Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection

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

Alberto Martín-Utrera
Department of Statistics, Universidad Carlos III de Madrid, 28903-Getafe (Madrid), Spain, amutrera@est-econ.uc3m.es

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

We carry out a comprehensive investigation of shrinkage estimators for portfolio selection. We study both portfolios computed from shrinkage estimators of the moments of asset returns (*shrinkage moments*), as well as *shrinkage portfolios* obtained by shrinking the portfolio weights directly. We make several contributions. First, we propose two novel shrinkage estimators of moments: one for mean returns and one for the inverse covariance matrix. Second, we propose a new calibration criterion for the shrinkage estimator of the covariance matrix that takes the condition number into account. Third, for shrinkage portfolios we introduce an additional scaling parameter that we adjust to minimize the bias of the *target portfolio*. Fourth, we study three novel calibration criteria for shrinkage portfolios. Fifth, we propose a parametric and a nonparametric approach to estimate the optimal shrinkage intensity. The parametric approach provides closed-form expressions under the assumption that returns are independent and identically distributed as a normal, and the nonparametric approach relies on bootstrapping. Finally, we carry out extensive tests on simulated and empirical datasets and characterize the performance of the different shrinkage estimators.

Key words: Portfolio choice, estimation error, shrinkage intensity, bootstrap.

---

1. Introduction

The classical mean-variance framework for portfolio selection proposed by Markowitz (1952) formalized the concept of investment diversification, and is widely used nowadays in the investment industry. To compute mean-variance portfolios, one needs to estimate the mean and covariance matrix of asset returns. One possibility is to replace these quantities with their sample estimators, but these are obtained from historical return data and contain substantial estimation error. As a result, mean-variance portfolios computed from sample estimators perform poorly out of sample; see, for instance, Jobson and Korkie (1981), Best and Grauer (1991), Broadie (1993), Britten-Jones (1999), DeMiguel et al. (2009).

One of the most popular approaches to combat the impact of estimation error in portfolio selection is to use shrinkage estimators, which are obtained by "shrinking" the sample estimator towards a target estimator.[1] The advantage is that while the shrinkage target is usually biased, it also contains less variance than the sample estimator. Thus it is possible to show under general conditions that there exists a shrinkage *intensity* for which the resulting *shrinkage* estimator contains less estimation error than the original sample estimator; see James and Stein (1961). The key then is to characterize the optimal trade-off between the sample estimator (low bias), and the target (low variance). In other words, shrinkage estimators can help reduce estimation error, but the shrinkage intensity (*size*) matters.

In this paper, we carry out a comprehensive investigation of shrinkage estimators for portfolio selection. We study both portfolios computed from shrinkage estimators of the moments of asset returns (*shrinkage moments*), as well as *shrinkage portfolios* obtained by shrinking directly the portfolio weights computed from the original (un-shrunk) sample moments. We make three contributions in the area of shrinkage moments. First, we propose two novel shrinkage estimators of moments: one for mean returns and one for the inverse covariance matrix. Second, we propose a new calibration criterion for the popular shrinkage estimator of the covariance matrix proposed by Ledoit and Wolf (2004b). Our calibration criterion is novel in that it captures not only the quadratic loss (as in Ledoit and Wolf (2004b)), but also the condition number of the shrinkage estimator of the covariance matrix. Third, we propose novel parametric and nonparametric approaches to estimate the optimal shrinkage intensity for the shrinkage moments. The parametric approach provides closed-form expressions for the optimal shrinkage intensity under the assumption that returns are independent and identically distributed as a normal distribution. The nonparametric approach makes no assumptions about the return distribution and relies on bootstrapping; see Efron (1979).

We consider three different shrinkage portfolios from the existing literature, and we make three contributions in this area. First, we introduce an additional scaling parameter that we adjust to minimize the bias of the *target portfolio*. Second, we consider three novel calibration criteria for the shrinkage portfolios (quadratic loss, portfolio variance, and Sharpe ratio), in addition to the standard criterion used in the exist-

[1] Other approaches proposed to combat estimation error in portfolio selection include: Bayesian methods (Barry (1974), Bawa et al. (1979)), Bayesian methods with priors obtained from asset pricing models (MacKinlay and Pastor (2000), Pastor (2000), Pástor and Stambaugh (2000)), robust optimization methods (Cornuejols and Tütüncü (2007), Goldfarb and Iyengar (2003), Garlappi et al. (2007), Rustem et al. (2000), Tutuncu and Koenig (2004)), Bayesian robust optimization (Wang (2005)), robust estimation methods (DeMiguel and Nogales (2009)), and imposing constraints (Best and Grauer (1992), Jagannathan and Ma (2003), and DeMiguel et al. (2009)).

ing literature (expected utility). Third, we propose a parametric and a nonparametric approach to estimate the optimal shrinkage intensity for the three shrinkage portfolios and four calibration criteria.

1.1. Related literature

The existing literature in portfolio selection has studied three different types of shrinkage estimators for portfolio selection: (i) estimators obtained by shrinking the mean of asset returns, (ii) the covariance matrix of asset returns, and (iii) the portfolio weights directly. Jorion (1986) proposes a shrinkage estimator of mean returns that he derives using an empirical Bayes-Stein approach.[2] His shrinkage estimator is the mean of a predictive density function obtained from an informative prior that belongs to the class of exponential distributions. Frost and Savarino (1986) study an empirical Bayesian approach to portfolio selection for an investor whose prior belief about the vector of means and the covariance matrix is jointly defined with a Normal-Wishart conjugate prior. This prior distribution results into a posterior shrinkage vector of means, which is calibrated using the maximum likelihood estimators of the parameters that define the strength of belief in the prior. Finally, Jobson et al. (1979) and Jorion (1985) also study the use of shrinkage estimators for the vector of means.

Ledoit and Wolf have proposed several shrinkage estimators of the covariance matrix of asset returns. Ledoit and Wolf (2003) propose a shrinkage estimator for the covariance matrix which is a weighted average of the sample covariance matrix and a single-index covariance matrix implied by the market factor model. Ledoit and Wolf (2004a) propose using as a shrinkage target the "constant correlation matrix", whose correlations are set equal to the average of all sample correlations. Finally, Ledoit and Wolf (2004b) propose using a multiple of the identity matrix as the shrinkage target. They show that the resulting shrinkage covariance matrix is well-conditioned, even if the sample covariance matrix is not. Ledoit and Wolf (2003, 2004a, 2004b) calibrate their shrinkage covariance matrices by minimizing their expected quadratic loss. They characterize the asymptotically optimal shrinkage intensity under the assumption that asset returns are independent and identically distributed (iid) with finite fourth moments.

Finally, several authors have proposed shrinkage portfolios obtained by shrinking directly the portfolio weights. Kan and Zhou (2007) study a three-fund portfolio that is a combination of the sample mean-variance portfolio, the sample minimum-variance portfolio, and the risk-free asset. They show that an optimal combination of the three funds diminishes the effects of estimation error in the investor's expected utility function. Tu and Zhou (2011) consider the optimal trade-off between the sample mean-variance portfolio and the equally-weighted portfolio. DeMiguel et al. (2009) study the empirical performance of a mixture of

[2] Note that there is a close relation between Bayesian portfolios and shrinkage estimators. Specifically, one can always view the target estimator as prior knowledge about the distribution of asset returns.

portfolios obtained as a combination of the sample minimum-variance portfolio and the equally-weighted portfolio. All these papers calibrate the shrinkage portfolios by maximizing the investor's expected utility, and they characterize the optimal shrinkage intensity under the assumption of iid normal returns.

1.2. Overview

We consider three shrinkage estimators of the moments of asset returns. First, we propose a new shrinkage estimator for the vector of means. Unlike the estimator proposed by Jorion (1986), our estimator is defined, a priori, as a convex combination of the sample mean and a target element, and we select the convexity parameter, or shrinkage intensity, to minimize the expected quadratic loss. Moreover, unlike Jorion (1986), we do not need to make any assumptions about the asset return distribution to obtain a closed-form expression for the optimal shrinkage intensity, instead we only need to assume that asset returns are iid. Second, we consider the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b) and consider the same calibration criterion (expected quadratic loss). Unlike Ledoit and Wolf (2004b), however, we provide a closed-form expression of the optimal shrinkage intensity *for finite samples* by assuming that returns are iid *normal*. Third, we propose a new shrinkage estimator for the inverse covariance matrix. This shrinkage estimator is constructed as a convex combination of the inverse of the sample covariance matrix and a scaled identity matrix. Under iid normal returns, we provide a closed-form expression of the true optimal shrinkage intensity that minimizes the expected quadratic loss.[3] Moreover, we propose a new calibration criterion for the shrinkage covariance matrix that takes into account not only the expected quadratic loss but also its condition number, and show how the corresponding shrinkage intensity can be obtained numerically. The condition number gives a bound for the sensitivity of the computed portfolio weights to estimation errors in the mean and covariance matrix of asset returns, and thus calibrating the shrinkage covariance matrix so that its condition number is relatively small helps to reduce the impact of estimation error in portfolio selection. Indeed, our experiments with simulated and empirical data demonstrate the advantages of using this criterion for the construction of minimum-variance portfolios. Finally, we show how the shrinkage estimators of the covariance and inverse covariance matrices can be calibrated using a nonparametric smoothed bootstrap procedure that makes no assumptions about the return distribution.

We consider three different shrinkage portfolios. The first is obtained by shrinking the sample mean-variance portfolio towards the sample minimum-variance portfolio and is closely related to the three-fund

[3] Frahm and Memmel (2010) and Kourstis et al. (2011) also propose shrinkage estimators for the inverse covariance matrix. The difference between our approach and their approaches is that while Frahm and Memmel (2010) and Kourstis et al. (2011) calibrate their estimators to minimize the out-of-sample variance of the minimum-variance portfolio estimator, we calibrate our estimator to minimize the expected quadratic loss of the inverse covariance matrix. We select this calibration criterion because Ledoit and Wolf (2004b) show that it results in good performance within the context of the covariance matrix.

portfolio of Kan and Zhou (2007), the second is obtained by shrinking the sample mean-variance portfolio towards the equally-weighted portfolio as in Tu and Zhou (2011), and the third is obtained by shrinking the sample minimum-variance portfolio towards the equally-weighted portfolio, similar to DeMiguel et al. (2009). We make three contributions in this area. First, we introduce an additional scaling parameter that we adjust to minimize the bias of the target portfolio. The advantage of introducing this scaling parameter is that by reducing the bias of the target, we also reduce the overall quadratic loss of the resulting shrinkage portfolio, and our empirical results show that, in general, this improves the out-of-sample performance of the shrinkage portfolios. Most of the existing literature calibrates the shrinkage portfolios by maximizing the investor's expected utility. Our second contribution is to consider, in addition to the expected utility criterion, three novel calibration criteria: expected quadratic loss minimization, portfolio variance minimization, and Sharpe ratio maximization. We consider the expected quadratic loss criterion because of its good performance in the context of shrinkage covariance matrices (see Ledoit and Wolf (2004a)). The other two calibration criteria (variance and Sharpe ratio) are relevant particular cases of the expected utility maximization criterion. Our third contribution is to study a parametric and a nonparametric approach to estimate the optimal shrinkage intensity for all three shrinkage portfolios and four calibration criteria. The parametric approach relies on the assumption that returns are iid normal, and results in closed-form expressions for the optimal shrinkage intensities for every calibration criterion, except for the Sharpe ratio criterion, for which it can be computed numerically. We also implement a nonparametric bootstrap approach to estimate the optimal shrinkage intensity of the shrinkage portfolios, which does not require any assumptions on the distribution of asset returns. To the best of our knowledge, this is the first work to consider such a nonparametric approach for shrinkage portfolios.

Finally, we evaluate the out-of-sample performance of the portfolios obtained from shrinkage moments, as well as that of the shrinkage portfolios on the six empirical datasets listed in Table 1. For portfolios computed from shrinkage moments, we identify two main findings. First, our proposed shrinkage estimator for the vector of means substantially reduces the estimation error, improving the out-of-sample performance of mean-variance portfolios. Second, taking the condition number of the estimated covariance matrix into account improves the quality of its shrinkage estimators. For shrinkage portfolios we identify two main findings. First, we find that for shrinkage portfolios that consider the vector of means, the best calibration criterion is the portfolio variance. Second, for shrinkage portfolios that do not consider the vector of means, the best calibration criterion is to minimize the expected quadratic loss. Finally, for both shrinkage moments and shrinkage portfolios, we find that the nonparametric approach to estimate the optimal shrinkage intensity tends to work better than the parametric approach.

Summarizing, we make the following contributions. First, we propose new shrinkage estimators of moments of asset returns (one for the mean and one for the inverse covariance matrix). Second, we consider new calibration criteria for both shrinkage moments and shrinkage portfolios. For the shrinkage covariance matrix we consider a calibration criterion that captures the condition number, and for the shrinkage portfolios we consider expected quadratic loss, portfolio variance, and portfolio Sharpe ratio. Third, for shrinkage portfolios we introduce a scaling parameter that we adjust to minimize the bias of the target portfolios. Fourth, we study a parametric approach to compute the optimal shrinkage intensity for the case with iid normal returns, and a nonparametric approach for the case where returns are just iid. Finally, we carry out a comprehensive empirical investigation of the characteristics of shrinkage estimators for portfolio selection on six empirical datasets.

The paper is organized as follows. Section 2 focuses on shrinkage moments, and Section 3 on shrinkage portfolios. Section 4 gives the results of the simulation experiment, and Section 5 compares the performance of the different shrinkage estimators on six empirical datasets. Section 6 concludes.

2. Shrinkage estimators of moments

In this section, we consider three shrinkage estimators of the moments of asset returns: a new shrinkage estimator of mean asset returns, the shrinkage estimator of the covariance matrix proposed by Ledoit and Wolf (2004b), and a new shrinkage estimator of the inverse covariance matrix.

We first consider the expected quadratic loss as the calibration criterion, like Ledoit and Wolf (2004b), and we show how to estimate the optimal shrinkage intensity for each of the three estimators. Specifically, we give a closed-form expression for the true optimal shrinkage intensity of our proposed shrinkage estimator for the mean, assuming only that returns are iid, but making no other distributional assumptions. For the shrinkage estimators of the covariance and the inverse covariance matrix, we give closed-form expressions for the true optimal shrinkage intensity when returns are iid normal, and we show how a nonparametric procedure can be used to estimate the optimal shrinkage intensity when returns are just iid.

Moreover, we propose a new calibration criterion for the covariance matrix that takes into account not only the expected quadratic loss, but also the condition number of the covariance matrix, which measures the sensitivity of the portfolio weights to estimation error in the mean and covariance matrix of asset returns. We show how the optimal shrinkage intensity can be estimated by solving numerically an optimization problem for both the cases with iid normal returns and just iid returns.

