OPERATIONS RESEARCH
Vol. 57, No. 3, May–June 2009, pp. 560–577
ISSN 0030-364X | EISSN 1526-5463 | 09 | 5703 | 0560

DOI 10.1287/opre.1080.0566

Portfolio Selection with Robust Estimation

Victor DeMiguel
Department of Management Science and Operations, London Business School,
London NW1 4SA, United Kingdom, avmiguel@london.edu

Francisco J. Nogales
Department of Statistics, Universidad Carlos III de Madrid, 28911-Leganés (Madrid),
Spain, fcojavier.nogales@uc3m.es

Mean-variance portfolios constructed using the sample mean and covariance matrix of asset returns perform poorly out of sample due to estimation error. Moreover, it is commonly accepted that estimation error in the sample mean is much larger than in the sample covariance matrix. For this reason, researchers have recently focused on the minimum-variance portfolio, which relies solely on estimates of the covariance matrix, and thus usually performs better out of sample. However, even the minimum-variance portfolios are quite sensitive to estimation error and have unstable weights that fluctuate substantially over time. In this paper, we propose a class of portfolios that have better stability properties than the traditional minimum-variance portfolios. The proposed portfolios are constructed using certain *robust* estimators and can be computed by solving a *single* nonlinear program, where robust estimation and portfolio optimization are performed in a single step. We show analytically that the resulting portfolio weights are less sensitive to changes in the asset-return distribution than those of the traditional portfolios. Moreover, our numerical results on simulated and empirical data confirm that the proposed portfolios are more stable than the traditional minimum-variance portfolios, while preserving (or slightly improving) their relatively good out-of-sample performance.

Subject classifications: finance: portfolio, investment; economics: econometrics; portfolio choice; minimum-variance portfolios; estimation error; robust statistics.
Area of review: Financial Engineering.
History: Received February 2007; revisions received August 2007, November 2007, December 2007; accepted December 2007. Published online in *Articles in Advance* February 26, 2009.

1. Introduction

An investor who cares only about the mean and variance of static portfolio returns should hold a portfolio on the mean-variance efficient frontier, which was first characterized by Markowitz (1952). To implement these portfolios in practice, one has to estimate the mean and the covariance matrix of asset returns. Traditionally, the *sample* mean and covariance matrix have been used for this purpose. However, because of estimation error, policies constructed using these estimators are extremely unstable; that is, the resulting portfolio weights fluctuate substantially over time. This has greatly undermined the popularity of mean-variance portfolios among portfolio managers, who are reluctant to implement policies that recommend such drastic changes in the portfolio composition. Moreover, the concerns of portfolio managers are reinforced by well-known empirical evidence, which shows that these unstable portfolios perform very poorly in terms of their *out-of-sample* mean and variance; see Michaud (1989), Chopra and Ziemba (1993), and Broadie (1993).

The instability of the mean-variance portfolios can be explained (partly) by the well-documented difficulties associated with estimating *mean* asset returns; see Merton (1980). For this reason, researchers have recently focused on the minimum-variance portfolio, which relies solely on estimates of the covariance matrix, and thus is not as sensitive to estimation error (Chan et al. 1999, Jagannathan and Ma 2003). Jagannathan and Ma, for example, state that "the estimation error in the sample mean is so large that nothing much is lost in ignoring the mean altogether" (p. 1652). This claim is substantiated by extensive empirical evidence that shows the minimum-variance portfolio usually performs better out of sample than any other mean-variance portfolio—even when Sharpe ratio or other performance measures related to *both* the mean and variance are used for the comparison; see Jorion (1986), Jagannathan and Ma (2003), and DeMiguel et al. (2005). Moreover, in this paper we provide numerical results that also illustrate the perils associated with using estimates of mean returns for portfolio selection. For all these reasons, herein our discussion focuses on the minimum-variance portfolios.

Although the minimum-variance portfolio does not rely on estimates of mean returns, it is still quite vulnerable to the impact of estimation error; see Chan et al. (1999) and Jagannathan and Ma (2003). The sensitivity of the minimum-variance portfolio to estimation error is surprising. These portfolios are based on the sample covariance matrix, which is the maximum likelihood estimator (MLE) for normally distributed returns. Moreover, MLEs are theoretically the most efficient for the assumed distribution; that is, these estimators have the smallest asymptotic variance provided the data follows the assumed distribution. So why does the sample covariance matrix give unstable portfolios? The answer is the efficiency of MLEs based on assuming normality of returns is highly sensitive to deviations of the asset-return distribution from the assumed (normal) distribution. In particular, MLEs based on the normality assumption are not necessarily the most efficient for data that depart even slightly from normality; see Example 1.1 in Huber (2004). This is particularly important for portfolio selection, where extensive evidence shows that the empirical distribution of returns usually deviates from the normal distribution.

To induce greater stability on the minimum-variance portfolio weights, in this paper we propose a class of policies that are constructed using *robust estimators* of the portfolio return characteristics. A robust estimator is one that gives meaningful information about asset returns even when the empirical (sample) distribution deviates from the assumed (normal) distribution (see Huber 2004, Hampel et al. 1986, Rousseeuw and Leroy 1987). Specifically, a robust estimator should have good properties not only for the assumed distribution, but also for any distribution in a neighborhood of the assumed one.

Classical examples of robust estimators are the median and the mean absolute deviation (MAD). The median is the value that is larger than 50% and smaller than 50% of the sample data points whereas the MAD is the mean absolute deviation from the median. The following example from Tukey (1960) illustrates the advantages of using robust estimators. Assume that all but a small fraction h of the data are drawn from a univariate normal distribution, whereas the remainder are drawn from the same normal distribution, but with a standard deviation three times larger. Then, a value of $h = 10\%$ is enough to make the median as efficient as the mean, whereas more sophisticated robust estimators are 40% more efficient than the mean with $h = 10\%$. Moreover, even $h = 0.1\%$ is enough to make the MAD more efficient than the standard deviation. The conclusion is that when the sample distribution deviates even slightly from the assumed distribution, the efficiency of classical estimators may be drastically reduced. *Robust estimators*, on the other hand, are not as efficient as MLEs when the underlying model is correct, but their properties are not as sensitive to deviations from the assumed distribution.

For this reason, we examine portfolio policies based on *robust* estimators. These policies should be less sensitive to deviations of the empirical distribution of returns from normality than the traditional policies. We focus on certain robust estimators known as M- and S-estimators, which have better properties than the classical median and MAD.

Our paper makes three contributions. Our first contribution is to show how one can compute the portfolio policy that minimizes a *robust* estimator of risk by solving a *single* nonlinear program. As mentioned above, we focus on minimum-risk portfolios because they usually perform better out of sample than portfolios that optimize the trade-off between in-sample risk and return. The proposed portfolios are the solution to a nonlinear program where portfolio optimization *and* robust estimation are performed in a single step. In particular, the decision variables of this optimization problem are the portfolio weights, and its objective is either the M- or S-estimator of portfolio risk.

Our second contribution is to characterize (analytically) the properties of the resulting portfolios. Specifically, we give an analytical bound on the sensitivity of the portfolio weights to changes in the distribution of asset returns. Our analysis shows that the portfolio weights of the proposed policies are less sensitive to changes in the distributional assumptions than those of the traditional minimum-variance policies. As a result, the portfolio weights of the proposed policies are more stable than those of the traditional policies. This makes the proposed portfolios a credible alternative to the traditional policies in the eyes of the investors, who are usually reticent to implement portfolios whose recommended weights fluctuate substantially over time.

Our third contribution is to compare the behavior of the proposed portfolios to that of the traditional portfolios on simulated and empirical data. The results confirm that minimum-risk portfolios (standard and robust) attain higher out-of-sample Sharpe ratios than return-risk portfolios (standard and robust). As mentioned above, this is because estimates of mean returns (standard and robust) contain so much estimation error that using them for portfolio selection worsens performance. Comparing the proposed minimum-risk portfolios to the traditional minimum-variance portfolios, we observe that the proposed portfolios have more stable weights than the traditional portfolios, while preserving (or slightly improving) their high out-of-sample Sharpe ratios.

Other researchers have proposed portfolio policies based on robust estimation techniques; see Cavadini et al. (2001), Vaz-de Melo and Camara (2003), Perret-Gentil and Victoria-Feser (2004), and Welsch and Zhou (2007). Their approaches, however, differ from ours. All three papers compute the robust portfolio policies in *two* steps. First, they compute a robust estimate of the covariance matrix of asset returns. Second, they solve the minimum-variance problem where the covariance matrix is replaced by its robust estimate. We, on the other hand, propose solving a single nonlinear program, where portfolio optimization and robust estimation are performed in *one step*.

The only other one-step approach to robust portfolio estimation is in Lauprete et al. (2002); see also Lauprete (2001). They consider a one-step robust approach based on the M-estimator of risk and give some numerical results. We, in addition, consider portfolios based on the S-estimators, give an analytical bound on the sensitivity of the M- and S-portfolio weights to changes in the distributional assumptions, and examine the behavior of both the M- and S-portfolios on simulated and empirical data sets.

Finally, a number of other approaches have been proposed in the literature to address estimation error. The robust portfolio optimization approach (see, for example, Goldfarb and Iyengar 2003, Tütüncü and Koenig 2004, Garlappi et al. 2007, Lu 2006) explicitly recognizes that the result of the estimation process is not a single-point estimate, but rather an uncertainty set, where the true mean and covariance matrix of asset returns lie with certain confidence. A robust portfolio is, then, one that optimizes the worst-case performance with respect to all possible values the mean and covariance matrix may take within their corresponding uncertainty sets. Bayesian portfolio policies are constructed using estimators that are generated by combining the investor's prior beliefs with the evidence obtained from historical return data; see Jorion (1986), Black and Litterman (1992), and Pástor and Stambaugh (2000). Finally, Jagannathan and Ma (2003) show that imposing short-selling constraints can help to reduce the impact of estimation error on the stability and performance of the minimum-variance portfolio.

The rest of this paper is organized as follows. Section 2 reviews the mean-variance and minimum-variance portfolios and highlights their lack of stability with a simple example. In §3, we show how to compute the M- and S-portfolios. In §4, we analyze the sensitivity of the proposed portfolio policies to changes in the empirical distribution of asset returns. In §5, we compare the different policies on simulated and empirical data. Section 6 concludes.

2. On the Instability of the Traditional Portfolios

In this section, we use a simple example to illustrate the instability of the portfolio weights of the mean-variance and minimum-variance policies. In particular, we consider two risky assets whose returns follow a normal distribution most of the time, but there is a small probability that the returns of the two risky assets follow a different *deviation* distribution. That is, we assume that the true asset-return distribution is

$$G = 99\% \times N(\mu, \Sigma) + 1\% \times D, \tag{1}$$

where $N(\mu, \Sigma)$ is a *normal* distribution with mean μ and covariance matrix Σ, and D is a *deviation* distribution. Specifically, we are going to consider the case where there is a 99% probability that the returns of the two assets are independently and identically distributed following a normal distribution with an annual mean of 12% and an annual standard deviation of 16%, and there is a 1% probability that the returns of the two assets are distributed according to a normal distribution with the same covariance matrix but with the mean return for the second asset equal to -50 times the mean return of the first asset. That is, we assume that $h = 1\%$,

$$\mu = \begin{pmatrix} 0.01 \\ 0.01 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 0.0021 & 0 \\ 0 & 0.0021 \end{pmatrix},$$

and $D = N(\mu_d, \Sigma_d)$, where $\Sigma_d = \Sigma$ and $\mu_d = (0.01, -0.50)$.

Finally, we would like to note that a basic assumption of our work is that the investor does know that the true asset-return distribution deviates from the normal but the investor does *not* know the parametric form of this deviation. If the investor knew the parametric form of the deviation distribution D, then the investor would be better off by estimating this distribution using, for example, maximum likelihood estimation. It is convenient for exposition purposes, however, to assume that the deviation distribution D does have a parametric (normal) form in our example.

