

A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms

Victor DeMiguel

London Business School, London NW1 4SA, United Kingdom, avmiguel@london.edu

Lorenzo Garlappi

McCombs School of Business, University of Texas at Austin, Austin, Texas 78712,
lorenzo.garlappi@mcombs.utexas.edu

Francisco J. Nogales

Universidad Carlos III de Madrid, 28911 Leganes, Madrid, Spain, fcojavier.nogales@uc3m.es

Raman Uppal

London Business School, London NW1 4SA, United Kingdom, ruppall@london.edu

We provide a general framework for finding portfolios that perform well out-of-sample in the presence of estimation error. This framework relies on solving the traditional minimum-variance problem but subject to the additional constraint that the norm of the portfolio-weight vector be smaller than a given threshold. We show that our framework nests as special cases the shrinkage approaches of Jagannathan and Ma (Jagannathan, R., T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance* **58** 1651–1684) and Ledoit and Wolf (Ledoit, O., M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance* **10** 603–621, and Ledoit, O., M. Wolf. 2004. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* **88** 365–411) and the 1/N portfolio studied in DeMiguel et al. (DeMiguel, V., L. Garlappi, R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financial Stud.* **22** 1915–1953). We also use our framework to propose several new portfolio strategies. For the proposed portfolios, we provide a moment-shrinkage interpretation and a Bayesian interpretation where the investor has a prior belief on *portfolio weights* rather than on *moments* of asset returns. Finally, we compare empirically the out-of-sample performance of the new portfolios we propose to 10 strategies in the literature across five data sets. We find that the norm-constrained portfolios often have a higher Sharpe ratio than the portfolio strategies in Jagannathan and Ma (2003), Ledoit and Wolf (2003, 2004), the 1/N portfolio, and other strategies in the literature, such as factor portfolios.

Key words: portfolio choice; covariance matrix estimation; estimation error; shrinkage estimator; norm constraints

History: Received July 14, 2007; accepted November 28, 2008, by David A. Hsieh, finance. Published online in *Articles in Advance* March 6, 2009.

1. Introduction

Markowitz (1952) showed that an investor who cares only about the mean and variance of static portfolio returns should hold a portfolio on the efficient frontier. To implement these portfolios in practice, one needs to estimate the means and covariances of asset returns. Traditionally, the sample means and covariances have been used for this purpose. But due to estimation error, the portfolios that rely on the sample estimates typically perform poorly out of sample.¹ In this paper, we provide a general frame-

work for determining portfolios with superior out-of-sample performance in the presence of estimation error. This general framework relies on solving the traditional minimum-variance problem (based on the sample covariance matrix) but subject to the additional constraint that the norm of the portfolio-weight vector be smaller than a given threshold.

It is well known that it is more difficult to estimate means than covariances of asset returns (see Merton 1980) and also that errors in estimates of means have a larger impact on portfolio weights than errors in estimates of covariances. For this reason, recent academic research has focused on minimum-variance portfolios, which rely solely on estimates of covariances and thus are less vulnerable to estimation error than

¹ For evidence of the poor performance of the Markowitz portfolio based on sample estimates of means and covariances, see Frost and Savarino (1986, 1988), Michaud (1989), Best and Grauer (1991), Chopra and Ziemba (1993), Broadie (1993), and Litterman (2003).

mean-variance portfolios.² Indeed, extensive empirical evidence shows that the minimum-variance portfolio usually performs better out of sample than any other mean-variance portfolio—even when using a performance measure that depends on *both* the portfolio mean and variance. For example, Jagannathan and Ma (2003, pp. 1652–1653) report:³

The estimation error in the sample mean is so large nothing much is lost in ignoring the mean altogether when no further information about the population mean is available. For example, the global minimum variance portfolio has as large an out-of-sample Sharpe ratio as other efficient portfolios when past historical average returns are used as proxies for expected returns. In view of this we focus our attention on global minimum variance portfolios in this study.

Just like Jagannathan and Ma (2003), we too focus on minimum-variance portfolios, even though the general framework we develop applies also to mean-variance portfolios. But even the performance of the minimum-variance portfolio depends crucially on the quality of the estimated covariances, and although the estimation error associated with the sample covariances is smaller than that for sample mean returns, it can still be substantial.

In the literature, several approaches have been proposed to deal with the problem of estimating the large number of elements in the covariance matrix. One approach is to use higher frequency data, say, daily instead of monthly returns (see Jagannathan and Ma 2003). A second approach is to impose some factor structure on the estimator of the covariance matrix (Chan et al. 1999, Green and Hollifield 1992), which reduces the number of parameters to be estimated. A third approach has been proposed by Ledoit and Wolf (2003, 2004), who use as an estimator a weighted average of the sample covariance matrix and another estimator, such as the 1-factor covariance matrix or the identity matrix. A fourth approach, which is often used in practice, is to impose shortsale constraints on the portfolio weights (see Frost and Savarino 1988, Chopra 1993). Jagannathan and Ma (2003) show that imposing a shortsale constraint when minimizing the portfolio variance is equivalent to shrinking the extreme elements of the covariance matrix. This simple remedy for dealing with estimation error performs very well. In fact, Jagannathan and Ma (2003) find that “the sample covariance matrix [with shortsale constraints] performs almost as well as those

[covariance matrices] constructed using factor models, shrinkage estimators or daily returns” (p. 1654). Finally, DeMiguel et al. (2009) demonstrate that even constraining shortsales may not mitigate completely the error in estimating the covariance matrix, and thus an investor may be better off (in terms of Sharpe ratio, certainty-equivalent returns, and turnover) ignoring data on asset returns altogether and using the naive $1/N$ rule to allocate an equal proportion of wealth across each of the N assets.

In this paper, we develop a new approach for determining the optimal portfolio weights in the presence of estimation error. Following Brandt (1999) and Britten-Jones (1999), we treat the *weights* rather than the *moments* of asset returns as the objects of interest to be estimated. So rather than shrinking the moments of asset returns, we introduce the constraint that the *norm* of the portfolio-weight vector be smaller than a given threshold.

Our paper contributes to the literature on optimal portfolio choice in the presence of estimation error in several ways. First, we show that our framework nests as special cases the shrinkage approaches of Jagannathan and Ma (2003) and Ledoit and Wolf (2003, 2004). In particular, we prove that if one solves the minimum-variance problem subject to the constraint that the sum of the absolute values of the weights (1-norm) be smaller than or equal to one, then one retrieves the shortsale-constrained minimum-variance portfolio considered by Jagannathan and Ma (2003). If one constrains the sum of the squares of the portfolio weights (2-norm) to be smaller than a given threshold, then one recovers the class of portfolios considered by Ledoit and Wolf (2004). Similarly, imposing a particular quadratic constraint allows one to recover the portfolios in Ledoit and Wolf (2003). Finally, constraining the squared 2-norm of the portfolio-weight vector to be smaller than or equal to $1/N$ gives the $1/N$ portfolio studied in DeMiguel et al. (2009).

Second, we use this unifying framework to develop new portfolio strategies. For example, we show how the shortsale-constrained portfolio considered in Jagannathan and Ma (2003) can be generalized. In particular, we show that by imposing the constraint that the 1-norm of the portfolio-weight vector be smaller than a threshold that is strictly larger than one, then one obtains a new class of shrinkage portfolios that limit the *total* amount of shortselling in the portfolio, rather than limit the shorting of each asset as in the traditional shortsale-constrained portfolio. To the best of our knowledge, this kind of portfolio has not been analyzed before in the academic literature, although it corresponds closely to the actual portfolio

² Note that although expected returns cannot be forecasted reasonably well from historical data, if a portfolio manager has the skill to forecast expected returns, then he or she may wish to use a mean-variance portfolio rather than a minimum-variance portfolio.

³ For additional evidence, see Jorion (1985, 1986, 1991) and DeMiguel et al. (2009).