We now give the mathematical statement of the three shrinkage estimators of moments we consider. Let the vector of excess returns $R_t \in \mathbb{R}^N$ be iid along time $t = 1, \ldots, T$, with vector of means μ and covariance matrix Σ. Then, the *sample* vector of means is $\mu_{sp} = (1/T)\sum_{t=1}^{T} R_t$, and the *sample* covariance matrix is $\Sigma_{sp} = (1/(T-1))\sum_{t=1}^{T}(R_t - \mu_{sp})(R_t - \mu_{sp})'$. The shrinkage estimators for the vector of means, the covariance matrix, and the inverse covariance matrix are defined as a convex combination of the sample estimator and a scaled shrinkage target:

$$\mu_{sh} = (1-\alpha)\mu_{sp} + \alpha\nu\mu_{tg} \tag{1}$$

$$\Sigma_{sh} = (1-\alpha)\Sigma_{sp} + \alpha\nu\Sigma_{tg} \tag{2}$$

$$\Sigma_{sh}^{-1} = (1-\alpha)\Sigma_{sp}^{-1} + \alpha\nu\Sigma_{tg}^{-1}, \tag{3}$$

where α is the shrinkage intensity and ν is a scaling parameter that we adjust to minimize the bias of the shrinkage target. The shrinkage intensity α determines the "strength" under which the sample estimator is shrunk towards the scaled shrinkage target, and it takes values between zero and one. When the "strength" is one, the shrinkage estimator equals the scaled shrinkage target, and when α is zero, the shrinkage estimator equals the sample estimator.

We introduce the scaling parameter ν for two reasons. First, the scaling parameter yields a more general type of combination between the sample estimator and the target than just a convex combination. Second, we adjust the scaling parameter to reduce the bias of the shrinkage target. In the case where the calibration criterion is the quadratic loss, this results in a higher optimal shrinkage intensity α than that for the case without scaling parameter. This is likely to result in more stable estimators that are more resilient to estimation error.

The remainder of this section is organized as follows. In Section 2.1, we give a closed-form expression for the shrinkage intensity that minimizes the expected quadratic loss of the shrinkage estimator of means under the assumption that returns are iid. In Sections 2.2 and 2.3, we show that the shrinkage intensity that minimizes the expected quadratic loss of the shrinkage estimator of the covariance and inverse covariance matrices can be written as the relative expected loss of the sample estimator. In Section 2.4, we propose a new calibration criterion for the shrinkage covariance matrix that takes into account not only the expected quadratic loss, but also the condition number of the estimated covariance matrix. Section 2.5 gives closed-form expressions for the expectations required to compute the optimal shrinkage intensities for the estimators of the covariance and inverse covariance matrices assuming returns are iid normal. Section 2.6 proposes an alternative nonparametric approach to estimate the optimal shrinkage intensities of the shrinkage covariance and inverse covariance matrices.

2.1. Shrinkage estimator of mean returns

In this section, we propose a new shrinkage estimator of mean returns and give a closed-form expression for the shrinkage intensity that minimizes the expected quadratic loss under the assumption that returns are iid, but without making any further assumptions about the return distribution.

Several Bayesian approaches from the finance literature result in estimators of mean returns that can be interpreted as shrinkage estimators. Frost and Savarino (1986) assume an informative Normal-Wishart conjugate prior where all stocks have the same expected values, variance and covariances. The predictive mean turns out to be a weighted average of the sample mean and a prior mean, defined as the historical average return for all stocks[4]. Jorion (1986) estimates the vector of means by integrating a predictive density function defined by an exponential prior which is only specified for the vector of means. The resulting estimator is defined as a weighted average of the sample mean μ_{sp} and the portfolio mean of the minimum-variance portfolio. Like DeMiguel et al. (2009), we use the Bayes-Stein estimator of means proposed by Jorion (1986) as a benchmark in our empirical evaluation.

We focus on a new shrinkage estimator for mean returns defined as a weighted average of the sample mean and the scaled shrinkage target $\nu\mu_{tg} = \nu\iota$, where ι is the vector of ones and ν is a scaling factor that we set to minimize the bias of the shrinkage target; that is, $\nu_\mu = \text{argmin}_\nu \|\nu\iota - \mu\|_2^2 = (1/N)\sum_{i=1}^N \mu_i = \overline{\mu}$. Then we choose the shrinkage intensity α to minimize the expected quadratic loss of the shrinkage estimator:

$$\min_{\alpha} \quad E\left[\|\mu_{sh} - \mu\|_2^2\right] \tag{4}$$

$$\text{s.t.} \quad \mu_{sh} = (1-\alpha)\mu_{sp} + \alpha\nu_\mu\iota, \tag{5}$$

where $\|x\|_2^2 = \sum_{i=1}^N x_i^2$. The rationale for choosing the shrinkage target $\iota = \mathbb{1} \in \mathbb{R}^N$ is that for the case where the shrinkage intensity is equal to one, the solution of the estimated mean-variance portfolio would be the minimum-variance portfolio, which is a common benchmark portfolio. We use the expected quadratic loss minimization rule because it is known that this criterion results in good estimators of covariance matrices within the context of portfolio optimization; see Ledoit and Wolf (2004a). Therefore, it is interesting to study the performance of this rule for the vector of means in the context of portfolio optimization. Furthermore, this flexible criterion allows us to make no assumption of the distribution of asset returns to obtain the optimal shrinkage intensity.

The following proposition gives the true optimal value of the shrinkage intensity α.

[4] For computational convenience, we do not consider this shrinkage estimator in the analysis as a benchmark. This shrinkage vector of means requires the definition of a parameter that determines the strength of belief in the prior mean. Frost and Savarino (1986) propose to estimate this parameter in an Empirical-Bayes fashion by maximizing the likelihood of the prior distribution. Since we do not have closed-form expression for the shrinkage intensity, we do not consider it as a benchmark, but rather we consider the shrinkage estimator proposed by Jorion (1986), which offers a closed-form expression of the shrinkage intensity.

PROPOSITION 1. *Assuming asset returns are iid, the shrinkage intensity α that minimizes the expected quadratic loss is:*

$$\alpha_\mu = \frac{E\left(\|\mu_{sp} - \mu\|_2^2\right)}{E\left(\|\mu_{sp} - \mu\|_2^2\right) + \|\nu_\mu \iota - \mu\|_2^2} = \frac{(N/T)\overline{\sigma^2}}{(N/T)\overline{\sigma^2} + \|\nu_\mu \iota - \mu\|_2^2}, \tag{6}$$

where $\overline{\sigma^2} = trace\,(\Sigma)\,/N$.

Note that the true optimal shrinkage intensity α is defined by the relative expected loss of the sample vector of means with respect to the total expected loss, defined by the expected loss of the sample vector of means plus the loss of the scaled vector of ones. We observe that the shrinkage intensity increases with the number of assets N, and decreases with the number of observations.

The main difference between the shrinkage estimator we consider and those proposed by Frost and Savarino (1986) and Jorion (1986) is that we use the expected quadratic loss as the calibration criterion. As a result, unlike the Bayes-Stein shrinkage intensity of Jorion (1986), our shrinkage intensity does not depend on the inverse covariance matrix. This is an advantage for the estimation of the shrinkage intensity, particularly for the case where there is a large number of assets and hence the sample covariance matrix is nearly singular.

2.2. Shrinkage estimator of the covariance matrix

We consider the shrinkage estimator of the covariance matrix defined by Ledoit and Wolf (2004b), who propose shrinking the sample covariance matrix towards a scaled identity matrix:

$$\Sigma_{sh} = (1-\alpha)\Sigma_{sp} + \alpha \nu I, \tag{7}$$

and they choose the scaling factor ν to minimize the bias of the shrinkage target I; that is, $\nu_\Sigma = \text{argmin}_\nu \|\nu I - \Sigma\|_F^2 = (1/N)\sum_{i=1}^N \sigma_i^2 = \overline{\sigma^2}$. Under the assumption of iid observations, they propose choosing the shrinkage intensity α to minimize the expected quadratic loss:

$$\min_\alpha \quad E\left[\|\Sigma_{sh} - \Sigma\|_F^2\right] \tag{8}$$

$$\text{s.t.} \quad \Sigma_{sh} = (1-\alpha)\Sigma_{sp} + \alpha \nu_\Sigma I, \tag{9}$$

where $\|X\|_F^2 = \text{trace}(X'X)$. Moreover, Ledoit and Wolf (2004b) show that plugging constraint (9) into problem (8), one can obtain the following equivalent problem:

$$\min_\alpha \quad E\left[\|\Sigma_{sh} - \Sigma\|_F^2\right] = (1-\alpha)^2 E\left[\|\Sigma_{sp} - \Sigma\|_F^2\right] + \alpha^2 \|\nu_\Sigma I - \Sigma\|_F^2. \tag{10}$$

The first-order optimality conditions for this problem show that the optimal shrinkage intensity is

$$\alpha_\Sigma = \frac{E\left(\|\Sigma_{sp}-\Sigma\|_F^2\right)}{E\left(\|\Sigma_{sp}-\Sigma\|_F^2\right)+\|\nu_\Sigma I-\Sigma\|_F^2}. \tag{11}$$

Like the shrinkage intensity for the vector of means, the shrinkage intensity for the covariance matrix is the relative expected loss of the sample covariance matrix with respect to the total expected loss of the sample covariance matrix and the scaled identity matrix. Ledoit and Wolf (2004b) obtain an estimator of α by giving consistent estimators of $E(\|\Sigma_{sp}-\Sigma\|_F^2)$, $\|\nu_\Sigma I-\Sigma\|_F^2$ and ν_Σ. We, on the other hand, will give in Section 2.5 a closed-form expression for $E(\|\Sigma_{sp}-\Sigma\|_F^2)$ under the assumption that returns are iid normal. This allows us to interpret the effects of estimation error on the true optimal shrinkage intensity as a function of the number of assets and observations considered in the sample. Furthermore, in Section 2.6, we propose an alternative nonparametric bootstrap procedure to estimate α for the case of iid returns.

2.3. Shrinkage estimator of the inverse covariance matrix

The inverse covariance matrix is required to compute the mean-variance and minimum-variance portfolio weights and thus it is a key component in portfolio selection. This is particularly the case when the number of observations T is not very large relative to the number of assets N, a common situation in portfolio selection, because in this case the covariance matrix is nearly singular and estimation error explodes when we invert a nearly singular matrix.

We propose the following shrinkage estimator of the inverse covariance matrix:

$$\Sigma_{sh}^{-1} = (1-\alpha)\Sigma_{sp}^{-1}+\alpha\nu I, \tag{12}$$

where the scaling factor ν is chosen to minimize the bias of the shrinkage target – that is, $\nu_{\Sigma^{-1}} = \text{argmin}_\nu \|\nu I-\Sigma^{-1}\|_F^2 = (1/N)\sum_{i=1}^N \sigma_i^{-2} = \overline{\sigma^{-2}}$ – and the shrinkage intensity α is chosen to minimize the expected quadratic loss:

$$\min_\alpha \quad E\left(\left\|\Sigma_{sh}^{-1}-\Sigma^{-1}\right\|_F^2\right) \tag{13}$$

$$\text{s.t.} \quad \Sigma_{sh}^{-1} = (1-\alpha)\Sigma_{sp}^{-1}+\alpha\nu_{\Sigma^{-1}} I. \tag{14}$$

Plugging constraint (14) in problem (13), we have

$$\begin{aligned}\min_\alpha \quad E\left[\left\|\Sigma_{sh}^{-1}-\Sigma^{-1}\right\|_F^2\right] &= (1-\alpha)^2 E\left[\left\|\Sigma_{sp}^{-1}-\Sigma^{-1}\right\|_F^2\right]+\alpha^2\left\|\nu_{\Sigma^{-1}}I-\Sigma^{-1}\right\|_F^2+\\ &+2(1-\alpha)\alpha E\left(<\Sigma_{sp}^{-1}-\Sigma^{-1},\nu_{\Sigma^{-1}}I-\Sigma^{-1}>\right),\end{aligned} \tag{15}$$

where $<A,B>=\text{trace}(A'B)$.

From the first-order optimality conditions for problem (15), we obtain that the optimal shrinkage intensity is

$$\alpha_{\Sigma^{-1}} = \frac{E\left(\left\|\Sigma_{sp}^{-1} - \Sigma^{-1}\right\|_F^2\right) - E\left(< \Sigma_{sp}^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right)}{E\left(\left\|\Sigma_{sp}^{-1} - \Sigma^{-1}\right\|_F^2\right) + \|\nu_{\Sigma^{-1}} I - \Sigma^{-1}\|_F^2 - 2E\left(< \Sigma_{sp}^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right)}. \tag{16}$$

Note that the optimal shrinkage intensity is given by the relative expected loss of the inverse of the sample covariance matrix with respect to the total expected loss of the inverse of the sample covariance matrix and the scaled identity matrix. In this case, the expected losses are smoothed by an element proportional to the bias of the inverse of the sample covariance matrix. Later, we give closed-form expressions for the expectations in (16) under the assumption that returns are iid normal, and we also propose a nonparametric approach to estimate these expectations assuming just iid returns.

Note that this shrinkage estimator may be very conservative because it is obtained by first inverting the sample covariance matrix, and then shrinking it towards the scaled identity matrix. If the sample covariance matrix is nearly singular, small estimation errors within the sample covariance matrix will become very large errors in the inverse covariance matrix and this will in turn result in very large shrinkage intensities (i.e. $\alpha \approx 1$). In this situation, the shrinkage inverse covariance matrix might lose valuable information about the variances and covariances of asset returns. To address this problem, we propose an alternative calibration criterion in the following section.

2.4. Shrinkage estimator of the covariance matrix considering the condition number

We propose an alternative calibration criterion for the covariance matrix that accounts for both the expected quadratic loss and the condition number of the shrinkage covariance matrix, which measures the impact of estimation error on the portfolio weights. [5]

To measure the expected quadratic loss we use the *relative improvement in average loss* (RIAL); see Ledoit and Wolf (2004b):

$$RIAL\left(\Sigma_{sh}\right) = \frac{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) - E\left(\|\Sigma_{sh} - \Sigma\|_F^2\right)}{E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right)}. \tag{17}$$

The RIAL is bounded above by one, and unbounded below. The maximum value is attained when the expected quadratic loss of the shrinkage estimator Σ_{sh} is negligible relative to the expected quadratic loss of the sample covariance matrix Σ_{sp}. The advantage of using RIAL with respect to using plain expected

[5] The condition number is a measure of the matrix singularity, and it provides a bound on the accuracy of the computed solution to a linear system. Mean-variance and minimum-variance portfolios can be interpreted as the solutions of a linear system and this is why the condition number of the estimated covariance matrix matters on the investor's portfolio. Some approaches have been already proposed to deal with this problem by shrinking the eigenvalues of the sample covariance matrix (see Stein (1975), Dey and Srinivasan (1985), Zumbach (2009)).

quadratic loss is that the RIAL is bounded above by one and thus it is easy to compare the RIAL and the condition number of the shrinkage covariance matrix. Note that to characterize the $RIAL\left(\Sigma_{sh}\right)$, it is enough to characterize the expectation $E\left(\|\Sigma_{sp}-\Sigma\|_F^2\right)$. In Section 2.5, we give a closed-form expression for this expectation. In Section 2.6 we give a nonparametric approach to estimate this expectation.

On the other hand, the condition number of the shrinkage covariance matrix Σ_{sh} is:

$$\delta_{\Sigma_{sh}} = \frac{(1-\alpha)\lambda_{\max} + \alpha\nu_{\Sigma}}{(1-\alpha)\lambda_{\min} + \alpha\nu_{\Sigma}}, \tag{18}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalues of the sample covariance matrix, respectively[6]. The smallest (and thus best) condition number is one, which is attained when α is one. In that case, the shrinkage covariance matrix coincides with the scaled identity matrix.