2.1. A Rolling-Horizon Simulation

We then perform a "rolling-horizon simulation." We first generate a time series of 240 asset returns by sampling from the true asset-return distribution G. Then, we carry out a rolling-horizon experiment based on this time series. Concretely, we use the first 120 returns in the time series to estimate the sample mean and covariance matrix of asset returns. We then compute the corresponding minimum-variance portfolio as well as the mean-variance portfolio for a risk aversion parameter $\gamma = 1$. We then repeat this procedure by "rolling" the estimation window forward one period at a time until we reach the end of the time series. Thus, after performing this experiment we have computed the portfolio policies corresponding to 120 different estimation windows of 120 returns each.

2.2. Computing the Mean-Variance and Minimum-Variance Portfolios

Given N risky assets, the mean-variance portfolio is the solution to the optimization problem

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \hat{\Sigma} \mathrm{w} - \frac{1}{\gamma} \hat{\mu}^\top \mathrm{w} \tag{2}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{3}$$

where $\mathrm{w} \in \mathscr{R}^N$ is the vector of portfolio weights, $\hat{\mu}^\top \mathrm{w}$ is the sample mean of portfolio returns, $\mathrm{w}^\top \hat{\Sigma} \mathrm{w}$ is the sample variance of portfolio returns, and γ is the risk-aversion parameter. The constraint $\mathrm{w}^\top e = 1$, where $e \in \mathscr{R}^N$ is the vector of ones, ensures that the portfolio weights sum to one. The sample covariance matrix of asset returns, $\hat{\Sigma}$, can be calculated as $\hat{\Sigma} = (1/(T-1)) \sum_{t=1}^{T} (r_t - \hat{\mu})(r_t - \hat{\mu})^\top$, where $r_t \in \mathscr{R}^N$ is the vector of asset returns at time t, T is the sample size, and $\hat{\mu} \in \mathscr{R}^N$ is the sample mean of asset returns, $\hat{\mu} = (1/T) \sum_{t=1}^{T} r_t$. Note that for different values of the risk aversion parameter γ, we obtain the different mean-variance portfolios on the efficient frontier. The minimum-variance portfolio is the mean-variance portfolio corresponding to an infinite risk aversion parameter ($\gamma = \infty$), and thus it can be computed by solving the following minimum-variance problem:

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \hat{\Sigma} \mathrm{w} \tag{4}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1. \tag{5}$$

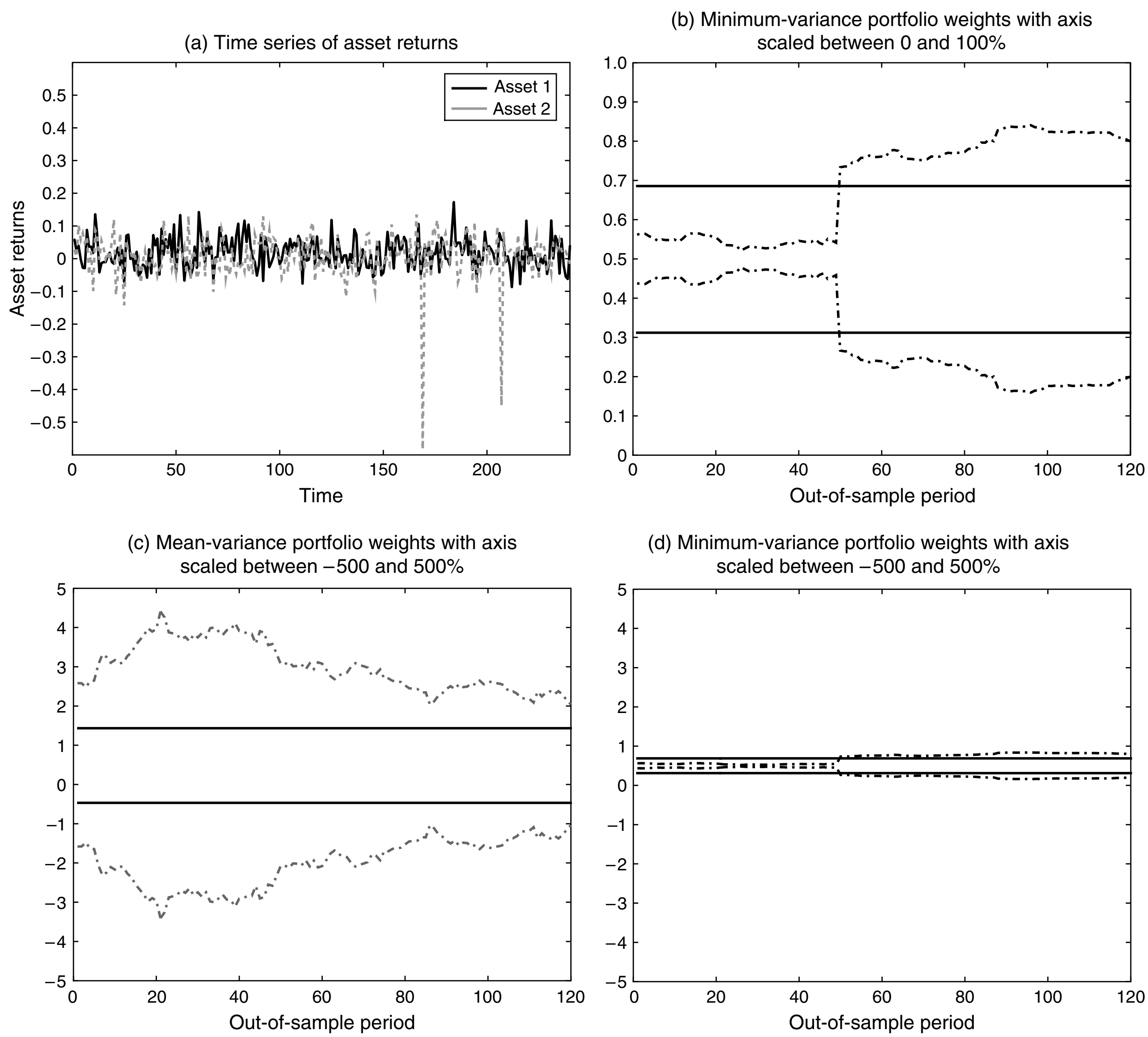
Note that the true asset return distribution G in our example is a mixture of normals, which is not normal in general. However, it is easy to compute the first and second moments of G from the first and second moments of the two normal distributions being mixed. Specifically, it is easy to see that $\mu_G = E(G) = (1-h)\mu + h\mu_d$ and $\Sigma_G = \mathrm{Var}(G) = (1-h)(\Sigma + (\mu-\mu_G)(\mu-\mu_G)^T) + h(\Sigma_d + (\mu_d-\mu_G)(\mu_d-\mu_G)^T)$. Hence, the *true* mean-variance and minimum-variance portfolios can be computed for our simple example by solving problems (2)–(3) and (4)–(5), respectively, replacing the sample mean and covariance matrix by μ_G and Σ_G.

2.3. Discussion

The time series of asset returns and the mean-variance and minimum-variance portfolio weights are depicted in Figure 1. Panel (a) depicts the times series of 240 returns for the two assets. Note that the two sample returns corresponding to dates 169 and 207 follow the deviation distribution D, whereas the rest of the returns follow the distribution $N(\mu, \Sigma)$.

Panel (c) in Figure 1 depicts the estimated mean-variance portfolio weights together with the true mean-variance portfolio weights, which are equal to 143% for the first asset and −43% for the second asset. Note that the estimated portfolio weights for the first asset range between 200% and 450% and the estimated weights for the second asset range between −325% and −100%. Clearly, the estimated mean-variance portfolio weights take extreme values that fluctuate substantially over time and tend to be very different from the true mean-variance portfolios.

Figure 1. Time series of asset returns and portfolio weights for the example in §2.



Notes. Panel (a) depicts the time series of asset returns for the two-asset example in §2. Panels (b)–(d) depict the 120 estimated mean-variance and minimum-variance portfolios. The panels also depict the "true" mean-variance and minimum-variance portfolio weights (corresponding to the true asset-return distribution). Panel (b) depicts the minimum-variance portfolio weights on an axis ranging between 0 and 100%, Panel (c) depicts the mean-variance portfolio weights on an axis ranging between −500% and 500%, and for comparison purposes, Panel (d) depicts the minimum-variance portfolio on an axis ranging between −500% and 500%.

Panel (d) shows the estimated minimum-variance portfolio weights. Comparing panels (c) and (d), it seems clear that the estimated mean-variance portfolio weights are more unstable than the estimated minimum-variance portfolio weights. This confirms the insight given by Merton (1980) that the error incurred when estimating mean asset returns is much larger than that incurred when estimating the covariance matrix. Specifically, Merton showed that although the estimation error in the sample covariance matrix can be reduced by increasing the frequency with which the return data is sampled, the estimation error in the sample mean can only be reduced by increasing the total duration of the time series. Consequently, for most real-world data sets, it is nearly impossible to obtain a time series long enough to generate reasonable estimates of mean asset returns. Our numerical results in §5 also confirm this point. For this reason, and following the same argument as in much of the recent literature (Chan et al. 1999, Jagannathan and Ma 2003), in this paper we focus on the minimum-variance policy.

The estimated minimum-variance portfolio weights are also depicted in panel (b) in Figure 1, but on a vertical axis that ranges only between 0% and 100%, whereas the vertical axis in panels (c) and (d) ranges between −500% and 500%. The panel also shows the true minimum-variance portfolio weights, which are equal to 69% for the first asset and 31% for the second asset. Note that the first 49 estimated minimum-variance portfolios are obtained from the first 168 return samples in the time series depicted in panel (a). None of these sample returns contains a negative jump for the second asset. Consequently, the estimated minimum-variance portfolio weights are close to 50%. The next 38 estimated portfolios are obtained from estimation windows containing exactly one negative jump. As a result, these 38 estimated minimum-variance portfolios assign a larger weight to the first asset. Comparing these 38 estimated portfolios to the true minimum-variance portfolio, we note, however, that these 38 portfolios *overestimate* the weight that should be assigned to the first asset. Finally, the rest of the estimated portfolios are obtained from estimation windows that contain exactly two negative jumps for the second asset. As a result, the corresponding estimated minimum-variance portfolios overestimate even more the weight that should be assigned to the first asset. Summarizing, the minimum-variance portfolios tend to underestimate the weight on the first asset when there are no jumps in the estimation window and they tend to overestimate the weight on the first asset when there are one or two jumps in the estimation window.

Hence, the example also shows that although the minimum-variance portfolio weights are more stable than those of the mean-variance portfolio, they are still quite unstable over time. This can be explained as follows. The minimum-variance portfolio is based on the sample covariance matrix, which is the MLE for normally distributed returns, and thus should be the most efficient estimator. However, although MLEs are very efficient for the assumed (normal) distribution, they are highly sensitive to deviations in the sample or empirical distribution from normality. Consequently, the minimum-variance portfolio is bound to be very sensitive to the two sample returns following the deviation distribution D. To understand this better, note that the sample variance of portfolio returns is

$$\mathrm{w}^\top \hat{\Sigma} \mathrm{w} = \frac{1}{T} \sum_{t=1}^{T} (\mathrm{w}^\top (r_t - \hat{\mu}))^2. \tag{6}$$

Although MLEs are very efficient for data that follow a normal distribution, the fast growth rate of the square function in (6) makes the sample variance (and thus the minimum-variance portfolio) highly sensitive to deviations in the empirical distribution from normality, such as jumps or heavy tails. This is particularly worrying in finance, where there is extensive evidence that the empirical return distributions often depart from normality.[1] In the next section we propose a class of portfolios that minimize robust estimates of risk. These robust estimates of risk are based on functions that grow more slowly than the square function.