Table 1 List of Data Sets Considered

No.	Data set	Abbreviation	N	Time period	Source
1	Ten industry portfolios representing the U.S. stock market	10Ind	10	07/1963–12/2004	K. French ^a
2	Forty-eight industry portfolios representing the U.S. stock market	48Ind	48	07/1963–12/2004	K. French
3	Six Fama and French (1992) portfolios of firms sorted by size and book-to-market	6FF	6	07/1963–12/2004	K. French
4	Twenty-five Fama and French (1992) portfolios of firms sorted by size and book-to-market	25FF	25	07/1963–12/2004	K. French
5	500 randomized stocks from CRSP ^b balanced monthly	500CRSP	500	04/1968–04/2005	CRSP

Notes. This table lists the various data sets of monthly asset returns analyzed, the abbreviation used to refer to each data set, the number of risky assets N in each data set, the time period spanned by the data set, and the source of the data. The data set of CRSP returns (500CRSP) is constructed in a way that is similar to Jagannathan and Ma (2003), with monthly rebalancing: in April of each year we randomly select 500 assets among all assets in the CRSP data set for which there is return data for the previous 120 months as well as for the next 12 months. We then consider these randomly selected 500 assets as our asset universe for the next 12 months.

^ahttp://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

^bCRSP, The Center for Research in Security Prices.

holdings allowed in personal margin accounts.⁴ We also propose a class of portfolios that we term “partial minimum-variance portfolios.” These portfolios are obtained by applying the classical *conjugate-gradient* method (see Nocedal and Wright 1999) to solve the minimum-variance problem. We show that these portfolios are related to the portfolios obtained by imposing a constraint on the 2-norm of the portfolio-weight vector.

Third, we give a Bayesian interpretation for the norm-constrained portfolios and also for the portfolios proposed by Jagannathan and Ma (2003) and Ledoit and Wolf (2003, 2004) that is in terms of a certain prior belief on *portfolio weights* rather than on moments of asset returns.

Fourth, our approach to minimum-variance portfolio selection is also related to a number of approaches proposed in the statistics and chemometrics literature to reduce estimation error in regression analysis. It is known in the literature that optimal portfolio weights in an unconstrained mean- or minimum-variance problem can be thought of as coefficients of an ordinary least squares regression (see, for example, Britten-Jones 1999). It then follows that, in general, constrained weights are the outcome of similarly specified restricted regressions. In particular, constraining the 1-norm of the portfolio vector to be less than a certain threshold is analogous to the statistical technique for regression analysis known as “least absolute shrinkage and selection operator” (lasso) (Tibshirani 1996); constraining the 2-norm of the portfolio vector to be less than a certain threshold corresponds to the statistical technique known as “ridge regression” (Hoerl and Kennard 1970); and computing the “partial minimum-variance portfolio” corresponds to the technique developed in chemometrics known as “partial least squares” (Wold 1975, Frank

and Friedman 1993). These regression techniques and the distribution theory associated with them have been used extensively in the statistics literature.

Fifth, the generalized framework allows one to calibrate the model using historical data to improve its out-of-sample performance. We compare empirically the out-of-sample performance of the norm-constrained portfolios to 10 strategies in the literature for five different data sets. The data sets we consider are listed in Table 1, and the portfolios we evaluate are listed in Table 2. In terms of out-of-sample variance, the norm-constrained portfolios often have a lower variance than the shortsale-constrained minimum-variance portfolio studied in Jagannathan and Ma (2003), the $1/N$ portfolio evaluated in DeMiguel et al. (2009), and also other strategies proposed in the literature, including factor portfolios and the parametric portfolios in Brandt et al. (2005); however, the variance of the norm-constrained portfolios is similar to that of the portfolios in Ledoit and Wolf (2003, 2004). In terms of out-of-sample Sharpe ratio, the portfolios we propose attain a Sharpe ratio that is higher than the shortsale-constrained minimum-variance portfolio, the $1/N$ portfolio, and the portfolios in Ledoit and Wolf (2003, 2004), although the higher Sharpe ratio is accompanied by higher turnover. Finally, the Sharpe ratio and turnover of the proposed portfolios is similar to that of Brandt et al. (2005) but without the need to use firm-specific characteristics.⁵

The remainder of this paper is organized as follows. Section 2 reviews the approaches considered in

⁴ These portfolios are quite popular among practitioners—see the articles in *The Economist* (Buttonwood 2007) and *The New York Times* (Hershey 2007) that describe “130–30” portfolios where investors are long 130% and short 30% of their wealth.

⁵ Because the parametric portfolios in Brandt et al. (2005) rely on firm-specific characteristics, they are not really comparable to the other portfolios we evaluate; however, we decided to include them in our empirical analysis because these portfolios achieve very high Sharpe ratios and hence are an important benchmark. Lauprete (2001) also considers the 1- and 2-norm-constrained portfolios; see also Lauprete et al. (2002) and Welsch and Zhou (2007). Our additional contribution is to relate these methods to the approaches in Jagannathan and Ma (2003) and Ledoit and Wolf (2003, 2004), propose the partial minimum-variance portfolios, and provide comprehensive empirical results.

Table 2 List of Portfolios Considered

No.	Model	Abbreviation
Panel A: Portfolio strategies developed in this paper		
1-norm-constrained minimum-variance portfolio		
	With δ calibrated using cross-validation over portfolio variance	NC1V
	With δ calibrated by maximizing portfolio return in previous period	NC1R
2-norm-constrained minimum-variance portfolio		
	With δ calibrated using cross-validation over portfolio variance	NC2V
	With δ calibrated by maximizing portfolio return in previous period	NC2R
$\hat{\Sigma}_F$ -norm-constrained minimum-variance portfolio		
	With δ calibrated using cross-validation over portfolio variance	NCFV
	With δ calibrated by maximizing portfolio return in previous period	NCFR
Partial minimum-variance portfolios		
	With k calibrated using cross-validation over portfolio variance	PARV
	With k calibrated by maximizing portfolio return in previous period	PARR
Panel B: Portfolio strategies from the existing literature used for comparison		
Simple benchmarks		
1	Equally-weighted (1/ N) portfolio	1/N
2	Value-weighted (market) portfolio	VW
Portfolios that use mean returns with shortsales unconstrained		
3	Mean-variance portfolio with risk aversion parameter $\gamma = 5$	MEAN
4	Bayesian mean-variance portfolio as in Jorion (1985, 1986) with risk aversion parameter $\gamma = 5$	BAYE
Minimum-variance portfolios		
5	Minimum-variance portfolio with shortsales unconstrained	MINU
6	Minimum-variance portfolio with shortsales constrained (Jagannathan and Ma 2003)	MINC
7	Minimum-variance portfolio with the market as the single factor	FAC1
Minimum-variance portfolios where covariance matrix is average of two estimators		
8	Weighted average of sample covariance and identity matrix (Ledoit and Wolf 2004)	LWID
9	Weighted average of sample covariance and 1-factor matrix (Ledoit and Wolf 2003)	LW1F
Parametric portfolios		
10	Parametric portfolio as in Brandt et al. (2005) with a risk-aversion parameter of $\gamma = 5$ using the factors size, book-to-market, and momentum	BSV3

Notes. This table lists the various portfolio strategies we consider. Panel A lists the norm-constrained portfolios developed in this paper, whereas panel B lists the portfolios from the literature. Note that δ is the threshold parameter that limits the norm of the portfolio-weight vector, whereas the order parameter k indicates which of the $N - 1$ partial minimum-variance portfolios to use. The last column gives the abbreviation that we use to refer to the strategy.

Jagannathan and Ma (2003) and Ledoit and Wolf (2003, 2004), which shrink some or all of the elements of the sample covariance matrix. In §3, we propose our general approach, which shrinks the portfolio-weight vector directly. In §4, we discuss the performance of the different portfolios on empirical data. Section 5 concludes. Our main results are highlighted in propositions, and proofs for all the propositions are available in the appendix. Details of how to compute the partial minimum-variance portfolios are available in the online appendix (provided in the e-companion).⁶

2. Existing Approaches: Shrinking the Sample Covariance Matrix

In the absence of shortsale constraints, the minimum-variance portfolio is the solution to the following

optimization problem:

$$\min_w w^T \hat{\Sigma} w, \quad (1)$$

$$\text{s.t. } w^T e = 1, \quad (2)$$

where $w \in \mathbb{R}^N$ is the vector of portfolio weights, $\hat{\Sigma} \in \mathbb{R}^{N \times N}$ is the estimated covariance matrix, $w^T \hat{\Sigma} w$ is the variance of the portfolio return, $e \in \mathbb{R}^N$ is the vector of ones, and the constraint $w^T e = 1$ ensures that the portfolio weights sum up to one. We denote the solution to this shortsale-unconstrained minimum-variance problem by w_{MINU} .