Therefore, we propose the following problem to find an optimal shrinkage intensity that accounts both for the expected quadratic loss and the condition number of the shrinkage covariance matrix:

$$\alpha = \text{argmin}\left\{\delta_{\Sigma_{sh}} - \phi RIAL\left(\Sigma_{sh}\right)\right\}, \tag{19}$$

where ϕ is a tuning parameter that controls for the trade-off between the $RIAL$ and the condition number. Then, if $\phi = 0$, the objective is to minimize the condition number of Σ_{sh}. In that case, the optimal shrinkage intensity would be one, since that value minimizes the condition number of Σ_{sh}. On the other hand, the larger the value of ϕ, the more important the $RIAL$ is. Then, if $\phi \to \infty$, the above formulation would be equivalent to minimize the expected quadratic loss of the shrinkage matrix.

The parameter ϕ must be exogenously specified. In our empirical analysis, we set ϕ as to minimize the portfolio variance, i.e. $\phi = \arg\min_{\phi}\sigma_{\phi}^2$, where σ_{ϕ}^2 is the portfolio variance of the minimum variance portfolio formed with the shrinkage covariance matrix Σ_{sh}, calibrated by criterion (19). To compute the portfolio variance, we use the nonparametric technique known as *cross-validation* (see Efron and Gong (1983))[7]. Since problem (19) is a highly nonlinear optimization problem, it is difficult to obtain a closed-form solution. Instead, we solve the problem numerically.

[6] See Ledoit and Wolf (2004b), equation (13), for the expression of the eigenvalues of the shrinkage covariance matrix. We use that equation to obtain the expression for the condition number of the shrinkage covariance matrix.

[7] For a given ϕ, we estimate the portfolio variance using cross-validation as follows. Let us define a sample with T observations. Then, we first delete the i-th observation from our estimation sample. Second, we compute the minimum-variance portfolio from the new sample with $T-1$ observations. This portfolio is computed with the shrinkage covariance matrix Σ_{sh} calibrated with the method defined in (19). Third, we evaluate that portfolio with the i-th observation, which was dropped out of the estimation sample. This is considered the i-th out-of-sample portfolio return. To compute the portfolio variance, we repeat the previous steps with the whole sample, obtaining a time series of T out-of-sample portfolio returns. We estimate the portfolio variance as the sample variance of the out-of-sample portfolio returns.

2.5. Parametric calibration of the shrinkage estimator of moments

Assuming iid normal returns, we now give closed-form expressions for the three different expectations that are required to compute the shrinkage intensities for the shrinkage estimators of covariance and inverse covariance matrices proposed before.

PROPOSITION 2. *Assume that asset returns are iid normal and $T > N+4$. Moreover, let us define the estimated inverse covariance matrix as $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$, which is the unbiased estimator of the inverse covariance matrix. Hence, the expected quadratic losses of the estimated covariance and inverse covariance matrices are:*

$$E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) = \frac{N}{T-1}\left(\frac{trace\left(\Sigma^2\right)}{N} + N\left(\overline{\sigma^2}\right)^2\right) \tag{20}$$

$$E\left(\left\|\Sigma_u^{-1} - \Sigma^{-1}\right\|_F^2\right) = trace\left(\Omega\right) - trace\left(\Sigma^{-2}\right), \tag{21}$$

and

$$E\left(< \Sigma_u^{-1} - \Sigma^{-1}, \nu_{\Sigma^{-1}} I - \Sigma^{-1} >\right) = 0, \tag{22}$$

where $\overline{\sigma^2} = trace\left(\Sigma\right)/N$ and $\Omega = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(trace\left(\Sigma^{-1}\right)\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right)$.

Note that the expected quadratic loss of the sample estimators increases with the number of assets and decreases with the number of observations. Also, note that we have modified the expression of the estimated inverse covariance matrix to obtain an unbiased estimator. This transformation can only be applied under the normality assumption. For the nonparametric framework, we estimate the inverse covariance matrix as the inverse of the sample covariance matrix.

Note that in order to compute the true optimal shrinkage intensity, we need the population moments of asset returns. In our empirical tests in Section 5 we instead use their sample counterparts to estimate the shrinkage intensity, which should also bear some estimation risk. Regardless of the estimation error within the estimated shrinkage intensity, (Tu and Zhou 2011, Table 5) show that this error is small in the context of shrinkage portfolios, and therefore the estimated optimal shrinkage may outperform the sample portfolio. Our empirical results show that it is also the case for shrinkage moments applied in the context of portfolio optimization.

2.6. Nonparametric calibration of the shrinkage estimators of the moments

In this section, we describe an alternative nonparametric bootstrap procedure to estimate optimal shrinkage intensities. We assume that stock returns are iid, but we do not specify any particular distribution. Efron (1979) introduced the bootstrap to study the distributional properties of any statistic of interest. Similarly, we use the bootstrap to approximate the expected value of the squared loss functions for the sample covariance matrix, and the inverse of the sample covariance matrix[8].

This methodology is very intuitive: we generate B bootstrap samples by drawing observations with replacement from the original sample. Then, for each bootstrap sample, we compute our statistic of interest, which in this case corresponds with any of the considered squared loss functions. Finally, we take the sample average among the B bootstrap statistics as an approximation to the expected value.

Contrary to the "simplest" version of bootstrap, we add an error term for each drawn observation. This is what is called *smoothed* bootstrap. We use the multivariate version of the smoothed bootstrap proposed by (Efron 1979, page 7), such that each extracted observation is defined as follows:

$$X_i^* = \mu_{sp} + (I + \Sigma_Z)^{-1/2} \left[X_i - \mu_{sp} + \Sigma_{sp}^{1/2} Z_i\right], \tag{23}$$

where I is the identity matrix, X_i is the i-th row observation from $X \in \mathbb{R}^{T\times N}$, μ_{sp} is the sample vector of means of X, Σ_{sp} is the sample covariance matrix of X, and Z_i is a multivariate random variable having zero vector of means and covariance matrix Σ_Z (in the empirical analysis, we set Z_i as a multivariate normal distribution with zero mean and covariance matrix Σ_{sp}, where Σ_{sp} is the sample covariance matrix). A positive feature of this technique is that X_i^* is a random variable which has mean μ_{sp} and covariance matrix Σ_{sp} under the empirical distribution $\widehat{F}$.

The advantage of using the smoothed bootstrap is that we draw observations from a continuous density function, instead of drawing from the set of sample observations, and in turn, the probability of having repeated observations is zero. The advantage is that in this manner we avoid the singularity in the estimated covariance matrix, which is likely to occur when there are many repeated observations.

Finally, we have also tested other nonparametric methods like the Jackknife or the d-Jackknife[9], but we find that the results are not as good as those from using the smoothed bootstrap, and we do not report the results to conserve space.

3. Shrinkage estimators of portfolio weights

We now focus on shrinkage portfolios defined as a convex combination of a *sample portfolio* and a scaled *target portfolio*:

$$w_{sh} = (1-\alpha)w_{sp} + \alpha\nu w_{tg}, \tag{24}$$

[8] Notice that we do not apply this technique to estimate the shrinkage intensity of the shrinkage vector of means. This is because our proposed technique in Section 2.1 is already a nonparametric technique that makes no assumption on the return distribution.

[9] For a detailed treatment of Jackknife techniques see Efron and Gong (1983) and Efron and Tibshirani (1993). For an application in finance see Basak et al. (2009).

where w_{sp} is the sample estimator of the true optimal portfolio w_{op}, w_{tg} is the target portfolio, α is the shrinkage intensity, and ν is a scale parameter that we adjust to minimize the bias of the target portfolio.

We consider three shrinkage portfolios obtained by shrinking the sample mean-variance portfolio towards the sample minimum-variance portfolio, the sample mean-variance portfolio towards the equally-weighted portfolio, and the sample minimum-variance portfolio towards the equally-weighted portfolio. Variants of these three shrinkage portfolios have been considered before by Kan and Zhou (2007), Tu and Zhou (2011), and DeMiguel et al. (2009), but there are three main differences between our analysis and the analysis in these papers. First, we introduce an additional scaling parameter that we adjust to minimize the bias of the target portfolio. The advantage of introducing this scaling parameter is that by reducing the bias of the target, we also reduce the overall quadratic loss of the resulting shrinkage portfolio, and our empirical results show that, in general, this improves the out-of-sample performance of the shrinkage portfolios. Second, the aforementioned papers consider expected utility maximization as the only calibration criterion for the shrinkage portfolios. Here, we consider (in addition to expected utility maximization) expected quadratic loss minimization, expected portfolio variance minimization, and Sharpe ratio maximization. Third, unlike previous work, we also show how the optimal shrinkage intensity can be estimated using nonparametric techniques.

Mathematically, we define each calibration criterion as follows:

$$\text{Expected quadratic loss (eql):} \qquad \min_{\alpha} \quad E\left(f_{\text{ql}}\left(w_{sh}\right)\right) = \min_{\alpha} \quad E\left(\left\|w_{sh} - w_{op}\right\|_2^2\right), \tag{25}$$

$$\text{Utility (ut):} \qquad \max_{\alpha} \quad E\left(f_{\text{ut}}\left(w_{sh}\right)\right) = \max_{\alpha} \quad E\left(w_{sh}'\mu - \frac{\gamma}{2} w_{sh}'\Sigma w_{sh}\right), \tag{26}$$

$$\text{Variance (var):} \qquad \min_{\alpha} \quad E\left(f_{\text{var}}\left(w_{sh}\right)\right) = \min_{\alpha} \quad E\left(w_{sh}'\Sigma w_{sh}\right), \tag{27}$$

$$\text{Sharpe ratio (SR):} \qquad \max_{\alpha} \quad E\left(f_{\text{SR}}\left(w_{sh}\right)\right) = \max_{\alpha} \quad \frac{E\left(w_{sh}'\mu\right)}{\sqrt{E\left(w_{sh}'\Sigma w_{sh}\right)}}, \tag{28}$$

where γ is the investor's risk aversion parameter. The expected utility and Sharpe ratio maximization criteria match the economic incentives of investors and thus the motivation to use them is straightforward. The expected variance minimization criterion also has an economic rationale because investors are often interested in finding the portfolios that minimize risk, or that minimize the variance of returns with respect to a given benchmark portfolio.[10] We consider the expected quadratic loss minimization criterion for two reasons. First, the expected quadratic loss criterion has been shown to work very well within the context of shrinkage estimators for the covariance matrix; see Ledoit and Wolf (2004b). Thus it is interesting to explore whether it will also result in shrinkage portfolios with good performance. Second, the quadratic loss

[10] In addition, it has been demonstrated in the literature that the estimation error in the mean is so large, that it is often more effective to focus on minimizing the variance of portfolio returns; see, for instance, Jagannathan and Ma (2003).

penalizes big errors over small ones, and this in turn is likely to result in more stable portfolio weights with lower turnover. Nevertheless, in practice the distribution of asset returns may vary with time, which implies that the true optimal portfolio w_{op} may also vary with time, and thus the quadratic loss criterion might fail to provide stable shrinkage portfolios. Presumably, the quadratic loss criterion may be more suitable for shrinkage portfolios that ignore the vector of means, which is more likely to change with time than the covariance matrix.

The remainder of this section is organized as follows. In Section 3.1, we show that the optimal shrinkage intensities for the four different calibration criteria can be rewritten as a function of certain expectations. In Section 3.2, we provide closed-form expressions for these expectations under the assumption that returns are iid normal. In Section 3.3 we discuss how a nonparametric approach can be used to estimate these expectations under the assumption that returns are just iid.

3.1. Characterizing the optimal shrinkage intensity

The following proposition characterizes the optimal shrinkage intensity α for the four calibration criteria. For the expected quadratic loss, utility, and variance criteria, the optimal shrinkage intensity can be written as a function of certain expectations involving the sample portfolio, the target portfolio, and the first and second moments of asset returns. For the Sharpe ratio criterion, the optimal shrinkage intensity is the maximizer to an optimization problem defined in terms of expectations.

PROPOSITION 3. *If asset returns are iid, then the shrinkage intensities for the optimal combination between the sample portfolio and the scaled target portfolio are:*

$$\alpha_{eql} = \frac{E\left(\|w_{sp} - w_{op}\|_2^2\right) - \tau_{sp-tg}}{E\left(\|w_{sp} - w_{op}\|_2^2\right) + E\left(\|\nu w_{tg} - w_{op}\|_2^2\right) - 2\tau_{sp-tg}}, \tag{29}$$

$$\alpha_{ut} = \frac{E\left(\sigma_{sp}^2\right) - \nu E\left(\sigma_{sp,tg}\right) - \frac{1}{\gamma}\left(E(\mu_{sp}) - \nu E(\mu_{tg})\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}, \tag{30}$$

$$\alpha_{var} = \frac{E\left(\sigma_{sp}^2\right) - \nu E\left(\sigma_{sp,tg}\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}, \tag{31}$$

$$\alpha_{SR} = arg\max_{\alpha} \frac{(1-\alpha)E(\mu_{sp}) + \alpha\nu E(\mu_{tg})}{\sqrt{(1-\alpha)^2 E\left(\sigma_{sp}^2\right) + \alpha^2\nu^2 E\left(\sigma_{tg}^2\right) + 2(1-\alpha)\alpha\nu E\left(\sigma_{sp,tg}\right)}}, \tag{32}$$

where $\tau_{sp-tg} = E\left((w_{sp} - w_{op})' (\nu w_{tg} - w_{op})\right)$, $E\left(\sigma_{sp}^2\right) = E\left(w_{sp}' \Sigma w_{sp}\right)$ is the expected sample portfolio variance, $E\left(\sigma_{tg}^2\right) = E\left(w_{tg}' \Sigma w_{tg}\right)$ is the expected target portfolio variance, $E\left(\sigma_{sp,tg}\right) = E\left(w_{sp}' \Sigma w_{tg}\right)$ is the expected covariance between the sample portfolio and the target portfolio, $E\left(\mu_{sp}\right) = E\left(w_{sp}' \mu\right)$ is the expected sample portfolio mean return, and $E\left(\mu_{tg}\right) = E\left(w_{tg}' \mu\right)$ is the expected target portfolio mean return.

A couple of comments are in order. First, note that, roughly speaking, the optimal shrinkage intensity is the ratio of the error of the sample portfolio, in terms of the specific calibration criterion, divided by the total error of the sample portfolio and the scaled target portfolio.

Second, from (30) and (31), we observe that the optimal shrinkage intensities of the utility and variance criteria satisfy:

$$\alpha_{ut} = \alpha_{var} - \frac{\frac{1}{\gamma}\left(E(\mu_{sp}) - \nu E(\mu_{tg})\right)}{E\left(\sigma_{sp}^2\right) + \nu^2 E\left(\sigma_{tg}^2\right) - 2\nu E\left(\sigma_{sp,tg}\right)}.$$

This implies that when the expected return of the sample portfolio is larger than the expected return of the scaled target portfolio ($E(\mu_{sp}) > \nu E(\mu_{tg})$), the utility criterion will result in a smaller shrinkage intensity than the variance criterion. This is likely to occur when the sample portfolio is the mean-variance portfolio—because it is (theoretically) more profitable than the minimum-variance and the equally weighted portfolios. Under these circumstances, the utility criterion will result in more aggressive shrinkage estimators (closer to the sample portfolio) than the variance criterion. This property of the utility criterion may backfire in practice as it is notoriously difficult to estimate mean returns from historical return data. Our empirical results in Section 5 confirm this by showing that, when the sample portfolio is the sample mean-variance portfolio, the variance criterion produces better out-of-sample performance than the utility criterion.