3. Robust Portfolio Estimation

In this section, we propose two classes of portfolio policies that are based on the robust M- and S-estimators, and we show how these policies can be computed by solving a nonlinear program where portfolio optimization and robust estimation are performed in one step.

3.1. M-Portfolios

For a given portfolio w, the M-estimator of portfolio risk s is

$$s = \frac{1}{T} \sum_{t=1}^{T} \rho(\mathrm{w}^\top r_t - m), \tag{7}$$

where the *loss function* ρ is a convex symmetric function with a unique minimum at zero, and m is the M-estimator of portfolio return:

$$m = \arg\min_{m} \frac{1}{T} \sum_{t=1}^{T} \rho(\mathrm{w}^\top r_t - m).$$

Particular cases of M-estimators are the sample mean and variance, which are obtained for $\rho(r) = 0.5r^2$, and the median and MAD, for $\rho(r) = |r|$. In our numerical experiments we focus on the M-estimators derived from Huber's loss function

$$\rho(r) = \begin{cases} r^2/2, & |r| \leqslant c, \\ c(|r| - c/2), & |r| > c, \end{cases} \tag{8}$$

where c is a constant. Note that for large values of $|r|$, all of these loss functions lie below the square function. This makes the M-estimators more robust with respect to deviations from normality of the empirical distribution than the traditional mean and variance.

We define the M-portfolio as the policy that minimizes the M-estimator of portfolio risk. The M-portfolio can then be computed as the solution to the following optimization problem:

$$\min_{\mathrm{w},\, m} \quad \frac{1}{T}\sum_{t=1}^{T} \rho(\mathrm{w}^\top r_t - m) \tag{9}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1. \tag{10}$$

Note that for fixed w, the minimum with respect to m of the objective function of problem (9)–(10) is equal to the M-estimator of risk s for the return of the portfolio w, as defined in (7). By including the portfolio weight vector w as a variable for the optimization problem, we compute the portfolio that minimizes the M-estimator of risk.

The M-portfolios generalize several well-known portfolio policies. For example, the minimum-variance portfolio is the M-portfolio corresponding to the square or L_2 loss function, $\rho(r) = 0.5r^2$. Also, the portfolio that minimizes the mean absolute deviation from the median (MAD) is the M-portfolio corresponding to the L_1 loss function $\rho(r) = |r|$. In our numerical experiments, we use the portfolios obtained from Huber's loss function because of their good out-of-sample performance.

3.2. S-Portfolios

The second class of portfolio policies we propose is based on the robust S-estimators. The main advantage of S-estimators is that they are equivariant with respect to scale; that is, multiplying the whole data set by a constant does not change the value of the S-estimator. This is not the case for the M-estimators. The S-estimators of portfolio return and risk are defined as the values of m and s that solve the following optimization problem:

$$\min_{m,\, s} \quad s \tag{11}$$

$$\text{s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} \rho\left(\frac{\mathrm{w}^\top r_t - m}{s}\right) = K, \tag{12}$$

where ρ is the loss function and K is the expectation of this loss function evaluated at a standard normal random variable z; that is, $K = E(\rho(z))$. Note that the portfolio return deviations, $\mathrm{w}^\top r_t - m$, are scaled by the S-estimator for risk s in Equation (12). Intuitively, this is what makes the S-estimators scale invariant.

The loss function ρ in (12) must satisfy two conditions: (i) it must be symmetric with a unique minimum at zero, and (ii) there must exist $c > 0$ such that ρ is strictly increasing on $[0, c]$ and constant on $[c, \infty)$. A crucial implication of these two conditions is that the loss function for S-estimators is bounded above. Consequently, the contribution of any sample return to the S-estimator of portfolio risk is bounded. In our experiments, we use Tukey's biweight function:

$$\rho(r) = \begin{cases} \dfrac{c^2}{6}(1 - (1 - (r/c)^2)^3), & |r| \leqslant c, \\ \dfrac{c^2}{6}, & \text{otherwise.} \end{cases} \tag{13}$$

S-estimators allow the flexibility to choose the breakdown point, which is the amount of data deviating from the reference model that an estimator can accept while giving meaningful information. For example, when using Tukey's biweight loss function, we can control the breakdown point by choosing the constant c. The S-estimators allow a breakdown point of up to 50%.

We define the S-portfolio as the policy that minimizes the S-estimate of risk; namely, the portfolio that solves the following optimization problem:

$$\min_{\mathrm{w},\, m,\, s} \quad s \tag{14}$$

$$\text{s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} \rho\left(\frac{\mathrm{w}^\top r_t - m}{s}\right) = K, \tag{15}$$

$$\mathrm{w}^\top e = 1. \tag{16}$$

3.3. Two-Step Approaches

Perret-Gentil and Victoria-Feser (2004), Vaz-de Melo and Camara (2003), Cavadini et al. (2001), and Welsch and Zhou (2007) propose a different procedure for computing portfolios based on robust statistics. Basically, they propose a two-step approach to robust portfolio estimation. First, they compute a robust estimate of the covariance matrix of asset returns. Second, they compute the portfolio policies by solving the classical minimum-variance problem (4)–(5), but replacing the sample mean and covariance matrix by their robust counterparts. The main difference among these three approaches is the type of robust estimator used. Perret-Gentil and Victoria-Feser (2004) use S-estimators, Vaz-de Melo and Camara (2003) use M-estimators, Cavadini et al. (2001) use the equivariant location and scale M-estimators (Maronna 1976 and Huber 1977), and Welsch and Zhou (2007) use the minimum covariance determinant estimator and winsorization.

Our approach differs from all of these two-step approaches because we propose solving a nonlinear program where robust estimation and portfolio optimization are performed in one step. Thus, our approach does not require the explicit computation of any estimate of the covariance matrix.

Just like we do, Cavadini et al. (2001) and Perret-Gentil and Victoria-Feser (2004) also derive analytical bounds on the sensitivity of their proposed portfolio weights to changes in the distributional assumptions. However, their method of analysis differs from ours. We derive these analytical bounds by studying how the solution to the optimality conditions of the M- or S-portfolio problem changes with changes in the asset-return distribution. Cavadini et al. (2001) and Perret-Gentil and Victoria-Feser (2004) show that the analytical bounds for their portfolios follow automatically from the influence function of the robust estimators they use for the covariance matrix of asset returns.

Finally, the approaches in Perret-Gentil and Victoria-Feser (2004), Vaz-de Melo and Camara (2003), Cavadini et al. (2001), and Welsch and Zhou (2007) can also be used to compute robust *mean*-variance portfolios. This can be done by simply replacing the sample mean and variance by their robust estimates in the classical mean-variance portfolio problem. Our out-of-sample evaluation results in §5, however, show that the resulting robust mean-variance portfolios are substantially outperformed (in terms of out-of-sample Sharpe ratio) by the robust minimum-variance portfolios. As argued before, the reason for this is that estimates of mean returns (both standard and robust) contain so much estimation error that using them for portfolio selection is likely to hurt the performance of the resulting portfolios. Also, the out-of-sample evaluation results show that the stability and performance of the two-step robust minimum-variance portfolios proposed in Perret-Gentil and Victoria-Feser (2004) are not as good as those of our proposed robust minimum M- and S-risk portfolios, but they are better than those of the traditional minimum-variance policy.

3.4. The Example Revisited

We have tried the minimum M-risk and S-risk portfolios on the time series of asset returns from the example in §2. The resulting portfolio weights are depicted in Figure 3 in the online appendix, which is available as part of the online version that can be found at http://or.pubs.informs.org/. The figure demonstrates that the weights of the robust portfolios are more stable than those of the traditional portfolios, and overall stay relatively close to the true minimum-variance portfolio weights. The reason for this is that by reducing the impact of the negative jumps on the estimated S-portfolios, the proposed policies manage to preserve the stability of the portfolio weights. In §5, we give numerical results on simulated and empirical data sets that confirm the insights from this simple example.

4. Analysis of Portfolio Weight Stability

In this section, we characterize (analytically) the sensitivity of the M- and S-portfolio weights to changes in the distribution of asset returns. To do so, we derive the *influence function* (IF) of the portfolio weights, which gives a first-order approximation to portfolio weight sensitivity. We also show that the IF of the proposed M and S-portfolios is smaller than that of the traditional minimum-variance policy. Specifically, we show that the sensitivity of the M-portfolio weights to a particular sample return grows *linearly* with the distance between the sample return and the location estimator of return, whereas the sensitivity of the S-portfolios to a particular sample return is bounded, and the sensitivity of the minimum-variance portfolios grows with the square of the distance between the sample return and the sample mean return.

The IF (see Hampel et al. 1986) measures the impact of small changes in the distributional assumptions on the value of an estimator θ. In our case, this estimator θ contains the vector of portfolio weights w, the robust estimators m and s, and the Lagrange multipliers of the constraints in problems (9)–(10) and (14)–(16). Given a cumulative distribution function (CDF) of returns $F(R)$, the IF measures the impact of a small perturbation $\hat{r}$ to this CDF on the value of the estimator θ. The formal definition of IF is the following:

$$\mathrm{IF}_\theta(\hat{r}, F) = \lim_{h\to 0} \frac{\theta((1-h)F + h\Delta_{\hat{r}}) - \theta(F)}{h}, \tag{17}$$

where $\theta(F)$ is the estimator corresponding to the cumulative distribution function F, and $\Delta_{\hat{r}}$ is a CDF for which $\hat{r}$ occurs with 100% probability; that is,

$$\Delta_{\hat{r}}(R) = \begin{cases} 0, & R < \hat{r}, \\ 1, & R \geqslant \hat{r}. \end{cases} \tag{18}$$

Thus, the IF measures the per-unit (standardized) effect of a sample return $\hat{r}$ on the value of an estimator. Mathematically, the IF may be interpreted as the directional derivative of the estimator θ, evaluated at the distribution function F, in the direction $\Delta_{\hat{r}}$. Finally, the IF function can be used to derive several statistical properties of an estimator such as the asymptotic variance and the gross-error sensitivity; see §1.3 of Hampel et al. (1986).

The IF of the portfolio weights is particularly informative in the context of portfolio selection. First, it is clear that if the IF of the portfolio weights of a given policy is relatively small or remains bounded for all possible values of $\hat{r}$, then this portfolio policy is relatively insensitive to changes in the distributional assumptions. Second, we can use the IF to give a first-order bound on the sensitivity of the portfolio weights to the introduction of an additional sample return in the estimation window. Concretely, assume that the empirical distribution of the historical data available at time T is given by F_T and that we then obtain a new sample return at time $T+1$, $\hat{r}$. Then, by Taylor's theorem, we know that the difference between the portfolio weights computed before and after $T+1$ is bounded as follows:

$$\begin{aligned} &\mathrm{w}\left(\frac{T}{T+1}F_T + \frac{1}{T+1}\Delta_{\hat{r}}\right) - \mathrm{w}(F_T) \\ &\quad \leqslant \frac{1}{T+1}\mathrm{IF}_\mathrm{w}(\hat{r}, F_T) + O(T^{-2}), \end{aligned} \tag{19}$$

where IF_w is the IF of the portfolio weights and $O(T^{-2})$ denotes the second-order (small) terms. The main implication of this bound is that if the IF of the portfolio weights corresponding to a particular policy is bounded (or relatively small) for all values of $\hat{r}$, then the effect of including a new sample point in the data is also bounded (or small), up to first-order terms.