Jagannathan and Ma (2003) study the shortsale-constrained minimum-variance portfolio, w_{MINC} , which is the solution to problem (1)–(2) with the additional constraint that the portfolio weights be nonnegative ($w \geq 0$). They show that the solution to the shortsale-constrained problem coincides with the solution to

⁶ An electronic companion to this paper is available as part of the online version that can be found at <http://mansci.journal.informs.org/>.

the *unconstrained* problem in (1)–(2) if the sample covariance matrix in (1) is replaced by the matrix

$$\hat{\Sigma}_{JM} = \hat{\Sigma} - \lambda e^T - e \lambda^T, \quad (3)$$

where $\lambda \in \mathbb{R}^N$ is the vector of Lagrange multipliers for the shortsale constraint. Because $\lambda \geq 0$, the matrix $\hat{\Sigma}_{JM}$ may be interpreted as the sample covariance matrix *after* shrinkage, because if the shortsale constraint corresponding to the i th asset is binding ($w_i = 0$), then the sample covariance of this asset with any other asset is reduced by λ_i , the magnitude of the Lagrange multiplier associated with its shortsale constraint.

Ledoit and Wolf (2003, 2004) propose replacing the sample covariance matrix with a weighted average of the sample covariance matrix and a low-variance target estimator, $\hat{\Sigma}_{\text{target}}$. Concretely, they propose solving problem (1)–(2), where the matrix $\hat{\Sigma}$ is replaced by

$$\hat{\Sigma}_{LW} = \frac{1}{1+\nu} \hat{\Sigma} + \frac{\nu}{1+\nu} \hat{\Sigma}_{\text{target}}, \quad (4)$$

and where ν is a positive constant. Ledoit and Wolf (2003, 2004) also show how one can estimate the value of ν that minimizes the expected Frobenius norm of the difference between the matrix $\hat{\Sigma}_{LW}$ and the true covariance matrix. They show that this method can be interpreted as shrinking the sample covariance matrix toward the estimator $\hat{\Sigma}_{\text{target}}$. They consider several candidates for $\hat{\Sigma}_{\text{target}}$, such as the identity matrix and the covariance matrix obtained from estimating a 1-factor model with the market as the factor.⁷

3. A Generalized Approach: Constraining the Portfolio Norms

In this section, we propose a class of portfolios that can be viewed as resulting from shrinking the *portfolio-weight vector* instead of shrinking the *moments* of asset returns. Specifically, we define the *norm-constrained minimum-variance portfolio* as the one that solves the traditional minimum-variance problem (1)–(2) subject to the additional constraint that the

norm of the portfolio-weight vector is smaller than a certain threshold δ ; that is, $\|w\| \leq \delta$, where $\|w\|$ is the norm of the portfolio-weight vector. In particular, we consider the 1-norm, $\|w\|_1 = \sum_{i=1}^N |w_i|$, and the A -norm, $\|w\|_A = (w^T A w)^{1/2}$, where $A \in \mathbb{R}^{N \times N}$ is a positive-definite matrix.

Note that the traditional shortsale-unconstrained minimum-variance portfolio, w_{MINU} , is the solution to the norm-constrained problem with $\delta = \infty$. Consequently, if $\delta < \|w_{\text{MINU}}\|$, then the norm of the portfolio that solves the norm-constrained minimum-variance portfolio problem must be strictly smaller than that of the unconstrained minimum-variance portfolio, w_{MINU} .⁸ Hence, imposing a constraint on the norm of the portfolio-weight vector can be interpreted as shrinking the shortsale-unconstrained sample-based minimum-variance portfolio toward a target portfolio. The target portfolio is the one that minimizes the norm of the weight vector subject to the condition that the weights add up to one.⁹ Shrinkage estimators have been a popular method for reducing estimation error ever since their introduction by James and Stein (1961). The idea behind shrinkage estimators is that shrinking an unbiased estimator toward a lower-variance target has a negative and a positive effect. The negative effect is that the shrinkage introduces bias into the resulting estimator, whereas the positive effect is that it reduces the variance of the estimator.

3.1. First Special Case: The 1-Norm-Constrained Portfolios

The *1-norm-constrained portfolio*, w_{NC1} , is the solution to the traditional minimum-variance portfolio problem (1)–(2) subject to the additional constraint that the 1-norm of the portfolio-weight vector be smaller than or equal to a certain threshold δ ; that is, $\|w\|_1 = \sum_{i=1}^N |w_i| \leq \delta$.

The following proposition shows that for the case $\delta = 1$, the solution to the 1-norm-constrained minimum-variance problem is the same as that for the shortsale-constrained minimum-variance portfolio analyzed in Jagannathan and Ma (2003).

PROPOSITION 1. *The solution to the 1-norm-constrained minimum-variance portfolio problem with $\delta = 1$ coincides with the solution to the shortsale-constrained problem.*

⁷ Note that Ledoit and Wolf (2004) actually consider a positive multiple of the identity matrix as their shrinkage target. Specifically, they consider the identity matrix multiplied by the average of the diagonal elements of the sample covariance matrix. This makes sense in the context of their work because their objective is to find the estimator that minimizes the Frobenius norm of the difference with the true covariance matrix. In the context of our manuscript, where the objective is to compute minimum-variance portfolios, it does not matter whether one uses the identity matrix as the shrinkage target, or a positive multiple of the identity matrix. The reason for this is that given a shrinkage target $\hat{\Sigma}_{\text{target}} = \alpha I$, with $\alpha > 0$, and a shrinkage weight $\nu_1 > 0$, the resulting minimum-variance portfolio coincides with that obtained using the identity matrix as a shrinkage target and a shrinkage weight $\nu_2 = \alpha \nu_1 > 0$.

⁸ Note that the norm-constrained portfolios are not obtained by shrinking every weight of the minimum-variance portfolio; instead, the norm-constrained portfolios are obtained by shrinking the total norm of the minimum-variance portfolio. This interpretation of shrinkage is the same as in Tibshirani (1996) in the context of lasso regression.

⁹ Let w_M be the portfolio that minimizes the norm. To see that this is the target portfolio in our approach, note that if we impose the constraint that the portfolio norm is smaller than $\|w_M\|$, then the only portfolio that is feasible with respect to this constraint is precisely w_M .

In contrast to the case where $\delta = 1$, for threshold values of $\delta > 1$ this approach generates a class of portfolios that generalize the shortsale-constrained minimum-variance portfolio. Specifically, some algebraic manipulation can be used to show that the 1-norm constraint $\|w\|_1 \leq \delta$ can be rewritten as

$$-\sum_{i \in \mathcal{N}(w)} w_i \leq \frac{\delta - 1}{2}, \quad (5)$$

where $\mathcal{N}(w)$ is the set of asset indexes for which the corresponding portfolio weight is negative, $\mathcal{N}(w) = \{i: w_i < 0\}$. The left-hand side of (5) is the total proportion of wealth that is sold short and $(\delta - 1)/2$ can be interpreted as the shortsale budget. This shortsale budget can then be freely distributed among all of the assets.

3.2. Second Special Case: The A -Norm-Constrained Portfolios

The A -norm-constrained minimum-variance portfolio is the solution to the traditional minimum-variance problem (1)–(2) subject to the additional constraint that the A -norm of the portfolio-weight vector be smaller than a particular threshold, $\hat{\delta}$. Because the squared A -norm is easier to analyze than the A -norm, we instead impose the equivalent constraint $w^\top A w \leq \delta$, where $\delta = \hat{\delta}^2$.

The following proposition shows the relation between the A -norm-constrained portfolios and the shrinkage portfolios proposed by Ledoit and Wolf.

PROPOSITION 2. *Provided $\hat{\Sigma}$ is nonsingular, for each $\nu \geq 0$ there exists a δ such that the solution to the minimum-variance problem in (1)–(2), with the sample covariance matrix, $\hat{\Sigma}$, replaced by $\hat{\Sigma}_{LW} = (1/(1+\nu))\hat{\Sigma} + \nu/(1+\nu)A$, coincides with the solution to the A -norm-constrained minimum-variance portfolio problem, which is the traditional minimum-variance problem subject to the additional constraint*

$$w^\top A w \leq \delta. \quad (6)$$

In particular, if we choose the matrix A equal to the identity matrix, I , then there is a one-to-one correspondence between the A -norm-constrained portfolios and the shrinkage portfolio proposed in Ledoit and Wolf (2004). On the other hand, if A equals the 1-factor covariance matrix, $\hat{\Sigma}_F$, then there is a one-to-one correspondence with the shrinkage portfolio in Ledoit and Wolf (2003).