3.2. Parametric calibration

Assuming returns are iid normal, we now give closed-form expressions for the expectations required to compute the optimal shrinkage intensities given in Proposition 3 for the three shrinkage portfolios and four calibration criteria. In Section 4 we exploit these closed-form expressions to improve our understanding of how the impact of estimation error depends on the number of assets N and the number of observations T.

PROPOSITION 4. *Assume returns are independent and normally distributed with mean μ and covariance matrix Σ, and let $T > N + 4$. Assume we use the following unbiased estimator of the inverse covariance matrix $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$, let the sample mean-variance portfolio be $w_{sp}^{mv} = \frac{1}{\gamma}\Sigma_u^{-1}\mu_{sp}$, and the sample minimum-variance portfolio be $w_{sp}^{min} = \Sigma_u^{-1}\iota$. Then, the expectations required to compute the optimal shrinkage intensities are given by the following closed-form expressions:*

The expected quadratic loss of the sample mean-variance portfolio:

$$E\left(\left\|w_{sp}^{mv}-w_{op}^{mv}\right\|_2^2\right)=\frac{a}{\gamma^2}\left[\textit{trace}\left(\Sigma^{-1}\right)\left(\frac{(T-2)}{T}+\mu'\Sigma^{-1}\mu\right)+(T-N-2)\mu'\Sigma^{-2}\mu\right] - \frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu. \tag{33}$$

The expected quadratic loss of the sample minimum-variance portfolio with respect to the true mean-variance portfolio:

$$E\left(\left\|\nu w_{sp}^{min}-w_{op}^{mv}\right\|_2^2\right)=\nu^2 a\left[\textit{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota+(T-N-2)\iota'\Sigma^{-2}\iota\right]+\frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu - 2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \tag{34}$$

The expected quadratic loss of the sample minimum-variance portfolio:

$$E\left(\left\|w_{sp}^{min}-w_{op}^{min}\right\|_2^2\right)=a\left[\textit{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota+(T-N-2)\iota'\Sigma^{-2}\iota\right]-\iota'\Sigma^{-2}\iota. \tag{35}$$

The expected value of the sample mean-variance portfolio variance:

$$E\left(\sigma_{mv}^2\right)=E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right)=\frac{1}{\gamma^2}\left(a(T-2)\left(\frac{N}{T}+\mu'\Sigma^{-1}\mu\right)\right). \tag{36}$$

The expected value of the sample minimum-variance portfolio variance:

$$E\left(\sigma_{min}^2\right)=E\left(w_{sp}^{min\prime}\Sigma w_{sp}^{min}\right)=a(T-2)\iota'\Sigma^{-1}\iota. \tag{37}$$

The expected value of the covariance between the sample mean-variance and sample minimum-variance portfolios:

$$E\left(\sigma_{mv,min}\right)=E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{min}\right)=a(T-2)\frac{1}{\gamma}\mu'\Sigma^{-1}\iota. \tag{38}$$

The term τ_{mv-min}^2:

$$\begin{aligned}\tau_{mv-min}&=E\left(\left(w_{sp}^{mv}-w_{op}^{mv}\right)'\left(\nu w_{sp}^{min}-w_{op}^{mv}\right)\right)\\&=\nu\left(\frac{a}{\gamma}\left[\textit{trace}\left(\Sigma^{-1}\right)\mu'\Sigma^{-1}\iota+(T-N-2)\mu'\Sigma^{-2}\iota\right]-\frac{1}{\gamma}\mu'\Sigma^{-2}\iota\right),\end{aligned} \tag{39}$$

where $a=\frac{(T-N-2)}{(T-N-1)(T-N-4)}$. Moreover, when asset returns are normally distributed, terms τ_{mv-ew} and τ_{min-ew} are equal to zero, $E(w_{sp}^{mv})'\mu=(w_{op}^{mv})'\mu$, and $E(w_{sp}^{min})'\mu=(w_{op}^{min})'\mu$.

Proposition 3 showed that, across every calibration criterion, the shrinkage intensity is higher when the expected quadratic loss or the expected portfolio variance of the sample portfolio are high. Proposition 4 shows that this is likely to occur when the sample covariance matrix is nearly singular. To see this, note that Σ^{-1} and Σ^{-2} appear in the expressions for the quadratic loss and variance of the mean-variance and minimum-variance portfolios, and a nearly singular covariance matrix results in large inverse covariance matrices.

Furthermore, we also see that the expected quadratic loss or the expected portfolio variance of the sample portfolio might be high when we have a low number of observations T compared with the number of assets N. On the other hand, we observe that a small ratio N/T reduces the expected quadratic loss and the portfolio variance. For instance, equation (33) converges to zero when the ratio N/T converges to zero. Also, formula (36) converges to the true variance of the mean-variance portfolio when the ratio N/T converges to zero.

3.3. Nonparametric calibration of portfolios

Although the closed-form expressions derived in the previous subsection are both convenient and insightful, the assumption of iid normal returns may not hold under certain market conditions. To address this issue, we also consider a nonparametric procedure to estimate the optimal shrinkage intensities in a more general setting. Specifically, similar to our analysis in Section 2.6, we implement a nonparametric approach based on a smoothed bootstrap (Efron (1979)) to estimate the expectations required to compute the optimal shrinkage intensities. In addition to the smoothed bootstrap approach, we have also tried other nonparametric methods based on the Jackknife, the d-Jackknife, and the cross-validation (see Efron and Gong (1983) and Efron and Tibshirani (1993)), but the performance of the shrinkage portfolios obtained from these methods is not as good as that from the smoothed bootstrap method and thus we do not report the results to conserve space.

4. Simulation Results

To understand the properties of the shrinkage moments and the shrinkage portfolios, we run a simulation experiment with return data generated by simulating from an iid multivariate normal distribution with sample moments calibrated to those of the 48 industry portfolio dataset from Ken French's website. Under iid normal returns, we can compute the true optimal shrinkage intensities using the closed-form expressions introduced in Sections 2 and 3. We then use simulated data to estimate the Sharpe ratio of the portfolios computed from the shrinkage estimators of moments[11]. Finally, we can compute the Sharpe ratio of the shrinkage portfolios using the closed-form expressions introduced in Section 3.

[11] We simulate 5000 samples of length T, and for each sample we compute any shrinkage moment using the true optimal shrinkage intensity. With the shrinkage moments, we compute the desired portfolios, and we compute the out-of-sample portfolio return and the out-of-sample portfolio variance of each portfolio. We approximate the expected portfolio return and the expected portfolio variance with the sample average among the 5000 generated values. We use the estimated expected portfolio return and the estimated expected portfolio variance to compute the Sharpe ratios.

Figure 1 shows how the true optimal shrinkage intensities for the shrinkage estimators of moments given in Section 2 and the Sharpe ratios of the portfolios obtained from these shrinkage moments change with the number of observations. Panel (a) depicts the shrinkage intensities for the Jorion (1986) vector of means and the shrinkage vector of means proposed in Section 2.1. We observe that the shrinkage intensity of the proposed estimator is larger than that of Jorion's estimator. The shrinkage intensity in both estimators represents a measure of the suitability (inadequacy) of the shrinkage target (sample estimator). Although the shrinkage intensities are obtained under different criteria[12], based on the meaning of shrinkage intensity, we can conclude that the scaled shrinkage target studied in Section 2.1 is more suitable than the shrinkage target of Jorion (1986).

Panel (b) in Figure 1 gives the shrinkage intensities for the covariance and inverse covariance matrices. The panel shows the optimal shrinkage intensities for both matrices for the quadratic loss criterion, together with the optimal shrinkage intensity for the covariance matrix for the criterion that takes into account both the expected quadratic loss and the condition number. Our first observation is that when we take the condition number into account, we obtain a larger shrinkage intensity than when we focus solely on the expected quadratic loss. Our second observation is that the shrinkage intensity for the inverse covariance matrix is quite large, specially in small samples. This is because for small samples, the sample covariance matrix is nearly singular, and thus the impact of estimation error explodes when we invert the sample covariance matrix. Consequently, the expected quadratic loss criterion results in very large shrinkage intensities.

Panel (c) in Figure 1 depicts the Sharpe ratios for the mean-variance portfolios formed with the Jorion (1986) vector of means, and our proposed shrinkage estimator of means. For the data simulated from a multivariate normal distribution, the Sharpe ratio of the mean-variance portfolio with the Jorion (1986) vector of means is larger than the Sharpe ratio of the mean-variance portfolio formed with the proposed shrinkage estimator of means. The reason for this is that the proposed estimator results in a larger shrinkage intensity. In Section 5, we will see that this conservative approach works well, however, when applied to the empirical data due to the instability of the sample vector of means.

Panel (d) in Figure 1 depicts the simulated Sharpe ratios for the minimum-variance portfolios formed with the shrinkage estimators of the covariance matrix and the inverse covariance matrix. We observe that the minimum-variance portfolio formed with the shrinkage covariance matrix that accounts for both the expected quadratic loss and the condition number attains the largest Sharpe ratio. This suggests that even for a sample size of $T = 250$ observations, it is important to take into account the singularity of the sample covariance matrix, and thus the condition number matters to calibrate the shrinkage intensity. On the other

[12] The shrinkage intensity of the Jorion (1986) vector of means is obtained from an empirical-Bayes approach, whereas for our proposed estimator it is chosen to minimize the expected quadratic loss.

hand, we observe that the minimum-variance portfolio formed with the shrinkage inverse covariance matrix attains the lowest Sharpe ratio. As mentioned before, minimizing the expected quadratic loss of the shrinkage inverse covariance matrix gives a very large shrinkage intensity that results in a portfolio too close to the equally-weighted portfolio.

Figure 2 depicts the shrinkage intensities and Sharpe ratios for the shrinkage portfolios calibrated with the methods described in Section 3. Panels (a) and (b) give the shrinkage intensities for the mv-min and mv-ew shrinkage portfolios, obtained by shrinking the sample mean-variance portfolio towards the minimum-variance and equal-weighted portfolios, respectively. For both shrinkage portfolios, the variance minimization criterion provides the largest shrinkage intensity. The reason for this is that, for both portfolios, the shrinkage targets are low-variance portfolios. Because the variance minimization criterion is not a utility-maximizing criterion, the portfolios calibrated with this criterion attain the lowest Sharpe ratios, as shown in panels (d) and (e).

Panel (c) depicts the shrinkage intensities for the min-ew shrinkage portfolio, obtained by shrinking the minimum-variance portfolio towards the equal-weighted portfolio. Since the expected quadratic loss minimization criterion seeks the stability of portfolio weights (see Section 3), and the equally weighted portfolio is a rather stable portfolio, this calibration criterion provides the largest shrinkage intensity. The stability of portfolio weights does not guarantee low portfolio variance and/or high expected return, as this is more dependent on market conditions. Therefore, this calibration criterion provides a slightly lower Sharpe ratio than the other calibration criteria, as we observe in panel (f).

Comparing panels (d), (e), and (f) in Figure 2, we see that among the shrinkage portfolios, the mv-min shrinkage portfolio obtains the worst Sharpe ratio for small samples. This is because this shrinkage portfolio is obtained from the sample mean-variance and sample minimum-variance portfolios, which contain substantial estimation error for small samples. Moreover, we see that there always exists a combination which beats the sample portfolio in terms of Sharpe ratio, although for large sample the difference becomes smaller.

We also study the case where the number of observations is lower than the number of assets[13]. This case is relevant in practice for portfolio managers who deal with a large number of assets, when the number of return observations that are relevant to the prevailing market conditions is small. We assume that returns follow an iid multivariate normal distribution defined with the sample moments of the 48IndP dataset, and we compute simulated Sharpe ratios as in the previous part. Then, we compare the evolution of the Sharpe ratios for the minimum-variance portfolios computed with the shrinkage covariance matrix calibrated by the

[13] For this part of the analysis, we only study minimum-variance portfolios computed from shrinkage covariance matrices, which are not singular.

method that minimizes the expected quadratic loss, and the method that accounts for the expected quadratic loss and the matrix condition number.

Figure 3 depicts the results of the experiment. When the sample size is very small, e.g. $T = 10$, the expected quadratic loss of the covariance matrix is very large. Consequently, the shrinkage intensity that minimizes the expected quadratic loss is high. Therefore, the resulting shrinkage estimator of the covariance matrix has a reasonable condition number. This is why, for very small samples, both shrinkage methods provide similar shrinkage intensities and, in turn, similar Sharpe ratios[14]. On the other hand, when the sample size is bigger than 20 observations, the Sharpe ratios of the minimum-variance portfolios computed with the shrinkage estimators from Sections 2.2 and 2.4 diverge. This is because for larger sample sizes ($T > 20$), the expected quadratic loss of the sample covariance matrix is lower and therefore, the shrinkage intensity that minimizes the quadratic loss is relatively small. As a result, although this shrinkage intensity is sufficient to reduce the quadratic loss, it is not large enough to keep the condition number small. By taking the condition number explicitly into consideration to compute the shrinkage intensity for the estimator proposed in Section 2.4, we therefore can improve the performance of the resulting portfolios for large T.

In general, we observe that the minimum-variance portfolios formed with the shrinkage covariance matrix studied in Section 2.4 with $\phi = 100$, have larger Sharpe ratios than the minimum-variance portfolio formed with the shrinkage covariance matrix studied in Section 2.2, specially for sample sizes larger than 20 observations. Therefore, we conclude that it is always beneficial to account for the matrix condition number, which is specially useful for managers that deal with large number of assets.

Finally, to study the robustness of our results with respect to the number of assets, we have repeated our simulations for the case where the number of observations is fixed to $T = 150$, but the number of assets changes. The robustness check analysis has been made across the five different datasets (5IndP, 10IndP, 38IndP, 48IndP and 100FF) listed in Table 1. We observe from the results, which we do not report to conserve space, that the insights from our experiment are robust to the number of assets.

5. Empirical Results

We now use the six empirical datasets listed in Table 1 to evaluate the out-of-sample performance of the different portfolios listed in Table 2. Section 5.1 describes the performance evaluation methodology. Section 5.2 discusses the main results of the empirical analysis.

Table 1 lists the six datasets considered in the analysis. We consider 4 industry portfolio datasets from Ken French's website. These are portfolios of all stocks from NYSE, AMEX and NASDAQ grouped in

[14] Notice that these results depend on parameter ϕ, which establishes the trade-off between expected quadratic loss and condition number.

terms of their industry. We use datasets with stocks grouped into 5, 10, 38, and 48 industries (5IndP, 10IndP, 38IndP, 48IndP). We also consider a dataset of 100 portfolios formed from stocks sorted by size and book-to-market ratio (100FF), downloaded from Ken French's website. The last dataset (SP100) is formed by 100 stocks, randomly chosen the first month of each new year from the set of assets in the S&P500 for which we have returns for the entire estimation window, as well as for the next twelve months.