4.1. M-Portfolio Influence Function

To compute the M-portfolio IFs, we study how the solution to the optimality conditions of the M-portfolio problem (9)–(10) is affected by changes in the distribution of asset returns. We denote the IF of the robust estimator m, the M-portfolio weights w, and the Lagrange multipliers of the M-portfolio problem λ as IF_m, IF_{w}, and IF_λ, respectively. Moreover, we formally define these IFs as $\mathrm{IF}_x \equiv \mathrm{IF}_x(\hat{r}, F) = (\partial/\partial h)x((1-h)F + h\Delta_{\hat{r}})|_{h=0}$ for $x = m, x, \lambda$. The first-order optimality conditions for the M-portfolio problem (9)–(10) are

$$-\frac{1}{T}\sum_{t=1}^{T}\psi(\mathrm{w}^\top r_t - m) = 0, \tag{20}$$

$$\frac{1}{T}\sum_{t=1}^{T}\psi(\mathrm{w}^\top r_t - m)r_t - \lambda e = 0, \tag{21}$$

$$\mathrm{w}^\top e - 1 = 0, \tag{22}$$

where $\psi(r) = \rho'(r)$ and λ is the Lagrange multiplier corresponding to the equality constraint $\mathrm{w}^\top e = 1$. The functional form of these first-order optimality conditions is the following:

$$\int \psi(\mathrm{w}^\top R - m)\, dF(R) = 0, \tag{23}$$

$$\int \psi(\mathrm{w}^\top R - m)R\, dF(R) - \lambda e = 0, \tag{24}$$

$$\mathrm{w}^\top e - 1 = 0, \tag{25}$$

where $F(R)$ is the CDF of asset returns.[2]

The following theorem gives a linear system whose solution gives the IFs of the M-portfolios. The proof to the theorem is given in the online appendix.

THEOREM 1. *Let (m, w, λ) be an M-estimate satisfying* (23)–(25) *and let the function $\psi(R)$ be measurable and continuously differentiable. Then, the influence functions of the M-portfolio are the solution to the following symmetric linear system*:

$$\begin{pmatrix} E(\psi'(w^\top R - m)) & -E(\psi'(w^\top R - m)R^\top) & 0 \\ -E(\psi'(w^\top R - m)R) & E(\psi'(w^\top R - m)RR^\top) & e \\ 0 & e^\top & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathrm{IF}_m \\ \mathrm{IF}_w \\ \mathrm{IF}_\lambda \end{pmatrix} = \begin{pmatrix} \psi(w^\top \hat{r} - m) \\ \lambda e - \psi(w^\top \hat{r} - m)\hat{r} \\ 0 \end{pmatrix}. \tag{26}$$

The following proposition gives an analytic expression for IF_{w}, the influence function for the M-portfolio weights. We use the following notation: $Z \equiv \mathrm{w}^\top R - m$, $\psi_Z \equiv \psi(\mathrm{w}^\top R - m)$, $\psi'_Z \equiv \psi'(\mathrm{w}^\top R - m)$, and $\psi_{\hat{z}} \equiv \psi(\mathrm{w}^\top \hat{r} - m)$. The proof of the proposition is given in the online appendix.

PROPOSITION 1. *If the following conditions hold*:
1. $E(\psi'_Z) \neq 0$,
2. *the return distribution $F(R)$ has finite first and second moments,*
3. *the following matrix is invertible*:

$$H = E(\psi'_Z RR^\top) - \frac{E(\psi'_Z R)E(\psi'_Z R^\top)}{E(\psi'_Z)}, \quad \textit{and}$$

4. $e^\top H^{-1} e \neq 0$,

then the matrix in (26) *is invertible and the M-portfolio weights influence function is*

$$\mathrm{IF}_w = \psi_{\hat{z}} H^{-1}\left(\frac{E(\psi'_Z R)}{E(\psi'_Z)} - \hat{r}\right). \tag{27}$$

REMARK 1. Conditions (1)–(4) in Proposition 1 are mild. To see this, note that for the square loss function the M-portfolio coincides with the minimum-variance portfolio and conditions (1)–(4) are necessary for the minimum-variance portfolio to be well defined. Concretely, for the square loss function $\rho(r) = 0.5r^2$, we have $\psi(r) = r$ and $\psi'(r) = 1$. Thus, $H = \Sigma$ if Condition (2) holds. Moreover, from the optimality Conditions (23)–(25), we have that $m = \mathrm{w}^\top \mu$ and $\mathrm{w} = (1/e^\top \Sigma^{-1} e)\Sigma^{-1} e$; that is, the M-portfolio coincides with the minimum-variance portfolio. Finally, clearly the minimum-variance portfolio is well defined only if Conditions (3) and (4) hold.

REMARK 2. The main implication of Equation (27) is that

$$\|\mathrm{IF}_{\mathrm{w}}\| \leqslant |\psi_{\hat{z}}| \times \|H^{-1}\| \times \left\|\frac{E(\psi'_Z R)}{E(\psi'_Z)} - \hat{r}\right\|. \tag{28}$$

We are particularly interested in comparing the IFs of the minimum-variance and M-portfolio weights. Note that the IF of the minimum-variance portfolio weights can be obtained from (28) by setting $\rho(r) = 0.5r^2$ or $\psi(r) = r$. Simple algebra yields the expression

$$\|\mathrm{IF}_{\mathrm{w}_{\mathrm{MV}}}\| \leqslant |\mathrm{w}_{\mathrm{MV}}^\top \hat{r} - \mu| \times \|\Sigma^{-1}\| \times \|\mu - \hat{r}\|, \tag{29}$$

where w_{MV} is the minimum-variance portfolio, μ is the vector of mean asset returns, and Σ is the covariance matrix. When comparing expressions (28) and (29), we note that the second and third factors on the right-hand side of (28) and (29) are roughly comparable in size for all loss functions mentioned in §3.1, including the squared or L_2 loss function $\rho(r) = 0.5r^2$. The main difference is that while the first factor in (29) (that is, $|\mathrm{w}_{\mathrm{MV}}^\top \hat{r} - \mu|$) is not bounded for all $\hat{r}$, the first factor in (28) (that is, $|\psi_{\hat{z}}|$) is bounded for the absolute value and Huber functions.

4.2. S-Portfolio Influence Function

To derive the influence function of the S-portfolios we follow a procedure similar to that we used to derive the M-portfolio IFs. In particular, we first state the optimality conditions of the S-portfolio problem (14)–(16), and then we analyze how the solution to this optimality conditions is affected by changes in the return distribution.

We denote the IFs of the robust estimators m and s, the S-portfolio weights w, and the Lagrange multipliers ν and λ as $\mathrm{IF}_x \equiv \mathrm{IF}_x(\hat{r}, F) = (\partial/\partial h)x((1-h)F + h\Delta_{\hat{r}})|_{h=0}$, where $x = \{m, s, \mathrm{w}, \nu, \lambda\}$. The functional form of the first-order optimality conditions for the S-portfolio problem (14)–(16) is

$$\int \frac{\nu}{s}\psi\left(\frac{\mathrm{w}^\top R - m}{s}\right) dF(R) = 0, \tag{30}$$

$$1 + \int \frac{\nu}{s}\psi\left(\frac{\mathrm{w}^\top R - m}{s}\right)\left(\frac{\mathrm{w}^\top R - m}{s}\right) dF(R) = 0, \tag{31}$$

$$-\int \frac{\nu}{s}\psi\left(\frac{\mathrm{w}^\top R - m}{s}\right) R\, dF(R) - \lambda e = 0, \tag{32}$$

$$\int \rho\left(\frac{\mathrm{w}^\top R - m}{s}\right) dF(R) - K = 0, \tag{33}$$

$$\mathrm{w}^\top e - 1 = 0, \tag{34}$$

where $\psi(r) = \rho'(r)$, ν is the Lagrange multiplier corresponding to the equality constraint (15), λ is the Lagrange multiplier for the equality constraint (16), and K is as defined in §3.2.

The following theorem gives the linear system whose solution gives the S-portfolio IFs. The proof to the theorem is given in the online appendix. We use the following notation: $Z \equiv (\mathrm{w}^\top R - m)/s$, $\hat{z} \equiv (\mathrm{w}^\top \hat{r} - m)/s$, $\psi_Z \equiv \psi((\mathrm{w}^\top R - m)/s)$, $\psi'_Z \equiv \psi'((\mathrm{w}^\top R - m)/s)$, $\psi_{\hat{z}} \equiv \psi((\mathrm{w}^\top \hat{r} - m)/s)$, and $\rho_{\hat{z}} \equiv \rho((\mathrm{w}^\top \hat{r} - m)/s)$.

Theorem 2. *Let (m, s, w, ν, λ) be an S-estimate satisfying* (30)–(34)*, and let the functions $\rho(r)$ and $\psi(r)$ be measurable and continuously differentiable. Then, the S-estimate influence functions are the solution to the following symmetric linear system*:

$$E(M)\mathrm{IF} = b, \tag{35}$$

where

$$M = \begin{pmatrix} -\frac{\nu}{s^2}\psi'_Z & -\frac{\nu}{s^2}(\psi_Z + \psi'_Z Z) & \frac{\nu}{s^2}\psi'_Z R^\top & \frac{1}{s}\psi_Z & 0 \\ -\frac{\nu}{s^2}(\psi_Z + \psi'_Z Z) & -\frac{\nu}{s^2}(2\psi_Z Z + \psi'_Z Z^2) & \frac{\nu}{s^2}(\psi'_Z Z R^\top + \psi_Z R^\top) & \frac{1}{s}\psi_Z Z & 0 \\ \frac{\nu}{s^2}\psi'_Z R & \frac{\nu}{s^2}(\psi_Z R + \psi'_Z Z R) & -\frac{\nu}{s^2}\psi'_Z R R^\top & -\frac{1}{s}\psi_Z R & e \\ \frac{1}{s}\psi_Z & \frac{1}{s}\psi_Z Z & \frac{1}{s}\psi_Z R^\top & 0 & 0 \\ 0 & 0 & e^\top & 0 & 0 \end{pmatrix},$$

$$\mathrm{IF} = \begin{pmatrix} \mathrm{IF}_m \\ \mathrm{IF}_s \\ \mathrm{IF}_w \\ \mathrm{IF}_\nu \\ \mathrm{IF}_\lambda \end{pmatrix}, \quad \textit{and} \quad b = \begin{pmatrix} -\frac{\nu}{s}\psi_{\hat{z}} \\ -\frac{\nu}{s}\psi_{\hat{z}}\hat{z} - 1 \\ \lambda e + \frac{\nu}{s}\psi_{\hat{z}}\hat{r} \\ \rho_{\hat{z}} - K \\ 0 \end{pmatrix}.$$

The following proposition shows that the S-portfolio weights influence function $\mathrm{IF_w}$ is bounded. The proof is given in the online appendix.

Proposition 2. *If the following conditions hold*:

1. *ρ is Tukey's biweight function* (13),
2. *the return distribution $F(R)$ has finite first and second moments,*
3. *the matrix $E(M)$ is invertible,*

then, the influence function of the S-risk portfolio weights is bounded.

Remark 3. Conditions (1)–(3) in Proposition 2 are mild. In particular, it is easy to show that for the square loss function, the S-portfolio coincides with the minimum-variance portfolio, and Conditions (2)–(3) are required if the minimum-variance portfolio is to be well defined.

Remark 4. The importance of the result in Proposition 2 is that it shows that the S-portfolio weights are more stable than the minimum-variance portfolio weights. In particular, although Proposition 2 shows that the IF of the S-portfolio weights is bounded for all $\hat{r}$, it is easy to see from (29) that the IF of the minimum-variance portfolio is unbounded.

