset



Panel C. 30 Momentum/Volatility Dataset



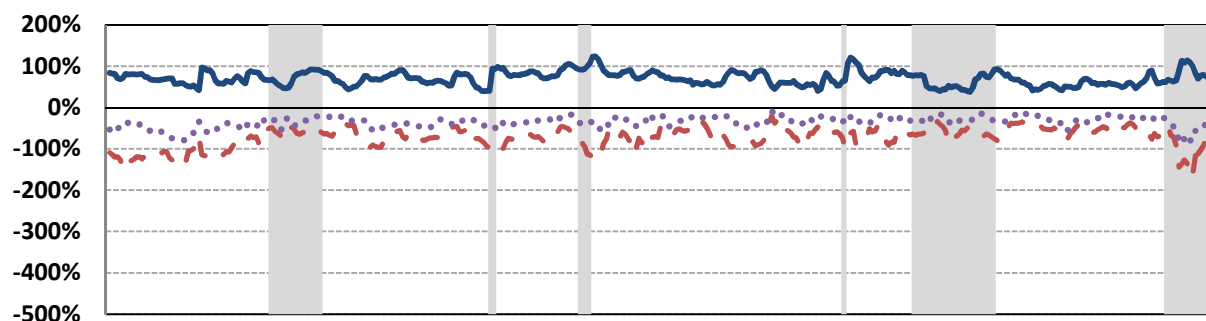
— ϕ (lhs)
 — φ
 - - - ψ (rhs)
 - - - $\delta * 120,000$ (rhs)
 . . . $\rho * 100$ (rhs)

Figure 3

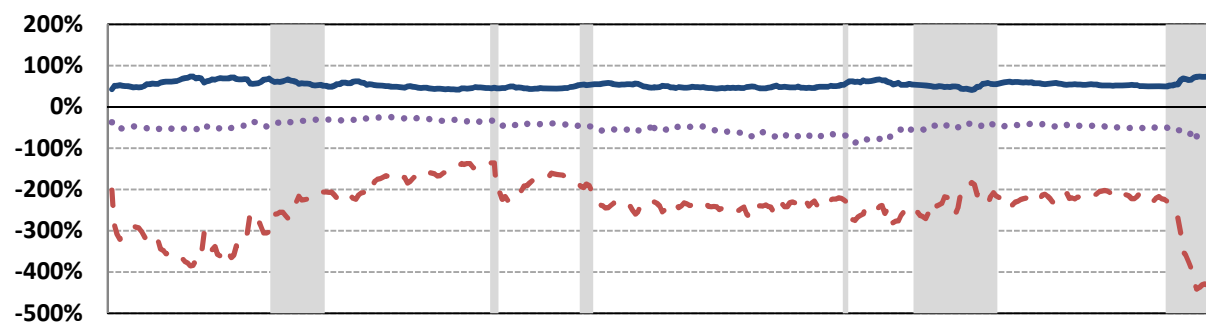
Characteristics of the Estimated Weights for the Sample UMVE Portfolios

The figure illustrates characteristics of the estimated optimal weights for the 10 Industry dataset (panel A), 25 Size/BTM dataset (panel B) and 30 Momentum/Volatility dataset (panel C). The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma=15$, proportional one-way transactions costs of $c=50$ basis points, partial-adjustment of the portfolio weights ($\delta \geq 0$), and shrinkage versions of the exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. Each graph plots the time series of the maximum estimated portfolio weight, the minimum estimated portfolio weight and the level of estimated portfolio leverage, defined as the total short position as a percent of portfolio value. The shaded rectangles indicate bear markets (10% or greater declines in the S&P 500 index). The data begin in January 1946. The performance evaluation period is January 1976 to December 2009 (408 months).