Table 2 lists all the portfolios considered. Panel A lists the portfolios from the existing literature that we consider as benchmarks. The first benchmark portfolio is the classical mean-variance portfolio of Markowitz (1952).[15] The second portfolio is the classical mean-variance portfolio composed with the shrinkage vector of means proposed by Jorion (1986). The next three portfolios are mixtures of portfolios proposed in the literature; the first one is the mixture of the mean-variance and minimum-variance portfolio of Kan and Zhou (2007); the second is the mixture of the mean-variance and equally weighted portfolios studied by Tu and Zhou (2011); the third is the mixture of the minimum-variance and equally weighted portfolio of DeMiguel et al. (2009). The sixth portfolio is the minimum-variance portfolio. The seventh portfolio is the minimum-variance portfolio formed with the shrinkage covariance matrix of Ledoit and Wolf (2004b), which shrinks the sample covariance matrix to the identity matrix. The eighth portfolio is the minimum-variance portfolio formed with the shrinkage covariance matrix of Ledoit and Wolf (2003), which shrinks the sample covariance matrix to the sample covariance matrix of a single-index factor model. The ninth portfolio is the equally weighted portfolio. Panel B lists the portfolios constructed with the shrinkage estimators studied in Section 2. The first portfolio in Panel B is the mean-variance portfolio with the shrinkage vector of means proposed in Section 2.1. The second portfolio is the minimum-variance portfolio formed with the shrinkage covariance matrix studied in Section 2.2. The third portfolio is the minimum-variance portfolio formed with the shrinkage inverse covariance matrix proposed in Section 2.3. The fourth portfolio is the minimum-variance portfolio formed with a shrinkage covariance matrix calibrated by accounting for the expected quadratic loss and the condition number. The shrinkage covariance matrices of the last three portfolios are calibrated under the parametric approach, assuming normality, and under the bootstrap nonparametric approach. Panel C lists the shrinkage portfolios proposed in Section 3. In the empirical analysis, we calculate the shrinkage intensities of these portfolios using the four calibration methods defined in Section 3. We compute the shrinkage intensities under a parametric approach, and also under a bootstrap nonparametric approach[16].

[15] For our empirical evaluation, we set the risk aversion coefficient $\gamma = 5$

[16] For the nonparametric approach, we generate B=500 bootstrap samples. We have also tried B=1000 and B=2000 bootstrap samples, but the results are similar to the case of B=500 samples.

5.1. Out-of-sample performance evaluation

We compare the out-of-sample performance of the different portfolios across three different criteria: (i) out-of-sample portfolio Sharpe ratio accounting for transaction costs, (ii) portfolio turnover (trading volume), and (iii) out-of-sample portfolio standard deviation. We use the "rolling-horizon" procedure to compute the out-of-sample performance measures. The "rolling-horizon" is defined as follows: first, we choose a window over which to estimate the portfolio. The length of the window is $M < T$, where T is the total number of observations of the dataset. In the empirical analysis, our estimation window has a length of $M = 150$, which corresponds with 12.5 years of data (with monthly frequency). Second, we compute the various portfolios using the return data over the estimation window. Third, we repeat the "rolling-window" procedure for the next month by including the next data point and dropping the first data point of the estimation window. We continue doing this until the end of the dataset. Therefore, at the end we have a time series of $T - M$ portfolio weight vectors for each of the portfolios considered in the analysis; that is $w_t^i \in \mathbb{R}^N$ for $t = M, \ldots, T-1$ and portfolio i.

The out-of-sample returns are computed by holding the portfolio weights for one month w_t^i and evaluate it with the next-month vector of excess returns: $r_{t+1}^i = R_{t+1}' w_t^i$, where R_{t+1} denotes the vector of excess returns at time $t+1$ and r_{t+1}^i is the out-of-sample portfolio return at time $t+1$ of portfolio i. We use the times series of portfolio returns and portfolio weights of each strategy to compute the out-of-sample standard deviation, Sharpe ratio and turnover:

$$(\sigma^i)^2 = \frac{1}{T-M-1} \sum_{t=M}^{T-1} \left(w_t^{i'} R_{t+1} - \overline{r}^i \right)^2, \tag{40}$$

$$\text{with } \overline{r}^i = \frac{1}{T-M} \sum_{t=M}^{T-1} \left(w_t^{i'} R_{t+1} \right), \tag{41}$$

$$SR^i = \frac{\overline{r}^i}{\sigma^i}, \tag{42}$$

$$\text{Turnover}^i = \frac{1}{T-M-1} \sum_{t=M}^{T-1} \sum_{j=1}^{N} \left(\left| w_{j,t+1}^i - w_{j,t^+}^i \right| \right), \tag{43}$$

where $w_{j,t}^i$ denotes the estimated portfolio weight of asset j at time t under policy i and w_{j,t^+}^i is the estimated portfolio weight of asset j accumulated at time $t+1$, which implies that the turnover is equal to the sum of the absolute value of the rebalancing trades across the N available assets over the $T-M-1$ trading dates, normalized by the total number of trading dates.

To account for transaction costs in the empirical analysis, the definition of portfolio return is slightly corrected by the implied cost of rebalancing the portfolio. Then, the definition of portfolio return, net of proportional transaction costs, is:

$$\underline{r}^i_{t+1} = (1 + R'_{t+1} w^i_t)\left(1 - \kappa \sum_{j=1}^{N} \left| w^i_{j,t+1} - w^i_{j,t^+} \right| \right) - 1, \tag{44}$$

where κ is the chargeable fee for rebalancing the portfolio. In the empirical analysis, expressions (40)-(42) are computed using portfolio returns discounted by transaction costs.

Finally, to measure the statistical significance of the difference between the adjusted Sharpe ratios, we use the stationary bootstrap of Politis and Romano (1994) with B=1000 bootstrap samples and block size b=1[17]. We use the methodology suggested in (Ledoit and Wolf 2008, Remark 2.1) to compute the resulting bootstrap p-values. Furthermore, we also measure the statistical significance of the difference between portfolio variances by computing the bootstrap p-values using the methodology proposed in Ledoit and Wolf (2011).

5.2. Discussion of the out-of-sample performance

Table 3 reports the annualized Sharpe ratio, adjusted by transaction costs, of the benchmark portfolios and the portfolios constructed with the shrinkage estimators studied in Section 2. We consider transaction costs of 50 basis points—Balduzzi and Lynch (1999) argue that 50 basis points is a good estimate of transaction costs for an investor who trades with individual stocks. From Panel A, which reports the Sharpe ratios for the benchmark portfolios, we observe that the minimum-variance portfolio with the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b) (lw) attains the highest out-of-sample Sharpe ratio among all benchmark portfolios. Panel B reports the Sharpe ratio for the portfolios formed with the shrinkage estimators studied in Section 2, calibrated under the assumption of iid normal returns. We observe that the minimum-variance portfolio formed from the shrinkage covariance matrix that accounts for the expected quadratic loss and the condition number (par-clw) outperforms the lw portfolio for medium and large datasets. This is because for medium and large datasets, the sample covariance matrix is more likely to be nearly singular. Furthermore, we can observe that the differences between par-clw and lw are statistically significant for the 38IndP, 48IndP and 100FF datasets. Consequently, for medium and large datasets it is important to use a calibration criterion that explicitly takes into account the condition number of the covariance matrix. Panel C reports the portfolios constructed with the shrinkage estimators studied in Section 2, calibrated without making any assumption about the distribution of stock returns[18]. First, we observe that the mean-variance portfolio obtained from the shrinkage vector of means studied in Section 2.1

[17] We have also computed the p-values when b=5. The interpretation of the results does not change for b=1 or b=5.

[18] For the vector of means, we use the calibration criterion developed in Section 2.1. To estimate the shrinkage covariance matrix and the inverse covariance matrix, we apply the bootstrap procedure described in Section 2.6.

beats the benchmark mean-variance portfolios (mv and bs) across every dataset. We also observe that the nonparametric calibration works better than the parametric approach to calibrate the shrinkage covariance matrix of the minimum-variance portfolios. In general, the nonparametric approach gives larger shrinkage intensities[19], which seems to imply that empirical data departs from the normality assumption, and therefore sample estimators require larger shrinkage intensities than those suggested by the parametric approach.

Table 4 reports the annualized Sharpe ratio, adjusted by transaction costs, of the shrinkage portfolios. Panel A reports the annualized adjusted Sharpe ratios of the shrinkage portfolios calibrated via parametric assumptions (see Section 3.2). Panel B reports the annualized adjusted Sharpe ratios of the shrinkage portfolios calibrated via bootstrap (see Section 3.3). Panel C reports the results of the shrinkage portfolios from the literature. From Panel A, we make two observations. First, the variance minimization criterion is the best calibration criterion for the portfolios that consider the vector of means, mv-min and mv-ew, whereas the expected quadratic loss is the best calibration criterion for the portfolio that does not consider the vector of means, min-ew. This result confirms the intuition about this criterion discussed in Section 3. Moreover, we observe that the best shrinkage portfolio is the min-ew portfolio. The explanation for this is that it is well-known that it is much harder to estimate the mean than the covariance matrix of asset returns from empirical data. Therefore, a mixture of the minimum-variance portfolio with the equally weighted portfolio always outperforms any other combination that considers the vector of means, which would provide more unstable portfolios with lower adjusted Sharpe ratios.

From Panel B of Table 4, we observe again that the best shrinkage portfolio is the mixture formed with the minimum-variance portfolio and the equally weighted portfolio. Furthermore, we also observe that the expected quadratic loss minimization criterion is, in general, the best calibration criterion in terms of Sharpe ratio. The results obtained under the nonparametric bootstrap approach are, in general, slightly better than the results obtained under the assumption of normally distributed returns because empirical returns seem to depart from the normality assumption.

Panel C of Table 4, shows the annualized Sharpe ratio of the existing mixture of portfolios from the literature. We observe that among the mixture of portfolios, the mixture formed by the minimum-variance portfolio and the equally weighted portfolio offers the best results. This mixture, however, performs worse than our studied shrinkage portfolio formed with the minimum-variance portfolio and the equally weighted portfolio across every dataset. Thus, our proposed framework to construct shrinkage portfolios turns out to hedge better the investor's portfolio against estimation error.

[19] We do not report the shrinkage intensities for the sake of brevity. Interested readers might have this information upon request.

Table 5 reports the turnover of the benchmark portfolios and the portfolios constructed with the shrinkage estimators studied in Section 2. Panel A shows the results of the benchmark portfolios. We observe that among all the benchmark portfolios, the naïve equally weighted portfolio attains the lowest turnover across every dataset. Panel B reports the turnover of portfolios formed with the shrinkage estimators studied in Section 2, calibrated under the assumption of normally distributed returns. We observe that the minimum-variance portfolio constructed with the shrinkage inverse covariance matrix of Section 2.3 gives the lowest turnover. The reason for this is that shrinking the inverse of the sample covariance matrix, under an expected quadratic loss criterion, turns out to offer an estimated inverse covariance matrix close to the inverse of the scaled identity matrix. Consequently, the resulting portfolio has a structural form similar to the equally weighted portfolio. Panel C of Table 5 reports the turnover of the portfolios constructed with the shrinkage estimators studied in Section 2, calibrated under a nonparametric approach. We observe that the mean-variance portfolio constructed with the shrinkage vector of means proposed in Section 2.1, provides smaller turnover than the benchmark mean-variance portfolios, mv and bs. Moreover, we observe that minimum-variance portfolios constructed with the shrinkage covariance matrices calibrated via bootstrap, attain smaller turnover than the same portfolios calibrated under the normality assumption. This is because the nonparametric bootstrap approach establishes a higher shrinkage intensity on the shrinkage covariance matrix. Consequently, this approach assigns more weight to the scaled identity matrix than the parametric approach. This issue gives more stable covariance matrices along time, which means that we obtain more stable portfolios.

Table 6 reports the turnover of the shrinkage portfolios. Panel A reports the results of the shrinkage portfolios calibrated under the assumption of normally distributed returns. Panel B reports the turnover of the shrinkage portfolios calibrated under a nonparametric bootstrap approach. Panel C reports the results of the existing mixture of portfolios from the literature. In general, we observe that the best calibration criterion for shrinkage portfolios that consider the vector of means is the variance minimization criterion. For the portfolio that do not consider the vector of means, the min-ew portfolio, the best calibration criterion is the expected quadratic loss minimization criterion. Moreover, we observe that among the shrinkage portfolios, the best shrinkage portfolio is the mixture of portfolios formed with the minimum-variance portfolio and the equally weighted portfolio. In Panel B we have the results for the nonparametric approach. We observe that in general, we obtain better results than with the parametric approach. This is because the nonparametric approach captures the departure from normality. It makes that the scaled target portfolios turn out to be more weighted under the nonparametric approach. Since the target portfolios are chosen to be stable along time, it makes that the nonparametric approach to calibrate shrinkage portfolios provides smaller turnovers than the parametric approach. Panel C shows the turnover of the existing mixture of portfolios studied in the literature. We observe that the mixture formed by the minimum-variance portfolio and the equally weighted portfolio offers the smallest turnover, except for the SP100 dataset.

Tables 7 and 8 report the results for the out-of-sample standard deviation of the studied portfolios. The results obtained from these tables are consistent with the results obtained for the Sharpe ratio and the turnover.

We now summarize the main findings from our empirical analysis. Our first observation is that portfolios computed from our proposed shrinkage vector of means outperform those computed from the Bayes-Stein vector of means of Jorion (1986). Second, we observe that controlling for the condition number of the shrinkage covariance matrix results in portfolio weights that are more stable, and this leads to better adjusted Sharpe ratios for medium and large datasets. Third, for shrinkage portfolios that consider the vector of means, the variance minimization criterion is the most robust criterion, whereas for shrinkage portfolios that do not consider the vector of means, the expected quadratic loss criterion works better. Finally, the studied nonparametric approach to calibrate shrinkage estimators captures the departure from normality in real return data and this results in more stable portfolios (small turnover) with reasonable Sharpe ratios.

6. Conclusions

We provide a comprehensive investigation of shrinkage estimators for portfolio selection. We first study and extend the existing shrinkage estimators of the moments of asset returns. We propose a new class of shrinkage estimator for the vector of means, for which we obtain a closed-form expression of the true optimal shrinkage intensity without making any assumptions on the distribution of stock returns. This new estimator for the vector of means turns out to perform better than the shrinkage vector of means proposed by Jorion (1986). We also propose a novel criterion to calibrate the shrinkage covariance matrix proposed by Ledoit and Wolf (2004b). The proposed criterion accounts for both the expected quadratic loss and the condition number of the covariance matrix, and our empirical results show that the shrinkage estimator based on this criterion results in portfolios with better Sharpe ratio and turnover for medium and large datasets.

For shrinkage portfolios, we have considered three novel calibration criteria (expected quadratic loss, portfolio variance, and portfolio Sharpe ratio) in addition to the expected utility criterion considered in most of the existent literature. Our empirical results show that the variance minimization criterion is the most robust to calibrate shrinkage portfolios that consider the vector of means. On the other hand, for portfolios that ignore the vector of means, the expected quadratic loss minimization criterion is the most robust.

Finally, we have shown that the smoothed bootstrap approach is a practical and simple procedure to calibrate shrinkage estimators, and portfolios computed using this approach perform well in medium and large datasets.

Appendix. Proofs, Tables and Figures

Appendix A. Proof of propositions

In this part, we proof all the propositions. Before going throughout all the propositions, we state two lemmas that will be used along the proofs:

LEMMA 1. *Let x be a random vector in $\mathbb{R}^N$ with mean μ and covariance matrix Σ, and let A be a definite positive matrix in $\mathbb{R}^{N\times N}$. Thus, the expected value of the quadratic form $x'Ax$ is:*

$$E(x'Ax) = trace\,(A\Sigma) + \mu' A\mu. \tag{45}$$

The proof for the expected value of quadratic forms is a standard result in econometrics. See, for instance, (Greene 2003, Page 49).