5. Out-of-Sample Evaluation

In this section, we use simulated and empirical data sets to illustrate the stability and performance properties of the proposed portfolios. Although our work focuses mainly on minimum-risk portfolios, for completeness we also evaluate the stability and performance of traditional and robust portfolios that optimize the trade-off between in-sample risk and return. We consider 12 portfolios: (1) the mean-variance (Mean-var) portfolio with risk aversion parameter $\gamma = 1$;[3] (2) the minimum-variance (Min-var) portfolio; (3) the two-step robust mean-variance portfolio (2-Mean) of Perret-Gentil and Victoria-Feser (2004) with risk aversion parameter $\gamma = 1$; (3) the two-step robust minimum-variance portfolio (2-var) of Perret-Gentil and Victoria-Feser (2004); (4) the minimum M-risk portfolio with Huber's loss function (M-Hub); and (5) the minimum S-risk portfolio with Tukey's biweight loss function (S-Tuk). The remaining six portfolios are the same but with the addition of short-selling constraints, i.e., with nonnegativity constraints on the portfolio weights.

This section is divided into three parts. In the first part, we explain our methodology for evaluating the different policies. In the second part, we give the results for the simulated data and, in the third part, for the empirical data.

5.1. Evaluation Methodology

We compare the different policies using a "rolling-horizon" procedure. First, we choose a window over which to perform the estimation. We denote the length of the estimation window by $T < L$, where L is the total number of samples in the data set. For our experiments, we use an estimation window of $T = 120$ data points, which for monthly data corresponds to 10 years.[4] Two, using the return data in the estimation window we compute the different portfolio policies. Three, we repeat this "rolling-window" procedure for the next period by including the data for the new date and dropping the data for the earliest period. We continue doing this until the end of the data set is reached. At the end of this process, we have generated $L - T$ portfolio weight vectors for each strategy; that is, w_t^k for $t = T, \ldots, L-1$ and for each strategy k.

In the remainder of this section, we explain how we use these $L - T$ portfolio weight vectors to compare the different policies in terms of their stability and out-of-sample performance.

5.1.1. Boxplots of Portfolio Weights. The boxplots of portfolio weights give a graphical representation of the stability of the different portfolio policies. As mentioned above, as a result of the application of the "rolling-horizon" methodology, we obtain $L-T$ portfolio weight vectors for each of the strategies. Each boxplot represents the variability of the portfolio weight assigned to a particular asset by a particular policy. Specifically, the plot gives a box that has lines at the 25th, 50th, and 75th percentile values of the time series $\{w_{j,t}^k\}$ for $t = T, \ldots, L-1$, where $w_{j,t}^k$ is the weight that strategy k assigns to asset j at time t. The boxplot also gives whiskers, which are lines extending from each end of the boxes to show the extent of the rest of the weights. Finally, the boxplot also depicts the extreme portfolio weights that have values beyond the whiskers. Clearly, stable policies should have relatively compact (short) boxplots.

5.1.2. Portfolio Turnover. To define portfolio turnover, let $\mathrm{w}_{j,t}^k$ denote the portfolio weight in asset j chosen at time t under strategy k, w_{j,t^+}^k the portfolio weight *before* rebalancing but at $t+1$, and $\mathrm{w}_{j,t+1}^k$ the desired portfolio weight at time $t+1$ (after rebalancing). Then, the turnover is defined as the sum of the absolute value of the rebalancing trades across the N available assets and over the $L-T-1$ trading dates, normalized by the total number of trading dates:

$$\text{turnover} = \frac{1}{L-T-1}\sum_{t=T}^{L-1}\sum_{j=1}^{N}\left(\left|\mathrm{w}_{j,t+1}^k - \mathrm{w}_{j,t^+}^k\right|\right).$$

Roughly speaking, the turnover is the average percentage of wealth traded in each period.

5.1.3. Out-of-Sample Mean, Variance, and Sharpe Ratio of Returns. Following the "rolling-horizon" methodology, for each strategy k we compute the portfolio weights w_t^k for $t = T, \ldots, L-1$. Holding the portfolio w_t^k for one period gives the following *out-of-sample* excess return at time $t+1$: $\hat{r}_{t+1}^k = \mathrm{w}_t^{k\top} r_{t+1}$, where r_{t+1} denotes the returns in excess of the benchmark (risk-free) rate. After collecting the time series of $L-T$ excess returns r_t^k, the out-of-sample mean, variance, and Sharpe ratio of excess returns are $\hat{\mu}^k = (1/(L-T))\sum_{t=T}^{L-1} \mathrm{w}_t^{k\top} r_{t+1}$, $(\hat{\sigma}^k)^2 = (1/(L-T-1))\sum_{t=T}^{L-1}(\mathrm{w}_t^{k\top} r_{t+1} - \hat{\mu}^k)^2$, and $\widehat{\mathrm{SR}}^k = \hat{\mu}_k/\hat{\sigma}_k$, respectively. We also report p-values that measure the statistical significance of the differences between the variance and Sharpe ratio of a particular strategy and those of the minimum-variance strategy (which serves as the benchmark).[5]

5.2. Simulation Results

In this section, we describe our simulation experiments and discuss the behavior of the different portfolio policies on simulated data. Section 5.3 discusses the results for the empirical data.

5.2.1. The Simulated Data Set. We use simulation to generate asset-return data following a distribution G that deviates slightly from the normal distribution. Concretely, we assume that G is a mixture of two different distributions:

$$G = (1-h)N(\mu, \Sigma) + hD,$$

where $N(\mu, \Sigma)$ is a *normal* distribution with mean μ and covariance matrix Σ, D is a *deviation* distribution, and h is the proportion of the data that follows the deviation distribution D.

We generate three different data sets with proportions of the data deviating from normality h equal to 0%, 2.5%, and 5%. This allows us to study how the different portfolios change when the asset-return distribution progressively deviates from the normal distribution. We generate monthly return data for 1,010 years ($L = 12{,}120$); we use an estimation window length of 10 years ($T = 120$), which matches our choice when analyzing the empirical data sets; and we leave the last 1,000 years ($L - T = 12{,}000$ months) for out-of-sample evaluation.

To generate the part of the data that follows the multivariate normal distribution $N(\mu, \Sigma)$, we sample from a factor model similar to that used in MacKinlay and Pastor (2000) and DeMiguel et al. (2005). Concretely, we consider a market composed of N risky assets and one risk-free asset. The N risky assets include K factors. The excess returns of the remaining $N-K$ risky assets are generated by the following model:

$$R_{a,t} = \alpha + BR_{b,t} + \epsilon_t, \tag{36}$$

where $R_{a,t}$ is the $(N-K)$ vector of excess asset returns; α is the $(N-K)$ vector of mispricing coefficients; B is the $(N-K)\times K$ matrix of factor loadings; $R_{b,t}$ is the K vector of excess returns on the factor ("benchmark") portfolios and is distributed as a multivariate normal distribution with mean μ_b and covariance matrix Ω_b, $R_{b,t} \sim N(\mu_b, \Omega_b)$, and ϵ_t is the $(N-K)$ vector of noise, $\epsilon \sim N(0, \Sigma_\epsilon)$, which is uncorrelated with the returns on the factor portfolios. We report the case where there are four risky assets ($N = 4$) and a single factor ($K = 1$). We have also tried the cases with $N = 10$, 25, and 50, but the insights are similar to those from the case with $N = 4$ and thus we do not report these cases to conserve space. We choose the factor return that has an annual average of 8% and standard deviation of 16%. The mispricing α is set to zero and the factor loadings B for each of the other three risky assets are randomly drawn from a uniform distribution between 0.5 and 1.5. Finally, the variance-covariance matrix of noise Σ_ϵ is assumed to be diagonal with each of the three elements of the diagonal drawn from a uniform distribution with support $[0.15, 0.25]$, that is, the cross-sectional average annual idiosyncratic volatility is 20%.

In our experiments, we consider several different deviation distributions D: (i) where D assigns a 100% probability to a constant asset-return vector whose return for each asset is equal to the expected return of the asset plus five times the standard deviation of the asset return; (ii) where D assigns a 100% probability to a constant vector equal to the expected asset return plus three times the standard deviation of returns; (iii) where D assigns a 50% probability to a constant vector equal to the expected asset return *plus* five times the standard deviation of returns and 50% probability to a constant vector equal to the expected asset return *minus* five times the standard deviation of returns; (iv) where D is a normal distribution $N(\hat{\mu}, \Sigma)$, where each component of $\hat{\mu}$ is equal to the corresponding component of μ plus five times the standard deviation of the asset return; and (v) where the deviation from the normal distribution occurs for each asset at different dates. The insights from the results for these alternative types of deviations are similar, and thus we only report the results for the first type of deviation.

Finally, the proportion of the data deviating from the normal distribution and the "size" of the deviation (i.e., 0%–5% and five standard deviations) are similar to those used in Perret-Gentil and Victoria-Feser (2004), where it is also argued that this proportion and size of the deviations are a good representation of the deviations present in the historical data sets they use for their analysis.[6]

5.2.2. Discussion of Portfolio Weight Stability. We now discuss the stability of the portfolio weights of the different policies on the simulated data sets with proportion of return data deviating from normality h equal to 0%, 2.5%, and 5%.[7]

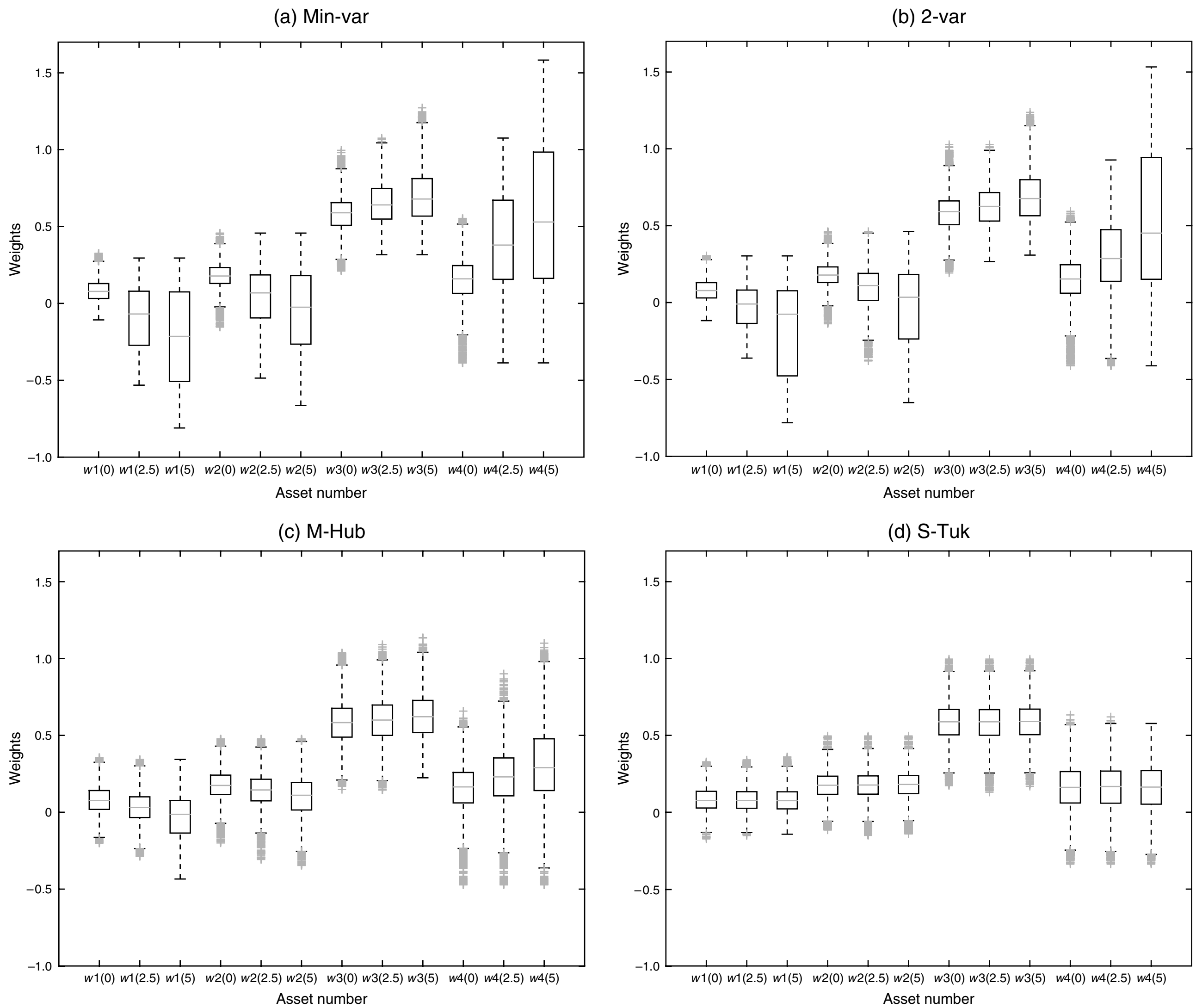
We first compare the stability of the portfolio weights of the mean-variance portfolio, the two-step robust mean-variance portfolio, and the minimum-variance portfolio on the simulated data set with 0% of the sample returns deviating from normality. We observe from our experiments that the mean-variance portfolios (traditional and robust) are much more unstable than the minimum-variance portfolio. For example, the weight assigned by the mean-variance policy to the fourth risky asset ranges between −600% and 1,200%, and the weight assigned to this same asset by the two-step robust mean-variance portfolio ranges between −700% and 1,350%, but the weight assigned to the same asset by the minimum-variance policy ranges only between −20% and 50%. This shows that mean-variance portfolios (traditional and robust) are highly unstable even for data that follow a normal distribution. Hence, the results confirm the finding in much of the recent financial literature that estimates of mean returns can be very noisy. Moreover, the experiment shows that using *robust* estimators of location for portfolio selection does not help to substantially improve the portfolio weight stability.