Note that for the special case where $A = I$, the A -norm is simply the 2-norm, $\|w\|_2$, and therefore

we refer to these as the *2-norm-constrained minimum-variance portfolios*. To gain intuition about these portfolios, note that the 2-norm constraint $\sum_{i=1}^N w_i^2 \leq \delta$ can be reformulated equivalently as follows:¹⁰

$$\sum_{i=1}^N \left(w_i - \frac{1}{N}\right)^2 \leq \left(\delta - \frac{1}{N}\right). \quad (7)$$

The reformulated constraint (7) demonstrates that imposing the 2-norm constraint on the portfolio weights is equivalent to imposing a constraint that the 2-norm of the difference between this portfolio and the $1/N$ portfolio is bounded by $\delta - 1/N$. Note also that the $1/N$ portfolio is a special case of the 2-norm-constrained portfolio with $\delta = 1/N$.

For the special case where $A = \hat{\Sigma}_F$, imposing a constraint on the $\hat{\Sigma}_F$ -norm of the portfolio is equivalent to imposing a constraint on the portfolio variance under the covariance estimator obtained from a 1-factor model; that is, imposing the constraint that $w^\top \hat{\Sigma}_F w$ is smaller than a certain threshold.

We now compare the 1-, 2-, and $\hat{\Sigma}_F$ -norm constrained portfolios. We have shown above that the 1-norm-constrained portfolios are a generalization of the shortsale-constrained portfolios in which the total amount of shortselling on all assets must remain below a shortsale budget of $(\delta - 1)/2$. Therefore, we may expect the 1-norm-constrained portfolios to have the well-known property of shortsale-constrained minimum-variance portfolios, which tend to assign a weight different from zero to only a few of the assets.¹¹ In contrast, we have shown that the 2-norm-constrained portfolio is the portfolio that minimizes the sample variance subject to the constraint that the square of the 2-norm of the difference with the $1/N$ portfolio is bounded by $\delta - 1/N$. Consequently, we would expect that the 2-norm-constrained portfolios will, in general, remain relatively close to the $1/N$ portfolio and thus will assign a positive weight to all assets. Similarly, we should expect the $\hat{\Sigma}_F$ -norm-constrained portfolios to remain close to the portfolio that minimizes the 1-factor model variance. Thus, investors who believe the total amount of shortselling of the optimal portfolio should not exceed a given budget would want to use a 1-norm constraint. On the other hand, investors who believe that the optimal portfolio is close to the well-diversified $1/N$ portfolio would want to use a 2-norm constraint when solving

¹⁰ To understand why these two formulations are equivalent, observe that $\sum_{i=1}^N (w_i - 1/N)^2 = \sum_{i=1}^N w_i^2 + \sum_{i=1}^N 1/N^2 - \sum_{i=1}^N 2w_i/N = \sum_{i=1}^N w_i^2 - 1/N$, where the last result follows from the fact that $\sum_{i=1}^N 1/N^2 = 1/N$ and $\sum_{i=1}^N 2w_i/N = 2/N$.

¹¹ The reason for this is that the 1-norm function has a kink at portfolios where one or more of the weights are zero, whereas this is not the case for the 2-norm function.

the minimum-variance problem to reduce estimation error. Finally, investors who believe a 1-factor structure holds in the asset-return distribution would use a $\hat{\Sigma}_F$ -norm constraint.

3.3. Third Special Case: The Partial Minimum-Variance Portfolios

The third class of portfolios we propose are obtained by applying the classical conjugate-gradient method (Nocedal and Wright 1999, Chap. 5) to solve the minimum-variance problem. The conjugate-gradient method takes as a starting portfolio some initial guess (in our implementation we use the $1/N$ portfolio) and then generates a sequence of $N - 1$ portfolios in which the terminal portfolio is the shortsale-unconstrained minimum-variance portfolio. We term each of these $N - 1$ intermediate portfolios a *partial minimum-variance portfolio*. In the online appendix, we provide a detailed description of how to compute these portfolios.

Even though the partial minimum-variance portfolios are not obtained by explicitly imposing a constraint on the norm of the minimum-variance portfolio, the following proposition shows that the 2-norm of the partial minimum-variance portfolios is smaller than the 2-norm of the shortsale-unconstrained minimum-variance portfolios.

PROPOSITION 3. *The 2-norm of the k th partial minimum-variance portfolio is smaller than or equal to the 2-norm of the shortsale-unconstrained minimum-variance portfolio for $k \leq N - 1$.*

Moreover, Proposition EC.3 in the online appendix shows that the partial minimum-variance portfolios can be viewed as a discrete first-order approximation to the 2-norm-constrained portfolios.

3.4. A Bayesian Interpretation of the Norm-Constrained Portfolios

Tibshirani (1996, §5) gives a Bayesian interpretation for the regression analysis techniques of lasso and “ridge” regressions; see also Hastie et al. (2001, Chap. 3) and Park and Casella (2008). Here we adapt his argument to give a Bayesian interpretation of the 1- and A -norm-constrained minimum-variance portfolios.

The following proposition shows that the 1-norm-constrained portfolio is the mode of the posterior distribution of portfolio weights for an investor whose prior belief is that the portfolio weights are independently and identically distributed as a Double Exponential distribution.

PROPOSITION 4. *Assume that asset returns are normally distributed. Moreover, assume that the investor believes a priori that each of the shortsale-unconstrained minimum-variance portfolio weights, w_i , is independently*

and identically distributed as a Double Exponential distribution with probability density function

$$\pi(w_i) = \frac{\nu}{2} e^{-\nu}. \quad (8)$$

Furthermore, assume that the investor believes a priori that the variance of the minimum-variance portfolio return, denoted by σ^2 , has an independent prior distribution $\pi(\sigma^2)$. Then there exists a threshold parameter δ such that the weights of the 1-norm-constrained minimum-variance portfolio are the mode of the posterior distribution of the minimum-variance portfolio weights.

The next proposition shows that the A -norm-constrained portfolio is the mode of the posterior distribution of portfolio weights for an investor whose prior belief is that the portfolio weights w_i have a multivariate Normal distribution with covariance matrix A .

PROPOSITION 5. *Assume that asset returns are normally distributed. Moreover, assume that the investor believes a priori that the shortsale-unconstrained minimum-variance portfolio weights are distributed as a multivariate Normal distribution with probability density function*

$$\pi(w) = (2\pi)^{-n/2} |A|^{1/2} e^{-\frac{1}{2} w^T A w}. \quad (9)$$

Furthermore, assume that the investor believes a priori that the variance of the minimum-variance portfolio return, denoted by σ^2 , has an independent prior distribution $\pi(\sigma^2)$. Then there exists a threshold parameter δ such that the weights of the A -norm-constrained minimum-variance portfolio are the mode of the posterior distribution of the minimum-variance portfolio weights.

Note that choosing the portfolio that maximizes the posterior distribution of the minimum-variance portfolio weights guarantees that the investor is choosing the portfolio with the highest probability of being the minimum-variance portfolio given the investor’s prior distribution on portfolio weights and the observed asset-return data. This interpretation is different from that in the traditional Bayesian portfolio choice literature (for instance, Jorion 1986) because in our framework the investor has a prior belief on the portfolio weights rather than on the asset-return distribution. Consequently, while the Bayesian investor in the traditional setting chooses the portfolio that maximizes expected utility with respect to the posterior distribution of asset returns, in our setting the investor chooses the portfolio that maximizes the posterior distribution of portfolio weights (see Tu and Zhou 2009).

3.5. A Moment-Shrinkage Interpretation of the Norm-Constrained Portfolios

The 1- and A -norm-constrained minimum-variance portfolios can also be interpreted as portfolios that result from shrinking some of the elements of the sample covariance matrix.

PROPOSITION 6. *Let the solution to the 1-norm-constrained minimum-variance problem be such that $(w_{\text{NC1}})_i \neq 0$ for $i = 1, \dots, N$. Then w_{NC1} is also the solution to the shortsale-unconstrained minimum-variance problem (1)–(2) if the sample covariance matrix, $\hat{\Sigma}$, is replaced by the matrix*

$$\hat{\Sigma}_{\text{NC1}} = \hat{\Sigma} - \nu n e^T - \nu e n^T, \quad (10)$$

where $\nu \in \mathbb{R}$ is the Lagrange multiplier for the 1-norm constraint at the solution to the 1-norm-constrained minimum-variance problem, and $n \in \mathbb{R}^N$ is a vector whose i th component is one if the weight assigned by the 1-norm-constrained portfolio to the i th asset is negative and zero otherwise.