LEMMA 2. *Given a sample $R \in \mathbb{R}^{T\times N}$ of independent and normally distributed observations, that is $R_t \sim N(\mu, \Sigma)$, the unbiased sample covariance matrix $\Sigma_{sp} = \frac{\sum_{t=1}^{T}\left(R_t - \overline{R}\right)^2}{T-1}$, where $\overline{R} = \frac{\sum_{t=1}^{T} R_t}{T}$, has a Wishart distribution $\Sigma_{sp} \sim \mathcal{W}\left(\frac{\Sigma}{T-1}, T-1\right)$. On the other hand, the unbiased estimator of the inverse covariance matrix $\Sigma_u^{-1} = \frac{T-N-2}{T-1}\Sigma_{sp}^{-1}$ has an inverse-Wishart distribution $\Sigma_u^{-1} \sim \mathcal{W}^{-1}\left((T-N-2)\Sigma^{-1}, T-1\right)$. Then, the expected values of $\Sigma_{sp}\Sigma_{sp}$, Σ_u^{-2} and $\Sigma_u^{-1}\Sigma\Sigma_u^{-1}$ are:*

$$E\left(\Sigma_{sp}\Sigma_{sp}\right) = \frac{T}{T-1}\Sigma^2 + \frac{1}{T-1} trace\,(\Sigma)\,\Sigma. \tag{46}$$

$$E\left(\Sigma_u^{-2}\right) = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(trace\left(\Sigma^{-1}\right)\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right), \tag{47}$$

$$E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right) = \frac{(T-N-2)(T-2)}{(T-N-1)(T-N-4)}\Sigma^{-1}. \tag{48}$$

The proof for $E\left(\Sigma_{sp}\Sigma_{sp}\right)$ can be found in Haff (1979), Theorem 3.1. The proof for $E\left(\Sigma_u^{-2}\right)$ and $E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right)$ are found in Haff (1979), Theorem 3.2.

Proof of Proposition 1

In this section, we prove the closed-form expression given in Proposition 1. In general, we consider that the asset returns are independent and identically distributed. Introducing constraint (5) in problem (4), we have that:

$$\min_{\alpha} \quad E\left[\left\|\mu_{sh} - \mu\right\|_2^2\right] = (1-\alpha)^2 E\left[\left\|\mu_{sp} - \mu\right\|_2^2\right] + \alpha^2 \left\|\nu_\mu \iota - \mu\right\|_2^2. \tag{49}$$

Now, developing the optimality conditions of problem (49), we can obtain the optimal α that minimizes the expected quadratic loss.

$$\alpha_\mu = \frac{E\left(\left\|\mu_{sp} - \mu\right\|_2^2\right)}{E\left(\left\|\mu_{sp} - \mu\right\|_2^2\right) + \left\|\nu_\mu \iota - \mu\right\|_2^2}, \tag{50}$$

where $\nu_\mu = \text{argmin}_\nu \|\nu\iota - \mu\|_2^2 = \overline{\mu}$. We develop the expected value given in (50) to derive the closed-form expressions:

$$E\left(\|\mu_{sp} - \mu\|^2\right) = E\left(\mu_{sp}'\mu_{sp}\right) - \mu'\mu. \tag{51}$$

Since μ_{sp} is a random variable with mean μ and covariance matrix $\frac{\Sigma}{T}$, we can use Lemma 1 to obtain the closed-form expression of $E\left(\|\mu_{sp} - \mu\|^2\right)$. Thus:

$$E\left(\|\mu_{sp} - \mu\|^2\right) = .(N/T)\overline{\sigma^2} \tag{52}$$

where $\overline{\sigma^2} = \text{trace}\,(\Sigma)\,/N$, and it completes the proof.

Proof of Proposition 2

In this section, we prove the closed-form expressions of the expected values considered in Proposition 2. In general, we consider that the vector of asset returns is iid normal. Thus, we can develop the expected values of Proposition 2 and use Lemma 2 to derive the closed-form expressions:

$$E\left(\|\Sigma_{sp} - \Sigma\|_F^2\right) = \text{trace}\left(E\left(\Sigma_{sp}'\Sigma_{sp}\right) - \Sigma'\Sigma\right) = \frac{N}{T-1}\left(\frac{\text{trace}\,(\Sigma^2)}{N} + N\left(\overline{\sigma^2}\right)^2\right) \tag{53}$$

$$E\left(\left\|\Sigma_u^{-1} - \Sigma^{-1}\right\|_F^2\right) = \text{trace}\left(E\left(\Sigma_u^{-2}\right) - \Sigma^{-2}\right) = \text{trace}\,(\Omega) - \text{trace}\left(\Sigma^{-2}\right) \tag{54}$$

$$E\left(< \Sigma_u^{-1} - \Sigma^{-1}, \nu I - \Sigma^{-1} >\right) = \text{trace}\left(E\left(\Sigma_u^{-1} - \Sigma^{-1}\right)'\left(\nu I - \Sigma^{-1}\right)\right) = 0 \tag{55}$$

being $\overline{\sigma^2} = \text{trace}\,(\Sigma)\,/N$ and $\Omega = \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left(\text{trace}\,(\Sigma^{-1})\,\Sigma^{-1} + (T-N-2)\Sigma^{-2}\right)$. It completes the proof.

Proof of Proposition 3

To make the proof of this proposition we simply develop the optimality conditions of the calibration functions defined by the shrinkage portfolio formed with the sample and the target portfolios. The scale parameter is defined as $\nu = \text{argmin}\left\{\|\nu E(w_{tg}) - w_{op}\|_2^2\right\}$ with respect to ν. Developing the optimality conditions, we obtain that the optimal scale factor is $\nu = \frac{E(w_{tg})'w_{op}}{E(w_{tg})'E(w_{tg})}$. First, the expected quadratic loss function of the considered shrinkage portfolio is:

$$\begin{aligned} E\left(\|w_{sh} - w_{op}\|_2^2\right) &= E\left(\|(1-\alpha)\,(w_{sp} - w_{op}) + \alpha\,(\nu w_{tg} - w_{op})\|_2^2\right) = \\ &= (1-\alpha)^2 E\left(\|w_{sp} - w_{op}\|_2^2\right) + \alpha^2 E\left(\|\nu w_{tg} - w_{op}\|_2^2\right) + \\ &+ 2(1-\alpha)\alpha E\left((w_{sp} - w_{op})'\,(\nu w_{tg} - w_{op})\right). \end{aligned} \tag{56}$$

Therefore, developing the optimality conditions of $E\left(\|w_{sh}-w_{op}\|_2^2\right)$, we obtain that the optimal α is:

$$\alpha_{eql}=\frac{E\left(\|w_{sp}-w_{op}\|_2^2\right)-\tau_{sp-tg}}{E\left(\|w_{sp}-w_{op}\|_2^2\right)+E\left(\|\nu w_{tg}-w_{op}\|_2^2\right)-2\tau_{sp-tg}}, \tag{57}$$

where $\tau_{sp-tg}=E\left((w_{sp}-w_{op})'\left(\nu w_{tg}-w_{op}\right)\right)$.

Second, the expected utility function of the shrinkage portfolio is:

$$\begin{aligned}E\left(U\left(w_{sh}\right)\right)=&(1-\alpha)E(w_{sp})'\mu+\alpha\nu E(w_{tg})'\mu-\\&-\frac{\gamma}{2}E\left((1-\alpha)^2w_{sp}'\Sigma w_{sp}+\alpha^2\nu^2w_{tg}'\Sigma w_{tg}+2(1-\alpha)\alpha\nu w_{sp}'\Sigma w_{tg}\right).\end{aligned} \tag{58}$$

Deriving the optimality conditions of the above expression, we obtain the optimal α:

$$\begin{aligned}\alpha_{ut}=&\frac{E\left(w_{sp}'\Sigma w_{sp}\right)-\nu E\left(w_{sp}'\Sigma w_{tg}\right)}{E\left(w_{sp}'\Sigma w_{sp}\right)+\nu^2E\left(w_{tg}'\Sigma w_{tg}\right)-2\nu E\left(w_{sp}'\Sigma w_{tg}\right)}-\\&-\frac{1}{\gamma}\frac{E(w_{sp})'\mu-\nu E(w_{tg})'\mu}{E\left(w_{sp}'\Sigma w_{sp}\right)+\nu^2E\left(w_{tg}'\Sigma w_{tg}\right)-2\nu E\left(w_{sp}'\Sigma w_{tg}\right)}.\end{aligned} \tag{59}$$

The proof of the variance is straightforward. The investor's portfolio variance is defined by the second addend of the utility, given by expression (58). Deriving the optimality conditions of that expression we have that the optimal α is:

$$\alpha_{var}=\frac{E\left(w_{sp}'\Sigma w_{sp}\right)-\nu E\left(w_{sp}'\Sigma w_{tg}\right)}{E\left(w_{sp}'\Sigma w_{sp}\right)+\nu^2E\left(w_{tg}'\Sigma w_{tg}\right)-2\nu E\left(w_{sp}'\Sigma w_{tg}\right)}. \tag{60}$$

Proof of proposition 4

Here, we illustrate how to prove Proposition 4. We develop each element mentioned in the Proposition. First, we show how to obtain $E\left(\left\|w_{sp}^{mv}-w_{op}^{mv}\right\|_2^2\right)$:

$$E\left(\left\|w_{sp}^{mv}-w_{op}^{mv}\right\|_2^2\right)=\frac{1}{\gamma^2}\left(E\left(\mu_{sp}\Sigma_u^{-2}\mu_{sp}\right)-\mu\Sigma^{-2}\mu\right). \tag{61}$$

Due to the fact that returns are assumed to be independent and normally distributed, μ_{sp} and Σ_{sp} are independent. Therefore, we can make use of Lemma 1 and Lemma 2 to compute the expected value of $E\left(\mu_{sp}\Sigma_u^{-2}\mu_{sp}\right)$. Thus, using the independence between μ_{sp} and Σ_{sp}, the expected value of Σ_u^{-2} given in Lemma 2 and the expected value of quadratic forms given in Lemma 1, we have:

$$\begin{aligned}E\left(\left\|w_{sp}^{mv}-w_{op}^{mv}\right\|_2^2\right)=&\frac{1}{\gamma^2}\left[\frac{\text{trace}\left(\Sigma^{-1}\right)(T-N-2)(T-2)}{(T-N-1)(T-N-4)T}+\right.\\&\left.+\frac{(T-N-2)}{(T-N-1)(T-N-4)}\left[\text{trace}\left(\Sigma^{-1}\right)\mu'\Sigma^{-1}\mu+(T-N-2)\mu'\Sigma^{-2}\mu\right]\right]-\\&-\frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu.\end{aligned} \tag{62}$$

The following element is $E\left(\left\|\nu w_{sp}^{min} - w_{op}^{mv}\right\|_2^2\right)$:

$$E\left(\left\|\nu w_{sp}^{min} - w_{op}^{mv}\right\|_2^2\right) = \nu^2 E\left(\iota'\Sigma_u^{-2}\iota\right) + \frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu - 2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \tag{63}$$

Using the value of $E\left(\Sigma_u^{-2}\right)$ given in Lemma 2, we have that:

$$\begin{aligned} E\left(\left\|\nu w_{sp}^{min} - w_{op}^{mv}\right\|_2^2\right) &= \nu^2 \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left[\text{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota + \right. \\ &\left. +(T-N-2)\iota'\Sigma^{-2}\iota\right] + \frac{1}{\gamma^2}\mu'\Sigma^{-2}\mu - 2\frac{\nu}{\gamma}\iota'\Sigma^{-2}\mu. \end{aligned} \tag{64}$$

Now, we prove how to obtain the closed-form expression of $E\left(\left\|w_{sp}^{min} - w_{op}^{min}\right\|_2^2\right)$. First, we expand the expression as usual:

$$E\left(\left\|w_{sp}^{min} - w_{op}^{min}\right\|_2^2\right) = E\left(\iota'\Sigma_u^{-2}\iota\right) - \iota'\Sigma^{-2}\iota. \tag{65}$$

Again, applying the value of $E\left(\Sigma_u^{-2}\right)$ given in Lemma 2, we obtain the following:

$$\begin{aligned} E\left(\left\|w_{sp}^{min} - w_{op}^{min}\right\|_2^2\right) &= \frac{(T-N-2)}{(T-N-1)(T-N-4)}\left[\text{trace}\left(\Sigma^{-1}\right)\iota'\Sigma^{-1}\iota + \right. \\ &\left. +(T-N-2)\iota'\Sigma^{-2}\iota\right] - \iota'\Sigma^{-2}\iota. \end{aligned} \tag{66}$$

The remaining elements are easy to prove. Understanding how to apply Lemma1 and Lemma 2, expressions $E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right)$, $E\left(w_{sp}^{min\prime}\Sigma w_{sp}^{min}\right)$ and $E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{min}\right)$ are simple to obtain. For instance,

$$E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right) = \frac{1}{\gamma^2}E\left(\mu_{sp}\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\mu_{sp}\right). \tag{67}$$

Since μ_{sp} and Σ_{sp} are independent, using Lemma 1 and the expression for $E\left(\Sigma_u^{-1}\Sigma\Sigma_u^{-1}\right)$ given in Lemma 2, we have:

$$E\left(w_{sp}^{mv\prime}\Sigma w_{sp}^{mv}\right) = \frac{1}{\gamma^2}\left(\frac{(T-N-2)(T-2)}{(T-N-1)(T-N-4)}\left(\frac{N}{T} + \mu'\Sigma^{-1}\mu\right)\right). \tag{68}$$

The proof of the remaining elements can be skipped having understood the steps of the previous proofs.

Appendix B. Tables

Table 1 List of Datasets:

This table list the various datasets analyzed, the abbreviation used to identify each dataset, the number of assets N contained in each dataset, the time period spanned by the dataset, and the source of the data. The dataset of CRSP returns (SP100) is constructed in a way similar to Jagannathan and Ma (2003), with monthly rebalancing: in January of each year we randomly select 100 assets as our asset universe for the next 12 months.

[a] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[b] CRSP, The Center for Research in Security Prices

#	Dataset	Abbreviation	N	Time Period	Source
1	5 Industry Portfolios representing the US stock market	5Ind	5	01/1972-06/2009	K. French[a]
2	10 Industry Portfolios representing the US stock market	10Ind	10	01/1972-06/2009	K. French
3	38 Industry Portfolios representing the U.S stock market	38IndP	38	01/1972-06/2009	K. French
4	48 Industry Portfolios representing the U.S. stock market	48Ind	48	01/1972-06/2009	K. French
5	100 Fama and French Portfolios of firms sorted by size and book to market	100FF	100	01/1972-06/2009	K. French
6	100 randomized stocks from S&P 500	SP100	100	01/1988-12/2008	CRSP[b]

Table 2 List of portfolio models:

This table lists the various portfolio strategies considered in the paper. Panel A lists the existing portfolios from the literature. Panel B lists portfolios where the moments are shrunk with the methods proposed in Section 2. Panel C lists the shrinkage portfolio defined in Section 3. The third column gives the abbreviation that we use to refer to each strategy.