We now turn to discuss the stability of the policies constructed using estimates of portfolio risk only. Each of the four panels of Figure 2 gives the boxplots of the portfolio weights for each of the following policies: (a) the minimum-variance portfolio (Min-var), (b) the two-step robust minimum-variance portfolio (2-var), (c) the minimum M-risk portfolio with Huber's loss function (M-Hub), and (d) the minimum S-risk portfolio with Tukey's biweight loss function (S-Tuk). Each panel contains 12 boxplots corresponding to each of the four assets and each of the data sets with h equal to 0%, 2.5%, and 5%. Each boxplot is labelled as $wk(h)$, where $k = 1, 2, 3, 4$ is the asset number and $h = 0, 2.5, 5$ is the proportion of the data deviating from the normal distribution. The boxplots for the portfolio weights for different values of h are given side by side to facilitate the understanding of the impact of the deviation from normality on portfolio weight stability.

It is clear from Panel (a) in Figure 2 that although the minimum-variance policy is reasonably stable for normally distributed returns, it is quite unstable when even a small proportion of the data deviates from the normal distribution. To see this, note that the minimum-variance portfolio weight on the fourth asset stays in-between −40% and 50% for the data set with $h = 0\%$, but it ranges between −40% and 160% for the data set with $h = 5\%$. That is, the width of the minimum-variance portfolio weight boxplots increases from 90% to 200% when the proportion of the data deviating from normality increases from 0% to 5%. From Panel (b), we can see that the 2-var portfolio remains reasonably stable for the case where 2.5% of the sample returns deviate from the normal distribution, but it becomes quite unstable for the case with 5% deviation from normality. Panel (c) gives the boxplots of the minimum M-risk portfolio (M-Hub). Note that this portfolio remains reasonably stable for all three data sets with h equal to 0%, 2.5%, and 5%; that is, the width of the boxplots corresponding to these three cases are quite similar to each other. Finally, panel (d) gives the boxplots of the minimum S-risk portfolio (S-Tuk). The boxplots of the S-Tuk portfolio have virtually the same width for all three data sets with h equal to 0%, 2.5%, and 5%. Thus, although the stability of the M-Hub and S-Tuk portfolio weights is not altered by the presence of return data deviating from normality, the stability of the 2-var and Min-var policies is quite sensitive to these deviations.

Figure 2 shows that although the minimum-variance portfolio computed from the sample covariance matrix is

Figure 2. Boxplots of unconstrained portfolio weights for simulated data.



Notes. This figure gives the boxplots of the portfolio weights for the unconstrained policies and for the simulated data sets with 0%, 2.5%, and 5% of the data deviating from normality. Each of the four panels gives the boxplots for one of the following four unconstrained policies: minimum-variance (Min-var), 2-step robust minimum-variance (2-var), minimum M-risk portfolio with Huber's loss function (M-Hub), and minimum S-risk portfolio with Tukey's biweight function (S-Tuk). Each panel contains 12 boxplots corresponding to each of the four assets and each of the tree degrees of deviation from normality. Each boxplot is labelled as $wk(h)$, where $k = 1, 2, 3, 4$ is the asset number and $h = 0, 2.5, 5$ is the proportion of the return data deviating from normality. The boxplots for the portfolio weights for different values of h are given side by side. Finally, the box for each portfolio weight has lines at the 25, 50, and 75 percentile values of the portfolio weights. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data. Extreme portfolio weights that have values beyond the whiskers are also depicted.

an efficient estimator when the asset-return distribution follows a normal distribution, it becomes a relatively inefficient estimator when even a small proportion of the sample returns deviate from the normal distribution. To see this, note that the width of the minimum-variance portfolio weight boxplots increases substantially when the proportion of the data deviating from normality increases. The traditional minimum-variance portfolios, however, are *unbiased* estimators of the "true" minimum-variance portfolios even when the asset-return distribution deviates from normality. To verify this, we have computed the "true" minimum-variance portfolios corresponding to the "true" asset-return distribution G for the data sets with h equal to 0%, 2.5%, and 5%, and we have observed that the "true" portfolios are very close to the 50th percentile of the distribution of the weights of the estimated minimum-variance portfolios.[8] The reason for this is that the minimum-variance portfolios computed from the sample covariance matrix assign equal importance to all sample returns (including those deviating from the normal distribution). Consequently, the estimated minimum-variance portfolios are unbiased, but inefficient, estimators of the true portfolios.

The M- and S-portfolios, on the other hand, assign a lower weight to sample returns that deviate from the

Table 1. Simulated data set: Out-of-sample variance, Sharpe ratio, and turnover for the different policies.

	Unconstrained policies						Constrained policies					
Statistic	Mean-var	Min-var	2-mean	2-var	M-Hub	S-Tuk	Mean-var	Min-var	2-mean	2-var	M-Hub	S-Tuk
Variance (0%)	0.03243	0.00163	0.04071	0.00163	0.00165	0.00164	0.00272	0.00163	0.00277	0.00162	0.00163	0.00163
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.27)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)	(0.08)	(0.00)	(0.00)
Variance (2.5%)	0.03766	0.00300	0.04527	0.00301	0.00296	0.00299	0.00478	0.00288	0.00475	0.00295	0.00293	0.00297
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.20)	(0.01)	(0.35)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(0.00)
Variance (5%)	0.04739	0.00410	0.05742	0.00428	0.00411	0.00426	0.00688	0.00399	0.00686	0.00407	0.00410	0.00424
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.00)	(0.37)	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sharpe rat. (0%)	0.05782	0.12592	0.05453	0.12555	0.12509	0.12602	0.11156	0.12681	0.11111	0.12648	0.12595	0.12679
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.06)	(0.19)	(0.43)	(0.01)	(1.00)	(0.01)	(0.08)	(0.14)	(0.49)
Sharpe rat. (2.5%)	0.09240	0.19845	0.08465	0.20222	0.20187	0.20216	0.18333	0.20237	0.18261	0.20352	0.20265	0.20268
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.00)	(0.05)	(0.08)	(0.00)	(1.00)	(0.00)	(0.00)	(0.39)	(0.40)
Sharpe rat. (5%)	0.12855	0.24333	0.11777	0.24694	0.25916	0.26006	0.23856	0.25776	0.23720	0.25828	0.26074	0.26052
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(0.02)
Turnover (0%)	2.13717	0.04555	3.41418	0.04735	0.05740	0.05167	0.14227	0.04047	0.14402	0.04169	0.04978	0.04575
Turnover (2.5%)	2.34580	0.06449	3.64195	0.05628	0.05997	0.05100	0.13756	0.03648	0.14240	0.03914	0.04815	0.04505
Turnover (5%)	2.64896	0.09055	4.09332	0.08814	0.06496	0.05073	0.12607	0.03616	0.13066	0.03731	0.04503	0.04436

Notes. This table reports the out-of-sample variance, p-value of the difference between the variance of each unconstrained or constrained policy to that of the unconstrained or constrained minimum-variance policy, the Sharpe ratio, p-value of the difference between the Sharpe ratio of each unconstrained or constrained policy to that of the unconstrained or constrained minimum-variance policy, and the turnover for the three simulated data sets with proportion of the data deviating from normality h equal to 0%, 2.5%, and 5% (shown in parentheses). The first column lists the particular out-of-sample statistic that is being reported and for which value of h (0%, 2.5%, or 5%). The remaining 12 columns report the values of the out-of-sample statistics for each of the portfolio policies considered.

normal distribution. We know from robust statistics that this is precisely what makes these portfolios efficient estimators for a reasonable range of possible deviations. However, because the robust portfolios assign a lower weight to some of the return samples, they are biased estimators of the true minimum-variance portfolios.[9] That is, the M- and S-portfolios are biased but efficient estimators of the true minimum-variance portfolios. Consequently, there is a trade-off between the traditional minimum-variance portfolios (which are unbiased but inefficient) and the M- and S-portfolios (which are biased but efficient).

We now study the impact of imposing short-selling constraints on the portfolio policies considered. Figure 4 in the online appendix gives the boxplots for the Min-var, 2-var, M-Hub, and S-Tuk policies *with* a constraint on short-selling. From panel (a), it is clear that constraints help to induce some further stability in the portfolio weights of the minimum-variance policy. However, it is also clear that, even in the presence of short-selling constraints, the minimum-variance portfolio weights become more unstable as the proportion of the data deviating from normality increases. Panel (b) shows that the portfolio weights of the constrained 2-var policy also become less stable as the proportion of the data deviating from normality increases. Panel (d), on the other hand, shows that the stability of the weights of the constrained S-Tuk portfolio does not change much as the proportion of the data deviating from normality increases. Finally, the stability of the constrained M-Hub policy seems less sensitive to deviations from normality than that of the constrained Min-var and 2-var policies, but a bit more sensitive than that of the constrained S-Tuk policy.

Summarizing, the stability of the S-Tuk portfolio weights (both unconstrained and constrained) is the least sensitive to the presence of deviations of the return distribution from normality. The stability of the M-Hub portfolio weights is also quite insensitive to the presence of deviations from normality, whereas the stability of the Min-var and 2-var portfolio weights is much more sensitive to deviations of the asset-return distribution from normality. Finally, the portfolio weights of the mean-variance and the two-step robust mean-variance portfolios are highly unstable even for normally distributed returns.

5.2.3. Discussion of Variance, Sharpe Ratio, and Turnover. Table 1 reports the out-of-sample variance, p-value of the difference between the variance of each policy and that of the minimum-variance policy, out-of-sample Sharpe ratio, p-value of the difference between the Sharpe ratio of each policy and that of the minimum-variance policy, and turnover of each of the 12 policies considered and for each of the three simulated data sets with proportion of the return data deviating from normality h equal to 0%, 2.5%, and 5%.

Our first observation is that portfolios that optimize the in-sample trade-off between risk and return (i.e., the mean-variance and the two-step robust mean-variance portfolios) perform much worse among all three out-of-sample criteria (variance, Sharpe ratio, and turnover) than the portfolios that minimize risk. This insight holds for all three data sets with h equal to 0%, 2.5%, and 5% and both in the absence and presence of shortselling constraints. This confirms that estimates of expected returns (standard and robust) are so inaccurate that using them for portfolio selection tends to worsen out-of-sample performance.