Proposition 6 shows that the 1-norm-constrained portfolios can also be interpreted as those obtained by shrinking some of the elements of the sample covariance matrix. Concretely, Equation (10) shows that the 1-norm-constrained portfolios can be seen as the result of shrinking by the constant amount ν the covariances of those assets that are being sold short with all the other assets. Note that the amount of shrinkage ν is the same for all assets that are being sold short. This is the main difference between the 1-norm-constrained and the shortsale-constrained portfolios. From Equation (3) it can be observed that for the shortsale-constrained portfolios, the amount of shrinkage applied to the covariances of each of the assets is equal to the Lagrange multiplier corresponding to its shortsale constraint λ_i , and these Lagrange multipliers may take different values for different assets.

Similarly, observe from Proposition 2 that the A -norm-constrained portfolios can be obtained also by shrinking all elements of the sample covariance matrix toward the elements of matrix A .

4. Out-of-Sample Evaluation of the Proposed Portfolios

In this section, we compare across five different data sets (listed in Table 1) the out-of-sample empirical performance of the norm-constrained portfolios to 10 portfolios from the existing literature using three performance metrics: the out-of-sample portfolio variance, the out-of-sample Sharpe ratio, and turnover.

4.1. Description of the Portfolios Evaluated

The norm-constrained minimum-variance portfolios and the partial minimum-variance portfolio that we consider are listed in panel A of Table 2. Note that to use the 1-, 2-, and $\hat{\Sigma}_F$ -norm-constrained minimum-variance portfolios, one needs to choose the value of the threshold parameter δ , which bounds the maximum value that the portfolio norm may take. Similarly, for the partial minimum-variance portfolios, one needs to choose the order parameter k that indicates which of the $N - 1$ partial minimum-variance portfolios to use. The parameters δ and k could be specified exogenously. But in our framework these can also be calibrated to achieve a particular objective, to exploit a particular feature of the returns data, or both. We use two different criteria to calibrate the norm-constrained portfolios: (i) minimizing the portfolio variance and (ii) maximizing the last period portfolio return to exploit positive autocorrelation in portfolio returns as opposed to autocorrelation in the return of individual securities.¹²

If the objective is to minimize the out-of-sample portfolio variance, then to choose δ and k we use the nonparametric technique known as *cross validation*—see Efron and Gong (1983) and Campbell et al. (1997, §12.3.2).¹³ On the other hand, if the objective is to maximize the portfolio return over the last period, we choose θ^* so that $\theta^* = \arg \max_{\theta} w_{\theta}^T r_{\tau}$, in which r_{τ} is the asset-return vector for the last period within the estimation window, and w_{θ} is the norm-constrained or partial minimum-variance portfolio computed using all the data over the estimation

¹² Our motivation for the portfolio autocorrelation criterion is the work by Campbell et al. (1997), who report: “Despite the fact that individual security returns are weakly negatively autocorrelated, portfolio returns—which are essentially averages of individual security returns—are strongly autocorrelated. This somewhat paradoxical result can mean only one thing: large positive cross-autocorrelations across individual securities across time” (p. 74). We consider the return in only the last period because this is where the autocorrelation is highest. We have also considered the average return in the last two to six months, but the out-of-sample performance is worse than when using only the return in the previous month.

¹³ Cross validation works as follows. Given an estimation window composed of τ sample returns, for each t ranging from 1 to τ perform the following four steps. First, delete the t th sample return from the estimation window and compute the sample covariance matrix corresponding to the data set without the t th sample return, $\hat{\Sigma}_{(t)}$. Second, compute the corresponding portfolio $(w_{\theta})_{(t)}$, where $\theta = \delta$ for the case of the norm-constrained portfolios and $\theta = k$ (with $1 \leq k \leq N - 1$) for the case of the partial minimum-variance portfolios. Third, compute the out-of-sample return attained by this portfolio on the t th sample asset return $(r_{\theta})_{(t)} = ((w_{\theta})_{(t)})^T r_t$. Then the variance of the out-of-sample portfolio return is given by the sample variance of the τ out-of-sample returns, $(r_{\theta})_{(t)}$; that is, $\hat{\sigma}_{\theta}^2 = \sum_{t=1}^{\tau} [(r_{\theta})_{(t)} - (\bar{r}_{\theta})_{(t)}]^2 / (\tau - 1)$, where $(\bar{r}_{\theta})_{(t)} = \sum_{t=1}^{\tau} (r_{\theta})_{(t)} / \tau$. Finally, choose the parameter θ^* that minimizes this out-of-sample return variance; that is, $\theta^* = \arg \min_{\theta} \hat{\sigma}_{\theta}^2$.

window with the parameter θ . That is, we choose the parameter $\theta = \{\delta, k\}$ to maximize the return in the last period within the estimation window.

Panel B of Table 2 lists the portfolios to which we compare the performance of the norm-constrained portfolios. The first two comparison portfolios are simple benchmarks that require neither estimation nor optimization: the $1/N$ portfolio and the value-weighted market portfolio, which we compute as the portfolio that assigns a weight to each asset equal to the market capitalization of that asset divided by the market capitalization of all the assets in the data set. We also consider two portfolios that rely on estimates of mean returns: the traditional shortsale-unconstrained mean-variance portfolio and the Bayesian mean-variance portfolio, which is selected using the approach in Jorion (1985, 1986); both of these portfolios are computed assuming a risk aversion of $\gamma = 5$.¹⁴ The next three portfolios we consider are the shortsale-unconstrained minimum-variance portfolio, the shortsale-constrained minimum-variance portfolio analyzed in Jagannathan and Ma (2003), and the minimum-variance portfolio under the assumption that returns are described by a 1-factor model with the market as the only factor. We also consider two minimum-variance portfolios that are formed using a covariance matrix that is a weighted average of two estimators. The first of these portfolios is formed from a combination of the sample covariance matrix and the identity matrix as in Ledoit and Wolf (2004), whereas the second is formed from a combination of the sample covariance matrix and the 1-factor covariance matrix with the market as the factor as in Ledoit and Wolf (2003). Finally, we consider the parametric portfolios of Brandt et al. (2005) that rely on firm-specific characteristics; again, we assume that the investor has a risk aversion equal to five.¹⁵

¹⁴ For the risk-aversion parameter we have considered also values of 1, 2, and 10, but because the insights are similar we do not report these results.

¹⁵ We do not consider several other portfolios. For instance, we do not consider estimators of the covariance matrix based on daily returns because Jagannathan and Ma (2003, §III) find that their performance is similar to that of the shortsale-constrained minimum-variance portfolio with monthly returns. Also, we do not consider portfolios based on the constant correlation model of Elton and Gruber (1973) because these are outperformed by portfolios proposed in Ledoit and Wolf (2003). We report results only for the single market-factor model because Chan et al. (1999) show that several factor models with 1, 3, 4, 8, and 10 factors (based on financial market variables and firm-specific characteristics) are not better than the 1-factor market model. Finally, we consider shortsale constraints only for the minimum-variance portfolio because Jagannathan and Ma (2003) find that “for the factor models and shrinkages estimators, imposing such constraints is likely to hurt” (pp. 1653–1654).

4.2. Description of the Methodology Used to Evaluate Performance

We compare the performance of the norm-constrained portfolios to the portfolios in the literature using three criteria: (i) out-of-sample portfolio variance, (ii) out-of-sample portfolio Sharpe ratio, and (iii) portfolio turnover (trading volume). We use the following “rolling-horizon” procedure for the comparison. First, we choose a window over which to perform the estimation. We denote the length of the estimation window by $\tau < T$, where T is the total number of returns in the data set. For our experiments, we use an estimation window of $\tau = 120$ data points, which for monthly data corresponds to 10 years.¹⁶ Second, using the return data over the estimation window, τ , we compute the various portfolios. Third, we repeat this “rolling-window” procedure for the next month by including the data for the next month and dropping the data for the earliest month. We continue doing this until the end of the data set is reached. At the end of this process, we have generated $T - \tau$ portfolio-weight vectors for each strategy; that is, w_t^i for $t = \tau, \dots, T - 1$ and for each strategy i .