#	Policy	Abbreviation
	Panel A: Benchmark portfolios	
1	Classical mean-variance portfolio	mv
2	Bayes-Stein mean-variance portfolio	bs
3	Kan-Zhou's (2007) three-fund portfolio	kz
4	Mixture of mean-variance and equally weighted (Tu and Zhou (2011))	tz
5	Mixture of minimum-variance and equally weighted DeMiguel et.al. (2009))	dm
6	Minimum-Variance portfolio	min
7	Minimum-variance portfolio with Ledoit and Wolf (2004) shrinkage covariance matrix, which shrinks the sample covariance matrix to the identity matrix	lw
8	Minimum-variance portfolio with Ledoit and Wolf (2003) shrinkage covariance matrix, which shrinks the sample covariance matrix to the sample covariance matrix of a single-index factor model	lw-m
9	Equally weighted portfolio	1/N or ew
	Panel B: Portfolios estimated with new calibration procedures to shrink moments	
	Shrinkage mean-variance portfolio	
10	Mean-variance portfolio formed with the shrinkage vector of means defined in Section 2.1	f-mv
	Shrinkage minimum-variance portfolio	
11	Formed with Ledoit and Wolf (2004) shrinkage covariance matrix: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-lw and npar-lw
12	Formed with the shrinkage inverse covariance matrix studied in Section 2.3: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-ilw and npar-ilw
13	Formed with a shrinkage covariance matrix that accounts for the expected quadratic loss and the condition number: calibrated under a parametric calibration assuming normality and calibrated under a bootstrap nonparametric approach	par-clw and npar-clw
	Panel C: Shrinkage portfolios	
14	Mixture of mean-variance and scaled minimum-variance portfolios	mv-min
15	Mixture of mean-variance and scaled equally weighted portfolios	mv-ew
16	Mixture of minimum-variance and scaled equally weighted portfolios	min-ew

Table 3 Annualized Sharpe ratio of benchmark portfolios and portfolios with shrinkage moments (κ =50 basis points):

This table reports the out-of-sample annualized Sharpe ratio of benchmark portfolios and portfolios constructed by using the shrinkage estimators studied in Section 2. We adjust the Sharpe ratio with transaction costs, where we assume that the chargeable fee is equivalent to 50 basis points. We consider an investor with a risk aversion level of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.593**	0.599***	0.004***	-0.003***	-0.463***	-0.102*
bs	0.733*	0.817*	0.250***	0.164***	-0.343***	0.113
Portfolios that do not consider the vector of means						
min	0.841	0.945	0.528***	0.378***	-0.014***	0.399
lw	0.863	0.955	0.731	0.651	1.003	0.687
lw-m	0.836	0.953	0.649***	0.607	0.843***	0.648
Naïve Portfolios						
1/N	0.761	0.780	0.695	0.688	0.712	0.328
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.845	0.945	0.643***	0.553***	0.762***	0.641
par-ilw	0.877	0.907	0.716	0.693	0.713	0.337
par-clw	0.853	0.948	0.824***	0.794***	1.164***	0.700
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.809	0.881	0.403***	0.270***	0.018***	0.257
Portfolios that do not consider the vector of means						
npar-lw	0.860	0.954	0.711***	0.622***	0.929***	0.667
npar-ilw	0.848	0.844	0.701	0.690	0.712	0.328
npar-clw	0.863	0.954	0.858**	0.825***	1.152*	0.702

Table 4 Annualized Sharpe ratio with transaction costs of shrinkage portfolios (κ =50 basis points):

This table reports the out-of-sample annualized Sharpe ratio of shrinkage portfolios studied in Section 3. We adjust the Sharpe ratio with transaction costs, where we assume that the chargeable fee is equivalent to 50 basis points. We consider an investor with a risk aversion level of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolio with parametric calibration						
EQL Minimization						
mv-min	0.702**	0.812**	0.313***	0.274**	-0.198***	0.307
mv-ew	0.647**	0.760*	0.483	0.437	0.028***	0.188*
min-ew	0.844	0.954	0.658	0.587	0.435***	0.509
Utility Maximization						
mv-min	0.708*	0.790*	0.303***	0.305**	-0.025***	0.388
mv-ew	0.691*	0.738*	0.386**	0.436	0.205***	0.234
min-ew	0.847	0.949	0.593***	0.472***	0.165***	0.483
Variance Minimization						
mv-min	0.820	0.925	0.527***	0.374***	-0.054***	0.414
mv-ew	0.773	0.796*	0.594	0.670	0.425***	0.277
min-ew	0.847	0.949	0.593***	0.472***	0.165***	0.483
Sharpe Ratio Maximization						
mv-min	0.714*	0.807**	0.284***	0.208**	-0.208***	0.273
mv-ew	0.685**	0.715**	0.249***	0.254**	-0.188***	0.111*
min-ew	0.817	0.955	0.606**	0.508***	0.104***	0.472
Panel B: Shrinkage portfolios with bootstrap calibration						
EQL Minimization						
mv-min	0.708**	0.795**	0.277***	-0.356***	-0.612***	0.398
mv-ew	0.657**	0.769*	0.575	0.596	0.713	0.329
min-ew	0.853	0.952	0.711	0.675	0.712	0.330
Utility Maximization						
mv-min	0.714*	0.771**	0.264***	0.158**	-0.471***	0.399
mv-ew	0.701**	0.750**	0.514	0.615	0.712*	0.328
min-ew	0.844	0.945	0.688	0.662	0.712	0.328
Variance Minimization						
mv-min	0.821	0.931	0.528***	0.374***	-0.093***	0.399
mv-ew	0.757	0.798	0.641	0.685	0.712	0.328
min-ew	0.855	0.948	0.684	0.660	0.712	0.328
Sharpe Ratio Maximization						
mv-min	0.702**	0.793**	0.261***	0.050***	-0.471***	0.153*
mv-ew	0.689**	0.714**	0.308***	0.401	0.714	0.331
min-ew	0.824	0.955	0.627	0.590	0.713	0.335
Panel C: Existing mixture of portfolios						
kz	0.714*	0.807**	0.284***	0.208**	-0.208***	0.273
tz	0.673**	0.704**	0.301***	0.360*	0.083***	0.195
dm	0.813	0.941	0.603**	0.508***	0.104***	0.472

Table 5 Turnover of benchmark portfolios and portfolios estimated with shrinkage moments:

This table reports the out-of-sample Turnover of benchmark portfolios and portfolios constructed by using the shrinkage estimators studied in Section 2. We consider an investor with a risk aversion level of $\gamma = 5$.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.287	0.369	1.336	2.311	8.791	2.075
bs	0.152	0.201	0.762	1.300	6.905	1.506
Portfolios that do not consider the vector of means						
min	0.081	0.124	0.395	0.567	2.775	1.156
lw	0.056	0.089	0.240	0.321	0.793	0.279
lw-m	0.080	0.112	0.266	0.320	0.952	0.245
Naïve Portfolios						
1/N	0.018	0.025	0.032	0.033	0.023	0.054
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.065	0.102	0.297	0.402	1.224	0.350
par-ilw	0.044	0.051	0.042	0.039	0.024	0.058
par-clw	0.054	0.079	0.173	0.216	0.361	0.210
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.096	0.149	0.490	0.805	3.679	1.270
Portfolios that do not consider the vector of means						
npar-lw	0.066	0.094	0.256	0.346	0.936	0.309
npar-ilw	0.041	0.036	0.033	0.034	0.023	0.054
npar-clw	0.053	0.072	0.155	0.192	0.308	0.194

Table 6 Turnover of shrinkage portfolios:

This table reports the out-of-sample Turnover of the shrinkage portfolios studied in Section 3. We consider an investor with a risk aversion level of $\gamma = 5$.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A:Shrinkage portfolios with parametric calibration						
EQL Minimization						
mv-min	0.173	0.228	0.716	1.004	4.772	1.260
mv-ew	0.177	0.173	0.516	0.726	3.416	0.621
min-ew	0.074	0.103	0.279	0.350	1.389	0.580
Utility Maximization						
mv-min	0.197	0.241	0.694	0.945	3.454	1.177
mv-ew	0.155	0.200	0.592	0.821	2.944	0.636
min-ew	0.072	0.112	0.340	0.474	2.165	0.822
Variance Minimization						
mv-min	0.119	0.132	0.398	0.571	2.965	1.178
mv-ew	0.076	0.115	0.279	0.356	1.970	0.517
min-ew	0.072	0.112	0.340	0.474	2.165	0.822
Sharpe Ratio Maximization						
mv-min	0.167	0.208	0.709	1.130	5.291	1.265
mv-ew	0.163	0.222	0.786	1.213	5.495	0.991
min-ew	0.109	0.113	0.308	0.424	2.334	0.847
Panel B: Shrinkage portfolios with bootstrap calibration						
EQL Minimization						
mv-min	0.181	0.250	0.786	4.119	7.627	1.157
mv-ew	0.171	0.159	0.371	0.399	0.023	0.054
min-ew	0.071	0.093	0.192	0.181	0.023	0.054
Utility Maximization						
mv-min	0.211	0.263	0.776	10.763	4.236	1.156
mv-ew	0.143	0.182	0.411	0.408	0.023	0.054
min-ew	0.093	0.097	0.228	0.222	0.023	0.054
Variance Minimization						
mv-min	0.117	0.135	0.395	0.589	3.003	1.156
mv-ew	0.098	0.107	0.181	0.146	0.023	0.054
min-ew	0.090	0.097	0.230	0.221	0.023	0.054
Sharpe Ratio Maximization						
mv-min	0.178	0.224	0.758	1.510	7.953	1.437
mv-ew	0.159	0.222	0.694	0.868	0.025	0.055
min-ew	0.102	0.111	0.275	0.317	0.024	0.056
Panel C: Existing mixture of portfolios						
kz	0.167	0.208	0.709	1.130	5.291	1.265
tz	0.175	0.242	0.717	1.009	3.626	0.708
dm	0.124	0.120	0.310	0.424	2.334	0.847

Table 7 Standard deviation of benchmark portfolios and portfolios with shrinkage moments:

This table reports the out-of-sample standard deviation of benchmark portfolios and portfolios constructed by using the shrinkage estimators studied in Section 2. We consider an investor with a risk aversion level of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Benchmark Portfolios						
Portfolios that consider the vector of means						
mv	0.161^{***}	0.157^{***}	0.244^{***}	0.336^{***}	0.417^{***}	0.267^{***}
bs	0.143^{**}	0.134^{***}	0.167^{***}	0.224^{***}	0.346^{***}	0.200^{***}
Portfolios that do not consider the vector of means						
min	0.138	0.126^{**}	0.131^{***}	0.137^{***}	0.179^{***}	0.171^{***}
lw	0.136	0.124	0.120	0.124	0.125	0.122
lw-m	0.138	0.126	0.121	0.123	0.132^{***}	0.120
Naïve Portfolios						
1/N	0.154^{***}	0.148^{***}	0.166^{***}	0.165^{***}	0.174^{***}	0.169^{***}
Panel B: Portfolios calibrated parametrically						
Portfolios that do not consider the vector of means						
par-lw	0.136	0.125	0.124^{***}	0.128^{***}	0.137^{***}	0.125^{**}
par-ilw	0.138^{**}	0.132^{*}	0.159^{***}	0.161^{***}	0.174^{***}	0.167^{***}
par-clw	0.136	0.124	0.119	0.122	0.122	0.121
Panel C: Portfolios calibrated nonparametrically						
Portfolios that consider the vector of means						
f-mv	0.138	0.128^{**}	0.139^{***}	0.169^{***}	0.204^{***}	0.178^{***}
Portfolios that do not consider the vector of means						
npar-lw	0.136	0.124	0.121^{***}	0.125^{***}	0.128^{***}	0.123^{*}
npar-ilw	0.143^{***}	0.140^{***}	0.164^{***}	0.164^{***}	0.174^{***}	0.169^{***}
npar-clw	0.137	0.124	0.119	0.122	0.123	0.121

Table 8 Standard deviation of shrinkage portfolios:

This table reports the out-of-sample standard deviation of the shrinkage portfolios studied in Section 3. We consider an investor with a risk aversion level of $\gamma = 5$. One, two and three asterisks indicate that the difference with the lw portfolio is statistically different from zero for a 90%, 95% and 99% confidence interval, respectively.

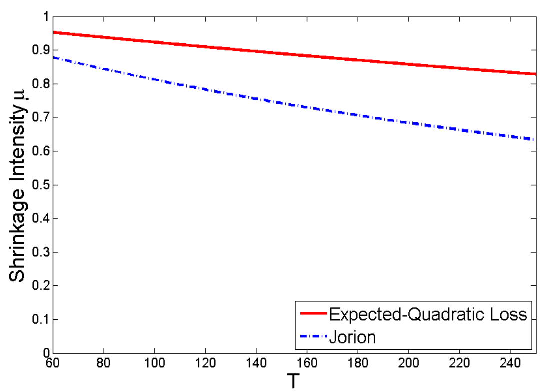
Policy	5IndP	10IndP	38IndP	48IndP	100FF	SP100
Panel A: Shrinkage portfolios with parametric calibration						
EQL Minimization						
mv-min	0.145**	0.137***	0.159***	0.183***	0.251***	0.173***
mv-ew	0.152***	0.140***	0.157***	0.177***	0.216***	0.149*
min-ew	0.138	0.125	0.126**	0.130**	0.146***	0.134**
Utility Maximization						
mv-min	0.145***	0.137***	0.158***	0.177***	0.195***	0.170***
mv-ew	0.149***	0.140***	0.165***	0.186***	0.198***	0.147*
min-ew	0.137	0.125	0.127***	0.132***	0.160***	0.142***
Variance Minimization						
mv-min	0.141**	0.127*	0.131***	0.137***	0.184***	0.172***
mv-ew	0.146***	0.138***	0.155***	0.162***	0.174***	0.145
min-ew	0.137	0.125	0.127***	0.132***	0.160***	0.142***
Sharpe Ratio Maximization						
mv-min	0.144***	0.135***	0.161***	0.203***	0.277***	0.176***
mv-ew	0.150***	0.142***	0.181***	0.228***	0.297***	0.162***
min-ew	0.136	0.124	0.126**	0.132***	0.166***	0.142***
Panel B: Shrinkage portfolios with nonparametric calibration						
EQL Minimization						
mv-min	0.144***	0.137***	0.160***	0.335***	0.264***	0.171***
mv-ew	0.151***	0.140***	0.153***	0.160***	0.174***	0.168***
min-ew	0.137	0.125	0.132**	0.142***	0.174***	0.168***
Utility Maximization						
mv-min	0.144**	0.137***	0.161***	0.971***	0.183***	0.171***
mv-ew	0.149***	0.140***	0.155***	0.161***	0.174***	0.169***
min-ew	0.138***	0.125	0.129**	0.140***	0.174***	0.169***
Variance Minimization						
mv-min	0.140*	0.127**	0.131***	0.138***	0.181***	0.171***
mv-ew	0.148***	0.140***	0.159***	0.162***	0.174***	0.169***
min-ew	0.138**	0.125	0.129**	0.140***	0.174***	0.169***
Sharpe Ratio Maximization						
mv-min	0.145***	0.135***	0.166***	0.231***	0.373***	0.189***
mv-ew	0.149***	0.142***	0.171***	0.193***	0.174***	0.168***
min-ew	0.137	0.124	0.127**	0.135**	0.174***	0.167***
Panel C: Existing mixture of portfolios						
kz	0.144***	0.135***	0.161***	0.203***	0.277***	0.176***
tz	0.150***	0.143***	0.174***	0.202***	0.222***	0.149**
dm	0.136	0.125	0.126**	0.132***	0.166***	0.142***