Panel A. 10 Industry Dataset



Panel B. 25 Size/Book-to-Market Dataset



Panel C. 30 Volatility/Momentum Dataset

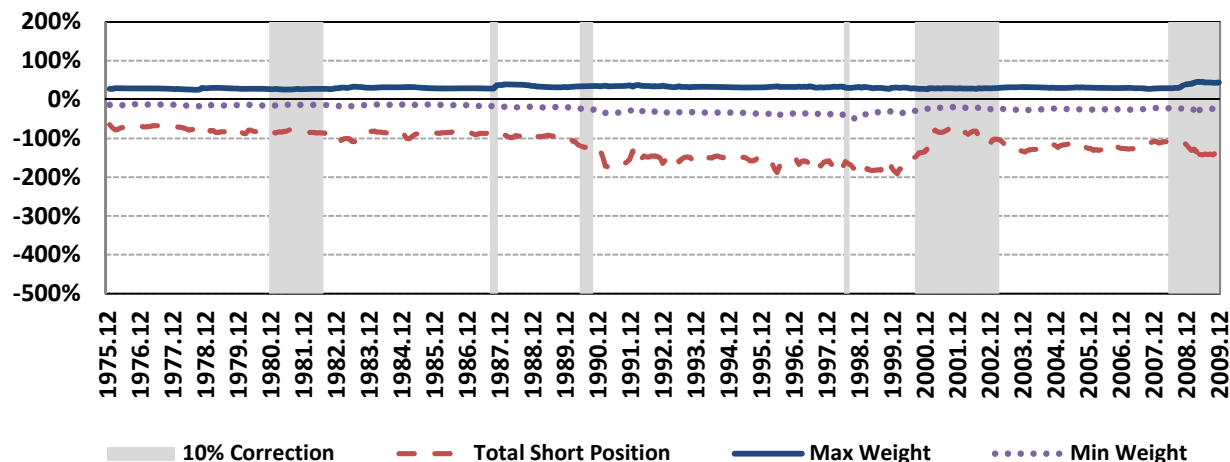
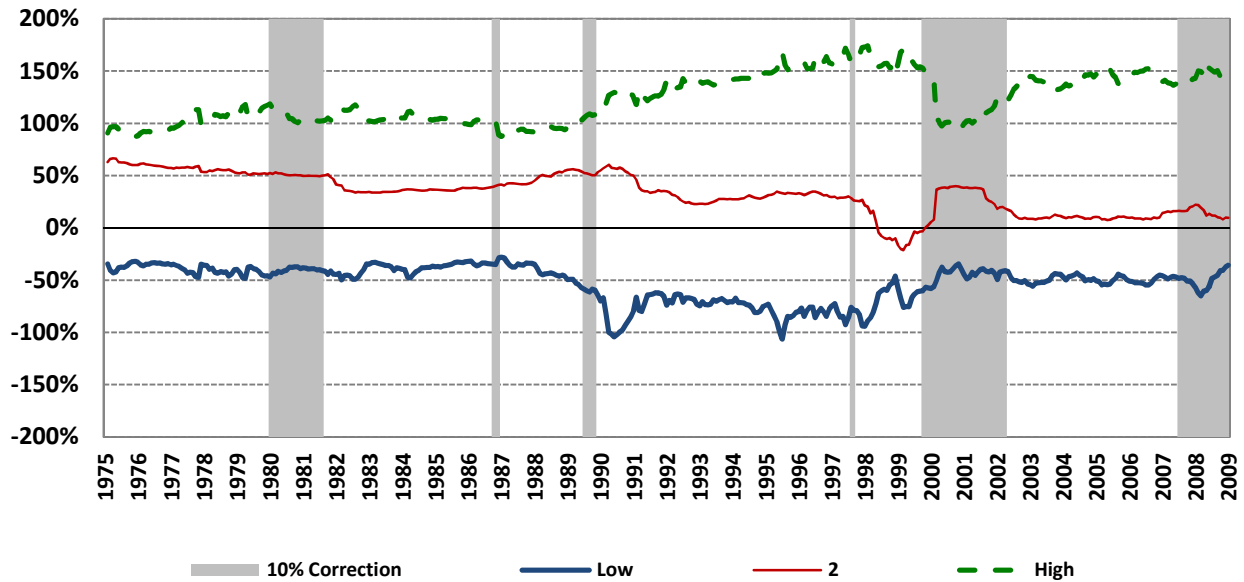


Figure 4

Estimated Weights as a Function of Momentum and Volatility Characteristics

The figure plots the time series of estimated optimal weights as a function of momentum and volatility characteristics for the 30 Momentum/Volatility dataset. The sample UMVE portfolios are constructed using a relative risk aversion of $\gamma=15$, proportional one-way transactions costs of $c=50$ basis points, partial-adjustment of the portfolio weights ($\delta \geq 0$), and shrinkage versions of the exponential smoothing estimators of the conditional mean vector and conditional second-moment matrix. All parameters in the expression for the optimal portfolio weights are estimated using an approach designed to minimize the adverse impact of estimation risk and rebalancing costs as described in the text. Panel A plots the time series obtained by aggregating the estimated weights within each of the three momentum categories. Panel B plots the time series obtained by aggregating the estimated weights within each of the ten volatility categories. The shaded rectangles indicate bear markets (10% or greater declines in the S&P 500 index). The data begin in January 1946. The performance evaluation period is January 1976 to December 2009 (408 months).

Panel A. Weights Aggregated by Momentum Category



Panel B. Weights Aggregated by Volatility Category