Holding the portfolio w_t^i for one month gives the out-of-sample return at time $t + 1$: $r_{t+1}^i = w_t^{i\top} r_{t+1}$, where r_{t+1} denotes the asset returns. We use the time series of returns and weights for each strategy to compute the out-of-sample variance, Sharpe ratio, and turnover:

$$(\hat{\sigma}^i)^2 = \frac{1}{T - \tau - 1} \sum_{t=\tau}^{T-1} (w_t^{i\top} r_{t+1} - \hat{\mu}^i)^2, \quad \text{with } \hat{\mu}^i = \frac{1}{T - \tau} \sum_{t=\tau}^{T-1} w_t^{i\top} r_{t+1}, \quad (11)$$

$$\widehat{\text{SR}}^i = \frac{\hat{\mu}^i}{\hat{\sigma}^i}, \quad (12)$$

$$\text{Turnover} = \frac{1}{T - \tau - 1} \sum_{t=\tau}^{T-1} \sum_{j=1}^N (|w_{j,t+1}^i - w_{j,t}^i|), \quad (13)$$

where in the definition of turnover, $w_{j,t}^i$ denotes the portfolio weight in asset j chosen at time t under strategy i , $w_{j,t}^i$ the portfolio weight before rebalancing but at $t + 1$, and $w_{j,t+1}^i$ the desired portfolio weight at time $t + 1$ (after rebalancing), implying that turnover is equal to the sum of the absolute value of the rebalancing trades across the N available assets and over the $T - \tau - 1$ trading dates, normalized by the total number of trading dates.

To measure the statistical significance of the difference between the variances and Sharpe ratios of the

¹⁶ We have tried other estimation window lengths, such as $\tau = 60$ and 240, but the results are similar; thus, we report the results only for the case $\tau = 120$.

returns for two given portfolios, we use bootstrapping methods, which are suitable when portfolio returns are not independently and identically distributed as a multivariate Normal. In particular, to compute the p -values for the Sharpe ratios we use a bootstrapping methodology proposed in Ledoit and Wolf (2008) that is designed for the case in which portfolio returns have fat tails and are of a time series nature (for instance, returns are serially correlated or exhibit volatility clustering). Specifically, to test the hypothesis that the Sharpe ratio of the return of portfolio i is equal to that of portfolio n , that is, $H_0: \mu_i/\sigma_i - \mu_n/\sigma_n = 0$, we compute a two-sided p -value using the studentized circular block bootstrap proposed in Ledoit and Wolf (2008) with $B = 1,000$ bootstrap resamples and a block size equal to $b = 5$. We do this using the code available at <http://www.iew.uzh.ch/chairs/wolf.html>. To test the hypothesis that the variance of the returns of two portfolios is equal, that is, $H_0: \sigma_i^2 - \sigma_n^2 = 0$, we use the (nonstudentized) stationary bootstrap of Politis and Romano (1994) to construct a two-sided confidence interval for the difference between the variances. We have used $B = 1,000$ bootstrap resamples and an expected block size $b = 5$. Then we use the methodology suggested in Ledoit and Wolf (2008, Remark 3.2) to generate the resulting bootstrap p -values.

4.3. Discussion of the Out-of-Sample Performance

Table 3 reports the out-of-sample variances for the different portfolios and the corresponding p -value that the portfolio variance for that strategy is different from that for the PARV portfolio. We have also computed all other pairwise p -values, and although we do not report them in the table, we use them when comparing portfolio strategies in our discussion below, and we say that the difference is significant if this p -value is smaller than 5%.

From panel A in Table 3 we see that the out-of-sample variance for the norm-constrained portfolios calibrated using cross validation over the return variance (NC1V, NC2V, NCFV, PARV) is similar across the five data sets. Also, not surprisingly, the out-of-sample variance is lower for these portfolios than for those that are calibrated using the criterion of maximizing the return of the portfolio in the previous period (NC1R, NC2R, NCFR, PARR).

Comparing the variances of the norm-constrained portfolios in panel A to those of the portfolios from the literature listed in panel B of Table 3, we see that the norm-constrained portfolios typically have lower out-of-sample variances than the portfolios from the literature. For instance, NC1V, NC2V, NCFV, and PARV always achieve out-of-sample variances that are lower than those of the 1/ N , value-weighted (VW), and mean-variance (MEAN) portfolios, and the differences are statistically significant. Similarly, the norm-constrained portfolios have lower variances

Table 3 Portfolio Variances

Strategy	10Ind	48Ind	6FF	25FF	CRSP
Panel A: Portfolio strategies developed in this paper					
NC1V	0.00134 (0.07)	0.00126 (0.01)	0.00156 (0.46)	0.00135 (0.65)	0.00074 (0.06)
NC1R	0.00138 (0.98)	0.00135 (0.26)	0.00159 (0.97)	0.00143 (0.10)	0.00080 (0.01)
NC2V	0.00134 (0.08)	0.00137 (0.21)	0.00156 (0.13)	0.00130 (0.43)	0.00066 (0.51)
NC2R	0.00149 (0.10)	0.00176 (0.00)	0.00163 (0.63)	0.00152 (0.02)	0.00087 (0.00)
NCFV	0.00135 (0.39)	0.00131 (0.03)	0.00162 (0.47)	0.00134 (0.82)	0.00052 (0.00)
NCFR	0.00144 (0.30)	0.00166 (0.01)	0.00171 (0.07)	0.00170 (0.00)	0.00068 (0.53)
PARV	0.00138 (1.00)	0.00141 (1.00)	0.00159 (1.00)	0.00133 (1.00)	0.00065 (1.00)
PARR	0.00153 (0.02)	0.00163 (0.01)	0.00161 (0.77)	0.00146 (0.12)	0.00085 (0.00)
Panel B: Portfolio strategies from existing literature					
1/ N	0.00179 (0.00)	0.00221 (0.00)	0.00230 (0.00)	0.00249 (0.00)	0.00169 (0.00)
VW	0.00158 (0.04)	0.00190 (0.00)	0.00191 (0.00)	0.00186 (0.00)	0.00157 (0.00)
MEAN	0.01090 (0.00)	0.38107 (0.00)	0.00353 (0.00)	0.00942 (0.00)	0.00626 (0.00)
BAYE	0.00264 (0.00)	0.06793 (0.00)	0.00221 (0.00)	0.00400 (0.00)	0.00066 (0.44)
MINU	0.00138 (0.68)	0.00186 (0.00)	0.00156 (0.25)	0.00143 (0.09)	0.00104 (0.00)
MINC	0.00134 (0.27)	0.00133 (0.33)	0.00186 (0.01)	0.00176 (0.00)	0.00087 (0.00)
FAC1	0.00145 (0.44)	0.00159 (0.10)	0.00202 (0.00)	0.00241 (0.00)	0.00075 (0.13)
LWID	0.00131 (0.00)	0.00143 (0.75)	0.00155 (0.28)	0.00126 (0.06)	0.00065 (0.75)
LW1F	0.00135 (0.04)	0.00140 (0.79)	0.00158 (0.86)	0.00134 (0.91)	0.00052 (0.00)
BSV3	0.00601 (0.00)	0.00392 (0.00)	0.00306 (0.00)	0.00344 (0.00)	0.00574 (0.00)

Notes. This table reports the monthly out-of-sample variances and the corresponding p -value that the portfolio variance for a strategy is different from that for the partial minimum-variance portfolio calibrated using cross validation over portfolio variance (PARV). The p -values are computed using the stationary bootstrap of Politis and Romano (1994) and the method in Ledoit and Wolf (2008, Remark 3.2).

than the Bayesian portfolio (BAYE) for all data sets except CRSP. The norm-constrained portfolios generally achieve lower out-of-sample variances than the unconstrained minimum-variance portfolio (MINU), and the differences are often statistically significant; for instance, PARV achieves a statistically significant lower variance than MINU for the 48Ind and CRSP data sets. Also, the 1-norm-constrained portfolios (NC1V) generally have lower variance than the shortsale-constrained minimum-variance portfolios (MINC) that NC1V nests, and the differences are

statistically significant for the 6FF, 25FF, and CRSP data sets. We also observe that the norm-constrained portfolios NC1V, NC2V, NCFV, and PARV all attain lower out-of-sample variances than the minimum-variance portfolio based on a 1-factor market model (FAC1), and the differences are often statistically significant; for example, the difference between the variances of PARV and FAC1 is statistically significant for the 6FF and 25FF data sets.