Appendix C. Figures

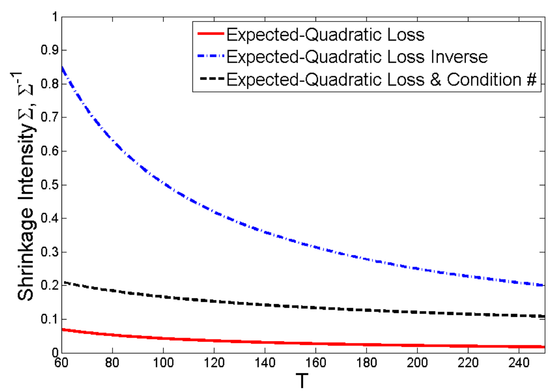
Figure 1 Shrinkage intensities and Sharpe ratios of portfolios computed with shrinkage moments

These plots show the evolution of the true optimal shrinkage parameters for the shrinkage estimators of μ, Σ and Σ^{-1}, as well as the Sharpe ratios of portfolios formed with the shrinkage moments. Plot (a) depicts the evolution of the shrinkage intensities for the vector of means of our studied shrinkage mean vector of returns (solid line) and the Jorion (1986) mean vector of returns (dot-dashed line). Plot (b) depicts the shrinkage intensities of the shrinkage covariance studied in Section 2.2 (solid line), the shrinkage inverse covariance matrix studied in Section 2.3 (dot-dashed line), and the shrinkage covariance matrix studied in Section 2.4 (dashed line). Plot (c) depicts the simulated Sharpe ratios of the mean-variance portfolios constructed with our studied shrinkage mean vector of returns (solid line), and the Jorion (1986) mean vector of returns (dot-dashed line). Plot (d) depicts the simulated Sharpe ratios of the minimum-variance portfolios constructed with the shrinkage covariance matrix studied in Section 2.2 (solid line), the shrinkage inverse covariance matrix studied in Section 2.3 (dot-dashed line), and the shrinkage covariance matrix studied in Section 2.4 (dashed line). To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution. The shrinkage estimator for the covariance matrix accounting for its expected quadratic loss and its condition number establishes $\phi = 100$. The experiment is made considering an investor with a risk aversion level of $\gamma = 10$.

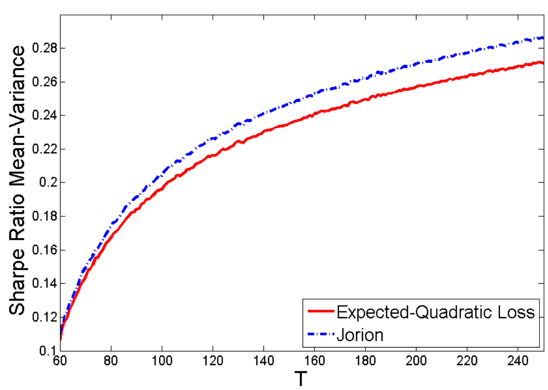
(a) Shrinkage intensity of μ_{sh}



(b) Shrinkage intensity of Σ_{sh} and Σ_{sh}^{-1}



(c) Sharpe ratio of w_{sp}^{mv} constructed with μ_{sh}



(d) Sharpe ratio of w_{sp}^{min} constructed with Σ_{sh} and Σ_{sh}^{-1}

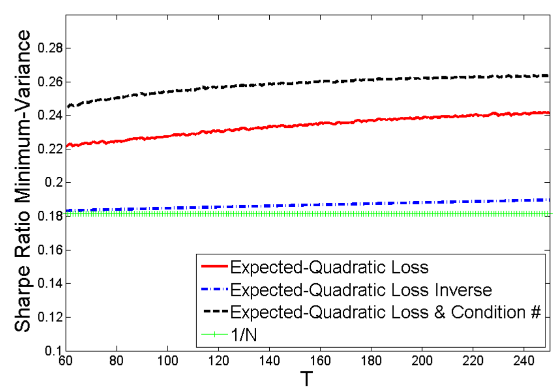
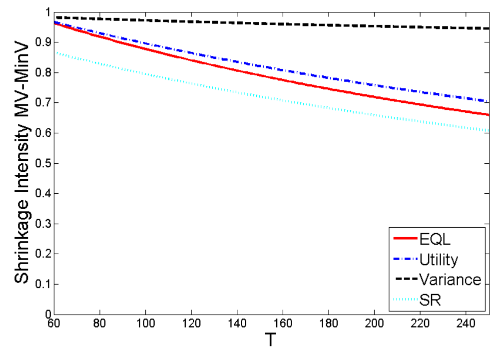


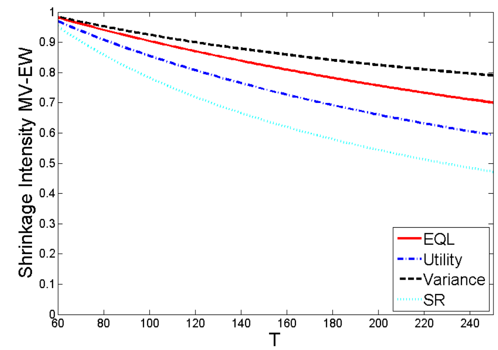
Figure 2 Shrinkage intensities and Sharpe ratios of shrinkage portfolios

These plots show the evolution of the true optimal shrinkage parameters for the shrinkage portfolios studied in Section 3, as well as their Sharpe ratios. For each shrinkage portfolio, we compute the corresponding considered value (the shrinkage intensity or the Sharpe ratio) under every calibration criterion, where EQL, Utility, Variance and SR stand for the expected quadratic loss minimization criterion (solid line), utility maximization criterion (dot-dashed line), variance minimization criterion (dashed line), and Sharpe ratio maximization criterion (dotted line), respectively. On the other hand, Mv and Min, in plots (d)-(f), stand for the Sharpe ratios of the sample mean-variance portfolio and the sample minimum-variance portfolio, respectively. To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution. The experiment is made considering an investor with a risk aversion level of $\gamma = 10$.

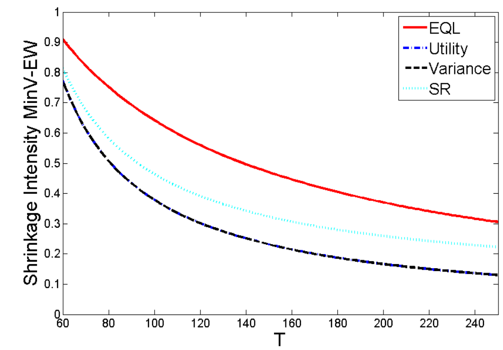
(a) Shrinkage parameter evolution for mv-min



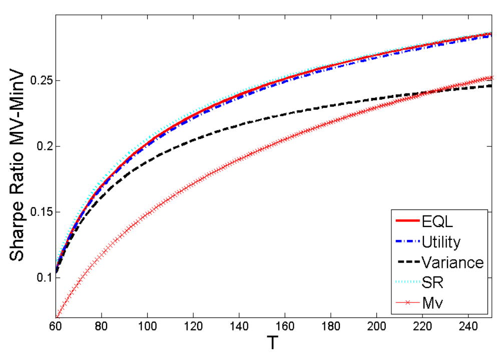
(b) Shrinkage parameter evolution for mv-ew



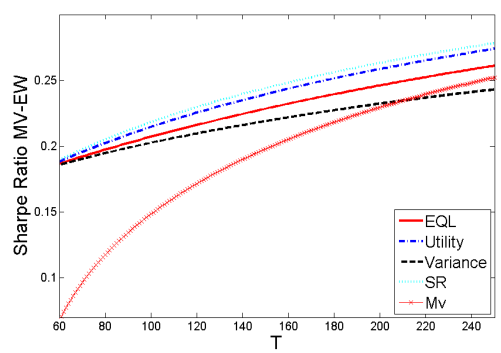
(c) Shrinkage parameter evolution for min-ew



(d) Sharpe ratio of mv-min



(e) Sharpe ratio of mv-ew



(f) Sharpe ratio of min-ew

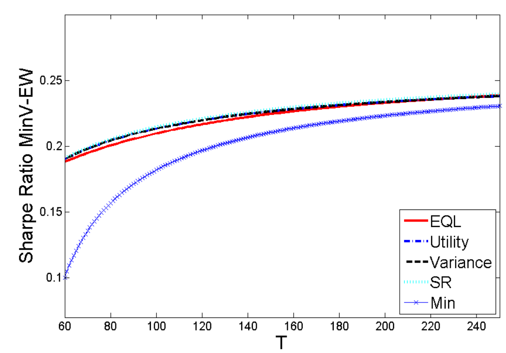
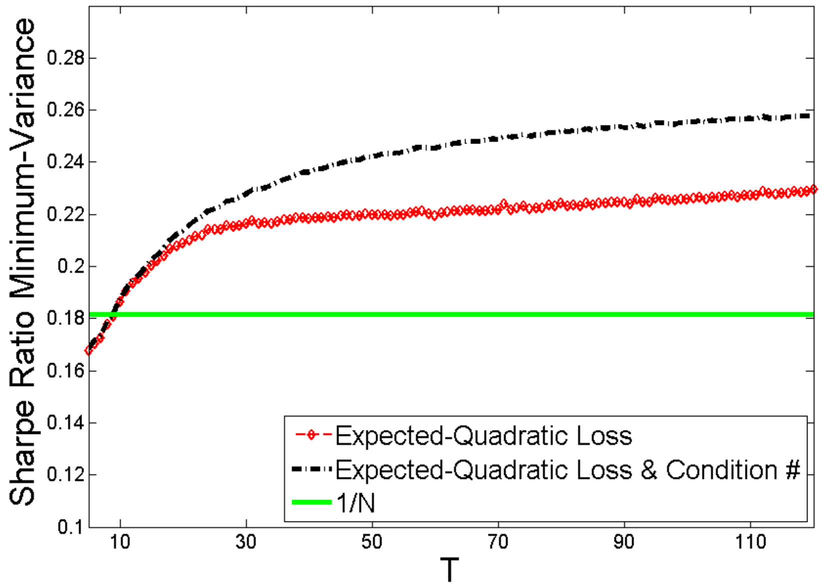


Figure 3 Sharpe ratios of portfolios formed with shrinkage covariance matrices

This plot shows the evolution of the Sharpe ratios for the minimum-variance portfolios composed with the shrinkage covariance matrix studied in Section 2.2 (dashed line with rhombus) and the shrinkage covariance matrix studied in Section 2.4 (dot-dashed line). For the sake of comparison, we also plot the results of the equally weighted portfolio (solid line). To compute the shrinkage covariance matrix studied in Section 2.4, we use $\phi = 100$, as in the previous simulations. To carry out the simulation, we use the sample moments of a dataset formed by 48 industry portfolios (48IndP) as the population moments of a multivariate normal distribution.



Acknowledgments

This work is supported by the Spanish Government through the project MTM2010-16519.

References

Balduzzi, P. and A. W. Lynch (1999). Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics 52*(1), 47–78.

Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *The Journal of Finance 29*(2), 515–522.

Basak, G. K., R. Jagannathan, and T. Ma (2009). Jackknife estimator for tracking error variance of optimal portfolios. *Management Science 55*(6), 990–1002.

Bawa, V. S., S. J. Brown, and R. W. Klein (1979). *Estimation Risk and Optimal Portfolio Choice*. North-Holland Pub. Co. (Amsterdam and New York and New York).

Best, M. J. and R. R. Grauer (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies 4*, 315–342.

Best, M. J. and R. R. Grauer (1992). Positively weighted minimum-variance portfolios and the structure of asset expected returns. *The Journal of Financial and Quantitative Analysis 27*(4), 513–537.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance 54*(2), 655–671.

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research 45*, 21–58.

Cornuejols, G. and R. Tütüncü (2007). *Optimization Methods in Finance*. Cambridge University Press.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science 55*, 798–812.

DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies 22*(5), 1915–1953.

DeMiguel, V. and F. J. Nogales (2009). Portfolio selection with robust estimation. *Operations Research 57*, 560–577.

Dey, D. K. and C. Srinivasan (1985). Estimation of a covariance matrix under Stein's loss. *The Annals of Statistics 13*(4), 1581–1591.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics 7*(1), 1–26.

Efron, B. and G. Gong (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician 37*(1), 36–48.

Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap*. Chapman & Hall.

Frahm, G. and C. Memmel (2010). Dominating estimators for minimum-variance portfolios. *Journal of Econometrics 159*(2), 289–302.

Frost, P. A. and J. E. Savarino (1986). An empirical bayes approach to efficient portfolio selection. *The Journal of Financial and Quantitative Analysis 21*(3), 293–305.

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

Goldfarb, D. and G. Iyengar (2003). Robust portfolio selection problems. *Mathematics of Operations Research 28*(1), 1–38.

Greene, W. H. (2003). *Econometrics Analysis, Fifth Edition*. Prentice Hall.

Haff, L. R. (1979). An identity for the wishart distribution with applications. *Journal of Multivariate Analysis 9*, 531–544.

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

James, W. and J. Stein (1961). Estimation with quadratic loss. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 361–379.

Jobson, J., B. Korkie, and V. Ratti (1979). Improved estimation for Markowitz portfolios using James-Stein type estimators. *Proceedings of the American Statistical Association*, 279–284.

Jobson, J. D. and B. Korkie (1981). Putting Markowitz theory to work. *Journal of Portfolio Management 7*, 70–74.

Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business 58*, 259–278.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis 21*(3), 279–292.

Kan, R. and G. Zhou (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis 42*, 621–656.

Kourstis, A., G. Dotsis, and R. N. Markellos (2011). Parameter uncertainty in portfolio selection: shrinking the inverse covariance matrix. *SSRN, eLibrary*.

Ledoit, O. and M. Wolf (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*, 603–621.

Ledoit, O. and M. Wolf (2004a). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management 30*, 110–119.

Ledoit, O. and M. Wolf (2004b). A well-conditioned estimaor for large-dimensional covariance matrices. *Journal of Multivariate Analysis 88*, 365–411.

Ledoit, O. and M. Wolf (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance 15*, 850–859.

Ledoit, O. and M. Wolf (2011). Robust performance hypothesis testing with the variance. *Wilmott Magazine forthcoming*.

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

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Pastor, L. (2000). Portfolio selection and asset pricing models. *The Journal of Finance 55*, 179–223.

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*(428), pp. 1303–1313.

Pástor, L. and R. F. Stambaugh (2000). Comparing asset pricing models: An investment perspective. *Journal of Financial Economics 56*, 335–381.

Rustem, B., R. G. Becker, and W. Marty (2000). Robust min-max portfolio strategies for rival forecast and risk scenarios. *Journal of Economic Dynamics and Control 24*, 1591–1621.

Stein, C. (1975). Estimation of a covariance matrix. *Rietz Lecture, 39th Annual Meeting IMS, Atlanta, GA*.

Tu, J. and G. Zhou (2011). Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics 99*(1), 204 – 215.

Tutuncu, R. H. and M. Koenig (2004). Robust asset allocation. *Annals of Operations Resarch 132*, 157–187.

Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies 18*(2), 673–705.

Zumbach, G. (2009). Inference on multivariate arch processes with large sizes. *SSRN, eLibrary*.