We now discuss the performance of the portfolio policies that are constructed using only estimates of portfolio risk: Min-var, 2-var, M-Hub, and S-Tuk. We start by studying the performance of these policies in the absence of short-selling constraints. It is clear from Table 1 that when the return data follows the normal distribution ($h = 0\%$), all minimum-risk portfolios (robust and classical) perform similarly and there are no big differences in their performance in terms of variance, Sharpe ratio, or turnover. Note also that as the proportion of the data deviating from normality increases, the Sharpe ratios of all portfolios increase too. This is not surprising because we are reporting results for the case where the returns that deviate from normality are equal to the expected return of each asset *plus* five times the standard deviation of the asset return—as we mentioned before, we have tried other types of deviation and the overall insights from the results are similar. Note, however, that the out-of-sample Sharpe ratio of the Min-var portfolio is worse than that of the M-Hub and S-Tuk portfolios for the data sets where 2.5% and 5% of the returns deviate from normality, and the difference is statistically significant for the data set with $h = 5\%$. That is, the performance of the Min-var portfolio gets worse than that of the M-Hub and S-Tuk portfolios as the proportion of data deviating from normality increases. The turnover of the minimum-variance policy also increases substantially when the asset-return distribution deviates from normality. The turnover of the unconstrained M-Hub and S-Tuk portfolios, on the other hand, is quite insensitive to the presence of return data deviating from normality. Finally, the 2-var policy is a bit more sensitive to the presence of asset-return data deviating from normality than the M and S-Tuk portfolios, but it is less sensitive than the traditional Min-var portfolio.

We now study the effect of imposing short-selling constraints on the portfolio policies constructed using only estimates of portfolio risk. From Table 1, we observe that for the data set that follows the normal distribution ($h = 0\%$), the introduction of constraints helps to slightly reduce the out-of-sample variance, increase the Sharpe ratio, and decrease the turnover of all of these policies, but the effect is quite mild. Also, in the absence of any deviations from normality, all short-selling constrained portfolios (traditional and robust) perform similarly in terms of variance, Sharpe ratio, and turnover. However, for the data sets that contain data deviating from normality, the Sharpe ratio of the constrained Min-var portfolio is worse than that of the M-Hub and S-Tuk portfolios, and the difference is statistically significant for the case with $h = 5\%$. Note, however, that the turnover of the constrained minimum-variance portfolios is not sensitive to deviations from normality. This is surprising because from the boxplots in Figure 4 in the online appendix we observe that the constrained minimum-variance portfolio does indeed change when h increases from 0% to 5%. Concretely, the median weight on the second asset is 18% for $h = 0\%$, whereas it is 0% for the data set with $h = 5\%$. A similar effect can be observed for the weight for the first asset. Thus, although the constrained minimum-variance portfolios are indeed sensitive to deviations from normality, their turnover is not because, on average, 50% of the dates this policy assigns a very stable (zero) weight to assets 1 and 2. This, however, has a negative impact on performance, as can be observed from the fact that the Sharpe ratio of the constrained minimum-variance portfolio gets worse than that of the constrained M-Hub and S-Tuk portfolios when h grows increases from 0% to 5%. The turnovers of the constrained M-Hub and S-Tuk portfolios are very insensitive to deviations of the return distribution from normality although they keep assigning positive weights to all four assets.

Summarizing, our results show that the M-Hub and S-Tuk portfolios attain higher out-of-sample Sharpe ratios and lower turnovers than the traditional Min-var portfolios when the asset-return distribution deviates from normality. The imposition of constraints helps to reduce the impact of the deviations from normality on the minimum-variance policy, but our proposed policies have slightly better Sharpe ratios even in the presence of short-selling constraints when the return distribution deviates from normality. Also, the performance of the 2-var portfolio is better than that of the minimum-variance portfolio, but worse than that of our proposed policies. Finally, the mean-variance and two-step robust mean-variance portfolios are substantially and significantly outperformed by all policies that ignore estimates of expected return.

Finally, the simulated data set allows us to explore how the different portfolios perform on those dates when the asset returns deviate from normality. We have explored this issue on a simulated data set containing 5% of returns deviating from normality, including returns deviating both positively and negatively from the normal distribution. Table 3 in the online appendix gives the results. Our main observation is that the returns of the proposed minimum M-risk and S-risk portfolios are not very different from those of the minimum-variance portfolios on market-crisis and market-boom dates. Nevertheless, the out-of-sample Sharpe ratio of the M- and S-portfolios over the totality of the data (including all dates) is usually higher than the Sharpe ratio of the traditional portfolios. This is because although the performance of the traditional and robust portfolios is similar on market-crisis (or market-boom) days, the robust portfolios tend to perform better on normal days. Also, we note that the traditional and robust mean-variance portfolios attain relatively high mean returns on market boom and market crisis. However, portfolio return variance for these portfolios is very large and, as a result, their out-of-sample Sharpe ratios are relatively small for the totality of the data in the absence of short-selling constraints, and only slightly larger than that of the minimum-variance portfolio in the presence of short-selling constraints. Finally,

Table 2. Ten S&P sector portfolios and market: Out-of-sample variance, Sharpe ratio, and turnover.

	Unconstrained policies						Constrained policies					
Statistic	Mean-var	Min-var	2-mean	2-var	M-Hub	S-Tuk	Mean-var	Min-var	2-mean	2-var	M-Hub	S-Tuk
Mean	0.03075	0.00304	0.03937	0.00280	0.00318	0.00306	0.00733	0.00322	0.00703	0.00332	0.00410	0.00372
Variance	0.21249	0.00138	0.29658	0.00135	0.00164	0.00135	0.00674	0.00142	0.00728	0.00141	0.00140	0.00136
p-val.-(min-var)	(0.00)	(1.00)	(0.00)	(0.05)	(0.00)	(0.30)	(0.00)	(1.00)	(0.00)	(0.45)	(0.39)	(0.08)
Sharpe ratio	0.06672	0.08200	0.07230	0.07608	0.07861	0.08330	0.08924	0.08552	0.08233	0.08835	0.10971	0.10059
p-val.-(min-var)	(0.45)	(1.00)	(0.46)	(0.14)	(0.46)	(0.44)	(0.50)	(1.00)	(0.50)	(0.27)	(0.05)	(0.15)
Turnover	43.41806	0.19964	106.38955	0.18824	0.14353	0.21483	0.13805	0.07276	0.14024	0.07078	0.06101	0.07206

Notes. This table reports the out-of-sample mean, variance, p-value of the difference between the variance of each unconstrained or constrained policy to that of the unconstrained or constrained minimum-variance policy, Sharpe ratio, p-value of the difference between the Sharpe ratio of each unconstrained or constrained policy to that of the unconstrained or constrained minimum-variance policy, and turnover for the data set corresponding to 10 S&P500 sector tracking portfolios and the market. The first column lists the particular out-of-sample statistic that is being reported. The remaining 12 columns report the values of the out-of-sample statistics for each of the portfolio policies considered.

because abnormal-market dates are those that really make a big difference in the evolution of wealth under management of any fund, it would be interesting to complement (in future research efforts) the results in Table 3 with an out-of-sample analysis of the terminal cumulative return implied by each portfolio considered. The relative importance of portfolios' value-insurance prowess and market-timing ability at abnormal-market dates is likely to be better understood in the context of such an analysis.

5.3. Empirical Results

We use an empirical data set with 11 assets. The first 10 assets are portfolios tracking the 10 sectors composing the S&P500 index and the eleventh asset is the U.S. market portfolio represented by the S&P500 index. The data span from January 1981 to December 2002. The returns are expressed in excess of the 90-day T-bill.[10]

5.3.1. Discussion of Portfolio Weight Stability. Our first observation is that the weights of the mean-variance and two-step robust mean-variance portfolios are much more unstable than those of the minimum-variance portfolio. For example, in our rolling-horizon experiment, the mean-variance portfolio weight on the eleventh asset ranges between −3,200% and −350%, and the two-step robust mean-variance portfolio weight on this same asset ranges between −3,600% and −350%, whereas the minimum-variance weight ranges only between −150% and 70%.

We now focus on the policies that use estimates of portfolio risk only. Figure 5 in the online appendix gives the portfolio weight boxplots for the unconstrained Min-var, 2-var, M-Hub, and S-Tuk policies. From the figure, it is clear that the M-Hub portfolio weights are the most stable, followed by the S-Tuk, 2-var, and Min-var portfolio weights, in this order. In particular, the portfolio weight corresponding to the eleventh asset ranges between −75% and 15% for M-Hub, −105% and 55% for S-Tuk, −145% and 70% for 2-var, and −150% and 70% for Min-var. Figure 6 in the online appendix gives the boxplots for the constrained policies. Although it is clear that the introduction of short-selling constraints substantially improves the stability of all policies, it can also be observed that the M-Hub policy is slightly more stable than the rest of the policies, even in the presence of constraints.

5.3.2. Discussion of Variance, Sharpe Ratio, and Turnover. Table 2 gives the out-of-sample results for all policies. Note that the variance and turnover of the mean-variance and the two-step robust mean-variance portfolios are much larger than those of the rest of the portfolios both in the presence and in the absence of short-selling constraints. In addition, the out-of-sample Sharpe ratio of the mean-variance policy is statistically indistinguishable from that of the minimum-variance policy. This again confirms that nothing much is lost by ignoring estimates (standard or robust) of mean returns.

We now focus on the rest of the policies that are constructed from estimates of portfolio risk only: Min-var, 2-var, M-Hub, and S-Tuk. The performance of the unconstrained versions of these four policies is quite similar—their Sharpe ratios are statistically indistinguishable. In terms of portfolio weight stability, the smallest turnover is that of the unconstrained M-Hub policy, which is coherent with the insights obtained from the boxplots. The imposition of short-selling constraints improves the performance of all policies, but the improvement is larger for the M-Hub and S-Tuk policies. In particular, the constrained M-Hub and S-Tuk policies have higher out-of-sample Sharpe ratios than the minimum-variance policy, and the p-values for the differences are relatively significant (5% and 15%, respectively). M-Hub and S-Tuk also have higher Sharpe ratios than the mean-variance and two-step robust mean-variance portfolios. Finally, as in the unconstrained case, the constrained M-Hub policy has the lowest turnover.

Summarizing, from our empirical results we conclude that the unconstrained and constrained M-Hub and S-Tuk policies have the most stable portfolio weights. Also, the out-of-sample Sharpe ratios of the constrained M-Hub and S-Tuk policies are larger than that of the constrained minimum-variance portfolio.

Finally, we would like to mention that passive indexes are a good practical alternative to the proposed M- and S-portfolios because they attain high Sharpe ratios while being even more stable than our proposed portfolios. For example, for the S&P500 sectors data set, the value-weighted index attains a Sharpe ratio of 0.1444, which is higher than that of the constrained M- and S-portfolios. Likewise, 5 out of the 10 sector portfolios attain higher out-of-sample Sharpe ratios than the constrained M- and S-portfolios; see Table 4 in the online appendix. This is not surprising given that there is a longstanding discussion in the literature on whether active portfolios offer any advantage over passive (index) portfolios. The evidence is mixed. DeMiguel et al. (2005) show that the constrained minimum-variance portfolio outperforms the value-weighted index in five out of the six data sets considered in their paper. Likewise, Wermers (2000) finds that equity mutual funds outperform the market (although expenses reduce this benefit). Bogle (1995), Malkiel (1995), and Gruber (1996), on the other hand, find that a large fraction of active equity managers have been outperformed by index funds. In this paper, however, we have focused on the situation where the investor is interested in active funds, and for this reason most of our discussion has focused on the comparison of the proposed M- and S-portfolios to other active portfolios.