However, the out-of-sample variances of the 2- and $\hat{\Sigma}_F$ -norm-constrained portfolios (NC2V, NCFV) are not always lower than those of the Ledoit and Wolf (2003, 2004) portfolios (LWID, LW1F). For instance, the LWID portfolio attains lower variances than the NC2V portfolio for the 10Ind and 25FF data sets, with the difference being significant for the 10Ind data set. On the other hand, the LW1F portfolio has a higher variance than the NCFV portfolio for the 48Ind data set, and the difference is statistically significant.

Table 4 reports the out-of-sample Sharpe ratios for the different portfolios and the corresponding p -value that the Sharpe ratio for each of these strategies is different from that for the PARR portfolio. We have also computed all other pairwise p -values, and when comparing portfolio strategies in our discussion below we say that the difference is significant if this p -value is smaller than 5%.

From panel A in Table 4 we see that the partial minimum-variance portfolio calibrated by maximizing the portfolio return in the last period (PARR) almost always attains higher Sharpe ratios than the 1-, 2-, and $\hat{\Sigma}_F$ -norm-constrained portfolios (NC1R, NC2R, NCFR), although the differences are significant only between PARR and NC1R, not between PARR and NC2R or NCFR.

Comparing the Sharpe ratios of the norm-constrained portfolios in panel A to those of the portfolios from the literature listed in panel B, we see that the norm-constrained portfolios have higher Sharpe ratios than both the equally weighted (1/N) and the value-weighted (VW) portfolios for all data sets, and the difference is substantial and significant in most cases; in fact, PARR attains a significantly higher Sharpe ratio for all data sets except 48Ind. The difference in performance is even more striking when the norm-constrained portfolios are compared to the traditional mean-variance (MEAN) and the Bayesian mean-variance strategies (BAYE).

The PARR portfolio attains a higher Sharpe ratio than MINU, MINC, and FAC1 for all the data sets, and the difference is often significant. Moreover, the 1-norm-constrained portfolio calibrated by maximizing the return of the portfolio in the previous period (NC1R) typically outperforms the shortsale-constrained minimum-variance portfolio (MINC) that

Table 4 Portfolio Sharpe Ratios

Strategy	10Ind	48Ind	6FF	25FF	CRSP
Panel A: Portfolio strategies developed in this paper					
NC1V	0.2854 (0.06)	0.2886 (0.32)	0.3385 (0.01)	0.3649 (0.00)	0.4013 (0.11)
NC1R	0.2890 (0.05)	0.2831 (0.19)	0.3374 (0.00)	0.3553 (0.00)	0.3706 (0.04)
NC2V	0.2919 (0.08)	0.2855 (0.22)	0.3527 (0.10)	0.4089 (0.22)	0.3994 (0.07)
NC2R	0.3193 (0.40)	0.2891 (0.05)	0.3922 (0.93)	0.4278 (0.36)	0.4672 (0.55)
NCFV	0.2927 (0.21)	0.2808 (0.27)	0.3479 (0.08)	0.3728 (0.03)	0.4463 (0.48)
NCFR	0.3114 (0.56)	0.2723 (0.22)	0.3186 (0.01)	0.3815 (0.11)	0.4243 (0.32)
PARV	0.2841 (0.07)	0.2823 (0.29)	0.3478 (0.10)	0.4077 (0.25)	0.3937 (0.05)
PARR	0.3293 (1.00)	0.3166 (1.00)	0.3912 (1.00)	0.4403 (1.00)	0.4768 (1.00)
Panel B: Portfolio strategies from existing literature					
1/N	0.2541 (0.02)	0.2508 (0.10)	0.2563 (0.00)	0.2565 (0.00)	0.3326 (0.00)
VW	0.2619 (0.02)	0.2698 (0.23)	0.2437 (0.00)	0.2558 (0.00)	0.2748 (0.00)
MEAN	0.0499 (0.01)	−0.0334 (0.01)	0.3214 (0.22)	0.2253 (0.02)	0.0723 (0.00)
BAYE	0.1685 (0.04)	−0.0121 (0.04)	0.3666 (0.57)	0.3151 (0.15)	0.4018 (0.08)
MINU	0.2865 (0.09)	0.2222 (0.01)	0.3640 (0.30)	0.4199 (0.50)	0.3820 (0.03)
MINC	0.2852 (0.06)	0.2914 (0.44)	0.2629 (0.00)	0.2720 (0.00)	0.3985 (0.08)
FAC1	0.3060 (0.54)	0.2674 (0.26)	0.2485 (0.00)	0.2486 (0.00)	0.4166 (0.37)
LWID	0.2962 (0.11)	0.2620 (0.05)	0.3226 (0.00)	0.3974 (0.10)	0.4086 (0.11)
LW1F	0.2902 (0.13)	0.2544 (0.04)	0.3296 (0.01)	0.3927 (0.09)	0.4500 (0.56)
BSV3	0.1157 (0.00)	0.3314 (0.82)	0.3907 (0.99)	0.4047 (0.62)	0.2674 (0.00)

Notes. This table reports the monthly out-of-sample Sharpe ratio and the corresponding p -value that the Sharpe ratio for each of these strategies is different from that for the partial minimum-variance portfolio calibrated by maximizing the portfolio return in the previous period (PARR). The p -values are computed using the studentized circular block bootstrapping methodology in Ledoit and Wolf (2008).

it nests, and the differences are significant for the 6FF and 25FF data sets.

The PARR portfolio attains a higher Sharpe ratio than LWID and LW1F for all the data sets, and the difference is statistically significant for the 48Ind and 6FF data sets, although as we will see below, the higher Sharpe ratio of PARR is accompanied by higher turnover. Also, the 2-norm-constrained portfolio (NC2R) always attains a higher out-of-sample Sharpe ratio than the LWID portfolio, although the difference is significant only for the 6FF data set.

Table 5 Portfolio Turnovers

Strategy	10Ind	48Ind	6FF	25FF	CRSP
Panel A: Portfolio strategies developed in this paper					
NC1V	0.1494	0.2680	0.1729	0.2407	0.6141
NC1R	0.6013	0.8232	1.0064	0.9767	0.9753
NC2V	0.1448	0.3266	0.1946	0.4570	0.5808
NC2R	1.0177	2.7556	1.6594	3.6275	1.0443
NCFV	0.1052	0.2469	0.2790	0.4134	0.5729
NCFR	0.6944	2.3117	1.9952	3.3560	1.0584
PARV	0.1689	0.3838	0.2600	0.4628	0.5743
PARR	1.0414	2.4846	1.6407	3.5657	1.0984
Panel B: Portfolio strategies from existing literature					
1/N	0.0232	0.0311	0.0155	0.0174	0.0595
VW	0.0382	0.0540	0.0213	0.0310	0.2316
MEAN	1.0135	105.6126	0.7987	4.2495	3.0014
BAYE	0.3565	6.6314	0.5388	2.1264	0.6191
MINU	0.1656	0.8286	0.2223	0.7953	0.7769
MINC	0.0552	0.0741	0.0461	0.0841	0.4222
FAC1	0.0935	0.2047	0.1152	0.2398	0.3650
LWID	0.1132	0.4029	0.0905	0.3144	0.5594
LW1F	0.1428	0.4034	0.1455	0.4893	0.5463
BSV3	0.4683	0.9066	0.5384	0.5564	2.1926

Notes. This table reports the monthly turnover of the various portfolios. Turnover is the average percentage of wealth traded in each period and, as defined in Equation (13), is equal to the sum of the absolute value of the rebalancing trades across the N available assets and over the $T - \tau - 1$ trading dates, normalized by the total number of trading dates.

However, the $\hat{\Sigma}_F$ -norm-constrained portfolio calibrated by maximizing the return of the portfolio in the previous period (NCFR) fails to achieve higher Sharpe ratios than the corresponding LW1F portfolio.

Finally, even though the PARR portfolio does not use firm-specific characteristics, it achieves Sharpe ratios that are at least as good as those for the parametric portfolios (BSV) developed in Brandt et al. (2005), and the differences are significant for the 10Ind and CRSP data sets.