6. Conclusion

We have characterized the influence functions for the weights of the M- and S-portfolio policies. These influence functions demonstrate that the weights of the robust policies are less sensitive to deviations of the asset-return distribution from normality than those of the traditional minimum-variance policy. Moreover, our numerical results confirm that the proposed policies are indeed more stable. The stability of the proposed portfolios makes them a credible alternative to the traditional portfolios because investors are usually reticent to implement policies whose recommended portfolio weights change drastically over time.

The numerical results also show that portfolios that optimize the trade-off between in-sample risk and return are usually outperformed by minimum-risk portfolios in terms of their out-of-sample Sharpe ratios. Also, the proposed M-Hub and S-Tuk portfolios improve the stability properties of the traditional minimum-variance portfolios while preserving (or slightly improving) their good out-of-sample Sharpe ratios. Finally, the explanation for the good behavior of the proposed policies is that because they are based on robust estimation techniques, they are much less sensitive to deviations of the asset-return distribution from normality than the the traditional portfolios.

7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

Endnotes

1. For example, Mandelbrot (1963) observed that asset-return distributions have heavier tails than the normal distribution. A number of papers study the use of *stable* distributions (instead of normal distributions) to model asset returns; see Simkowitz and Beedles (1980), Tucker (1992), Ortobelli et al. (2002), and the references therein. Also, see Das and Uppal (2004) and the references therein for evidence on jumps in the returns of international equities.
2. If a solution to the functional form of the first-order optimality conditions is uniquely defined, then the estimators based on the optimality conditions (20)–(22) are consistent, see Huber (2004).
3. We have tried other risk aversion parameters such as $\gamma = 2$ and 5, but the insights from the results are similar, and thus we report the results only for the case $\gamma = 1$.
4. We have tried other estimation window lengths such as $T = 60$ and 240, but the results are similar, and thus we report the results only for the case $T = 120$.
5. To compute the p-values, we use the bootstrapping methodology described in Efron and Tibshirani (1993). Specifically, consider two portfolios i and n, with μ_i, μ_n, σ_i, σ_n as their true means and variances. We wish to test the hypothesis that the Sharpe ratio of portfolio i is equal to that of portfolio n, that is, H_0: $\mu_i/\sigma_i - \mu_n/\sigma_n = 0$. To do this, we obtain B pairs of size $T - \tau$ of the portfolio returns i and n by resampling with replacement. If $\hat{F}$ denotes the empirical distribution function of the B bootstrap pairs corresponding to $\hat{\mu}_i/\hat{\sigma}_i - \hat{\mu}_n/\hat{\sigma}_n$, then a two-sided p-value for the previous null hypothesis is given by $\hat{p} = 2\hat{F}(0)$. In a similar way, to test the hypothesis that the variances of two portfolio returns are identical, H_0: $\sigma_i^2/\sigma_n^2 = 1$, if $\hat{F}$ denotes the empirical distribution function of the B bootstrap pairs corresponding to $\hat{\sigma}_i^2/\hat{\sigma}_n^2$, then, a two-sided p-value for this null hypothesis is given by $\hat{p} = 2\hat{F}(0)$. For a nice discussion of the application of other bootstrapping methods to test the significance of Sharpe ratios, see Ledoit and Wolf (2008).
6. Das and Uppal (2004) calibrate a jump diffusion process to historical returns on the indexes for six countries. Their estimates imply that on average there will be a jump on stock returns every 20 months. This is similar to the 5% amount of data deviating from the normal distribution we use in our experiments.
7. We use the nonlinear programming code KNITRO (Byrd et al. 1999, Waltz 2004) to solve the portfolio problems. Also, the following policies need to be calibrated: Mean-Var, 2-Mean, 2-Var, M-Hub, and S-Tuk. For Mean-Var and 2-Mean, we set the risk aversion parameter $\gamma = 1$ (we have also tried $\gamma = 2$ and 5, and the insights from the results are similar, and thus we only report the results for the case $\gamma = 1$). In addition, for 2-Mean we choose the breakdown point (the amount of data deviating from the nominal distribution that a robust estimator can accept while still giving meaningful information) to 20%. For 2-Var, we choose the breakdown point to 20%. For M-Hub, we set

$c = 0.01$ in the Huber loss function. For S-Tuk, we calibrate the c constant in Tukey's biweight function so that the corresponding breakdown point is also 20%. To keep computational time short, we calibrate the 2-Mean, 2-Var, M-Hub, and S-Tuk portfolios "offline;" that is, we consider several values of the breakdown point and the parameter c and keep the values that work best for each policy.

8. Note that the true asset-return distribution G for the data sets with $h = 2.5\%, 5\%$ is a *mixture* of normals, which is not normal in general. It is easy, however, to calculate the covariance matrix of G and hence compute the corresponding "true" minimum-variance portfolios.

9. The M- and S-portfolios are unbiased estimators of the true M- and S-portfolios, which differ in general from the true minimum-variance portfolio, but should be relatively close to the true minimum-variance portfolio provided h is small.

10. We thank Roberto Wessels for creating this data set and making it available to us. The data set can be downloaded from http://faculty.london.edu/avmiguel/SPSectors.txt. For this data set, we have chosen the value of $c = 0.0001$ for the Huber loss function and the value of 0.2 for the breakdown point of the 2-Mean, 2-Var, and S-Tuk policies. For Mean-Var and 2-Mean, we choose the risk aversion parameter $\gamma = 1$ (we have also tried $\gamma = 2$ and 5, and the insights from the results are similar, and thus we only report the results for the case $\gamma = 1$).

Acknowledgments

The authors are grateful to the area editor (Mark Broadie), the associate editor, and two referees for their detailed feedback, which helped to improve this manuscript substantially. They thank Lorenzo Garlappi and Raman Uppal for their help. They are also grateful for comments from Suleyman Basak, Zeger Degraeve, Guillermo Gallego, Francisco Gomes, Garud Iyengar, Catalina Stefanescu, Garrett J. van Ryzin, Maria-Pia Victoria-Feser, Bruce Weber, Roy Welsch, William T. Ziemba, and seminar participants at London Business School, London School of Economics, Southampton University, Columbia University, Northwestern University, University of Illinois at Urbana Champaign, MIT, the 19th International Symposium on Mathematical Programming, and the 2006 INFORMS Annual Meeting. This research was partially supported by project MTM2004-02334 of the Ministerio de Educación y Ciencia de España.

References

Black, F., R. Litterman. 1992. Global portfolio optimization. *Financial Analysts J.* **48**(5) 28–43.

Bogle, J. 1995. The triumph of indexing. The Vanguard Group. (April) 1–45.

Broadie, M. 1993. Computing efficient frontiers using estimated parameters. *Ann. Oper. Res.* **45** 21–58.

Byrd, R., M. E. Hribar, J. Nocedal. 1999. An interior point method for large scale nonlinear programming. *SIAM J. Optim.* **9**(4) 877–900.

Cavadini, F., A. Sbuelz, F. Trojani. 2001. A simplified way of incorporating model risk, estimation risk and robustness in mean variance portfolio management. Working paper, Tilburg University, Tilburg, The Netherlands.

Chan, L. K. C., J. Karceski, J. Lakonishok. 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *Rev. Financial Stud.* **12**(5) 937–974.

Chopra, V. K., W. T. Ziemba. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *J. Portfolio Management* **19**(2) 6–11.

Das, S. J., R. Uppal. 2004. Systemic risk and international portfolio choice. *J. Finance* **59**(6) 2809–2834.

DeMiguel, V., L. Garlappi, R. Uppal. 2005. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financial Stud.* Forthcoming.

Efron, B., R. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Garlappi, L., R. Uppal, T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Rev. Financial Stud.* **20**(1) 41–81.

Goldfarb, D., G. Iyengar. 2003. Robust portfolio selection problems. *Math. Oper. Res.* **28**(1) 1–38.

Gruber, M. J. 1996. Another puzzle: The growth in actively managed mutual funds. *J. Finance* **51** 783–810.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel. 1986. *Robust Statistics*: *The Approach Based on Influence Functions*. John Wiley & Sons, New York.

Huber, P. J. 1977. Robust covariances. S. Gupta, D. S. Moore, eds. *Statistical Decision Theory and Related Topics*, Vol. 2. Academic Press, New York, 1753–1758.

Huber, P. J. 2004. *Robust Statistics*. John Wiley & Sons, New York.

Jagannathan, R., T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance* **58**(4) 1651–1684.

Jorion, P. 1986. Bayes-Stein estimation for portfolio analysis. *J. Financial Quant. Anal.* **21**(3) 279–292.

Lauprete, G. J. 2001. Portfolio risk minimization under departures from normality. Ph.D. thesis, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

Lauprete, G. J., A. M. Samarov, R. E. Welsch. 2002. Robust portfolio optimization. *Metrika* **25** 139–149.

Ledoit, O., M. Wolf. 2008. Robust performance hypothesis testing with the Sharpe ratio. *J. Empirical Finance* **15** 850–859.

Lu, Z. 2006. A new cone programming approach for robust portfolio selection. Working paper, Department of Mathematics, Simon Fraser University, Vancouver, British Columbia, Canada.

MacKinlay, A. C., L. Pastor. 2000. Asset pricing models: Implications for expected returns and portfolio selection. *Rev. Financial Stud.* **13**(4) 883–916.

Malkiel, B. 1995. Returns from investing in equity mutual funds 1971 to 1991. *J. Finance* **50** 549–572.

Mandelbrot, B. B. 1963. The variation of certain speculative prices. *J. Bus.* **36** 417–492.

Markowitz, H. M. 1952. Mean-variance analysis in portfolio choice and capital markets. *J. Finance* **7** 77–91.

Maronna, R. 1976. Robust m-estimators of multivariate location and scatter. *Ann. Statist.* **4** 51–67.

Merton, R. C. 1980. On estimating the expected return on the market: An exploratory investigation. *J. Financial Econom.* **8** 323–361.

Michaud, R. O. 1989. The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts J.* **45**(1) 31–42.

Ortobelli, S., I. Huber, E. S. Schwartz. 2002. Portfolio selection with stable distributed returns. *Math. Methods Oper. Res.* **55**(2) 265–300.

Pástor, L., R. F. Stambaugh. 2000. Comparing asset pricing models: An investment perspective. *J. Financial Econom.* **56**(3) 335–381.

Perret-Gentil, C., M.-P. Victoria-Feser. 2004. Robust mean-variance portfolio selection. FAME Research Paper 140. International Center for Financial Asset Management and Engineering, Geneva.

Rousseeuw, P. J., A. M. Leroy. 1987. *Robust Regression and Outlier Detection*. John Wiley & Sons, New York.

Simkowitz, M. A., W. L. Beedles. 1980. Asymmetric stable distributed security returns. *J. Amer. Statist. Assoc.* **75** 306–312.

Tucker, A. L. 1992. A reexamination of finite and infinite-variance distributions as models of daily stock returns. *J. Bus. Econom. Statist.* **10** 73–81.

Tukey, J. W. 1960. A survey of sampling from contaminated distributions. I. Olkin, ed. *Contributions to Probability and Statistics*. Stanford University Press, Stanford, CA, 448–485.

Tütüncü, R. H., M. Koenig. 2004. Robust asset allocation. *Ann. Oper. Res.* **132**(1) 157–187.

Vaz-de Melo, B., R. P. Camara. 2003. Robust modeling of multivariate financial data. Coppead Working Paper Series 355, Federal University at Rio de Janeiro, Rio de Janeiro, Brazil.

Waltz, R. A. 2004. *KNITRO User's Manual. Version* 4.0. Ziena Optimization, Evanston, IL.

Welsch, R. E., X. Zhou. 2007. Application of robust statistics to asset allocation models. *Revstat. Statist. J.* **5**(1) 97–114.

Wermers, R. 2000. Mutual fund performance: An empirical decomposition into stock picking talent, style, transaction costs and expenses. *J. Finance* **55** 1655–1703.