From panel A in Table 5 we see that the turnover of the norm-constrained portfolios calibrated using cross validation over the portfolio variance is much lower than that of the portfolios calibrated by maximizing the portfolio return. Panel B of this table shows, not surprisingly, that the best portfolios in terms of turnover are the 1/N and value-weighted portfolios.¹⁷ The turnover of these portfolios is followed by that of the shortsale-constrained minimum-variance portfolio (MINC). The turnovers of the 1-, 2-, and $\hat{\Sigma}_F$ -norm-constrained portfolios and the partial minimum-variance portfolio calibrated with cross validation

¹⁷ We compute the value-weighted portfolio for each data set as the portfolio that assigns a weight to each asset equal to the market capitalization of that asset divided by the market capitalization of all the assets in the data set. Note that because the composition of the “market portfolio” may be changing over time, the strategy of holding the value-weighted portfolio may have a turnover that is different from zero.

over variance, and the LWID and LW1F portfolios proposed in Ledoit and Wolf (2003, 2004), are higher than that of MINC. The shortsale-unconstrained minimum-variance portfolio and the portfolios based on factor models have higher turnover than MINC and the norm-constrained strategies calibrated to minimize portfolio variance. The partial minimum-variance portfolio calibrated by maximizing the portfolio return in the last month (PARR) and the parametric portfolios based on the work by Brandt et al. (2005) have similar turnovers, which are much higher than those of the rest of the portfolios.

5. Conclusion

We provide a general unifying framework for determining portfolios in the presence of estimation error. This framework is based on shrinking directly the portfolio-weight vector rather than some or all of the elements of the sample covariance matrix. This is accomplished by solving the traditional minimum-variance problem (based on the sample covariance matrix) but subject to the additional constraint that the norm of the portfolio-weight vector be smaller than a given threshold. We show that our general framework nests as special cases the shrinkage approaches of Jagannathan and Ma (2003), Ledoit and Wolf (2003, 2004), and the 1/N portfolio studied in DeMiguel et al. (2009). We also compare empirically the out-of-sample performance of the new portfolios we have proposed to 10 strategies in the literature across five data sets. We find that the norm-constrained portfolios often have a higher Sharpe ratio than the portfolio strategies in Jagannathan and Ma (2003), Ledoit and Wolf (2003, 2004), the 1/N portfolio, and other strategies in the literature, such as factor portfolios.

6. Electronic Companion

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

Acknowledgments

The authors gratefully acknowledge financial support from INQUIRE-UK; however, this article represents the views of the authors and not of INQUIRE. The third author is partially supported by the Ministry of Education and Science and Technology of Spain through CICYT Project MTM2007-63140. The authors are very grateful to Department Editor David Hsieh, the associate editor, and two anonymous referees for detailed and extensive comments. The authors thank Mike Chernov, Wayne Ferson, Vito Gala, Michael Gallmeyer, Francisco Gomes, Burton Hollifield, Garud Iyengar, Tongshu Ma, Igor Makarov, Spencer Martin, Javier Peña, Roberto Rigobon, Pedro Santa Clara, Chester Spatt, Catalina Stefanescu, and Bruce Weber for comments. The

authors are also grateful for suggestions from seminar participants at Carnegie Mellon University, Columbia University, Goethe University, INSEAD, Universidade Nova, University of North Carolina, University of Southern California, INQUIRE Conference, 2007 INFORMS Annual Meeting, 2007 ICCOPT-MOPTA Conference, Follow-Up Workshop on Optimization in Finance, Third Asset Management Forum sponsored by Schroder Bank, Fifth World Congress of the Bachelier Society, Second Tremblant Conference on Risk Management, and 2008 American Finance Association Meetings.

Appendix. Proofs of Propositions

PROOF OF PROPOSITION 1. Note that the 1-norm-constrained minimum-variance portfolio problem with $\delta = 1$ can be obtained from the shortsale-constrained problem by replacing the shortsale constraint $w \geq 0$ with the 1-norm constraint $\|w\|_1 \leq 1$. Straightforward algebra shows that, in the presence of the additional constraint that the portfolio weights add up to one, these two constraints are indeed equivalent. \square

PROOF OF PROPOSITION 2. Note that the portfolio that minimizes the A -norm subject to the constraint that the weights add up to one can be written as $w_F = A^{-1}e/(e^T A^{-1}e)$. Hence, if $\delta \geq 1/(e^T A^{-1}e)$, then there exists a feasible point for the A -norm-constrained minimum-variance portfolio problem. If, in addition, $\hat{\Sigma}$ is nonsingular, then there is a unique global minimizer to the problem, and this unique minimizer is characterized by the first-order optimality conditions, which imply that there exist Lagrange multipliers ψ and ν such that

$$2\hat{\Sigma}w + 2\nu Aw - \psi e = 0, \quad (14)$$

$$w^T e = 1, \quad (15)$$

$$w^T A w \leq \delta, \quad (16)$$

$$\nu \geq 0. \quad (17)$$

These conditions imply that the first-order optimality conditions hold for the minimum-variance problem with $\hat{\Sigma}_{LW} = (1/(1+\nu))\hat{\Sigma} + (\nu/(1+\nu))A$. Moreover, because $\hat{\Sigma}$ is nonsingular and positive definite, A is positive definite, and $\nu \geq 0$, we know that the matrix $\hat{\Sigma}_{LW}$ is positive definite, and the first-order optimality conditions for the minimum-variance problem with $\hat{\Sigma}_{LW}$ are then also sufficient for optimality. \square

PROOF OF PROPOSITION 3. Note that the $1/N$ portfolio is the portfolio with minimum 2-norm out of all portfolios whose weights sum up to one. Moreover, we use the $1/N$ portfolio as the starting point for the conjugate-gradient method when we apply this method to solve the minimum-variance portfolio problem. Then, it trivially follows from the analysis in Phatak and De Hoog (2003) and Steihaug (1983) that the 2-norm of the iterates generated by the conjugate-gradient method is nondecreasing. \square

PROOF OF PROPOSITION 4. Because the prior distributions of the portfolio weights are independent, the prior distribution for a portfolio w is

$$\pi(w) = \prod_{i=1}^N \frac{\nu}{2} e^{-\nu} = \frac{\nu^N}{2^N} e^{-\nu \|w\|_1}. \quad (18)$$

Then Bayes theorem implies that the posterior distribution of the portfolio weights conditional on the observed sample returns is

$$\begin{aligned} \pi(w, \sigma^2 | \{r_t\}_{t=1}^T) \\ = \frac{\pi(w)\pi(\sigma^2)L(w, \sigma^2 | \{r_t\}_{t=1}^T)}{\int_{W, \sigma^2} \pi(w)\pi(\sigma^2)L(w, \sigma^2 | \{r_t\}_{t=1}^T) dw d\sigma^2}, \end{aligned} \quad (19)$$

where

$$L(w, \sigma^2 | \{r_t\}_{t=1}^T) = \frac{1}{\sigma^N \sqrt{2^N \pi^N}} \exp\left(-\frac{\sum_{t=1}^T (w^T r_t - w^T \hat{\mu})^2}{2\sigma^2}\right)$$

is the likelihood function of w and σ^2 given the observed sample returns $\{r_t\}_{t=1}^T$.¹⁸ Hence we have that

$$\pi(w, \sigma^2 | \{r_t\}_{t=1}^T) \propto \pi(w)\pi(\sigma^2)L(w, \sigma^2 | \{r_t\}_{t=1}^T), \quad (20)$$

where the symbol \propto means that the distribution is proportional to the right-hand-side term. Simple algebra then shows that

$$\begin{aligned} \pi(w, \sigma^2 | \{r_t\}_{t=1}^T) \\ \propto \frac{\pi(\sigma^2)}{\sigma^N} \exp\left(-\frac{\sum_{t=1}^T (w^T r_t - w^T \hat{\mu})^2}{2\sigma^2} - \nu \|w\|_1\right). \end{aligned} \quad (21)$$

Hence for any value of σ , the portfolio that maximizes the posterior distribution of the minimum-variance portfolio weights, subject to the condition that the portfolio weights add up to one, is the solution to the following optimization problem:

$$\min_w w^T \hat{\Sigma} w + \rho \|w\|_1, \quad (22)$$

$$\text{s.t. } w^T e = 1, \quad (23)$$

where $w^T \hat{\Sigma} w = \sum_{t=1}^T (w^T r_t - w^T \hat{\mu})^2 / (T - 1)$ and $\rho = 2\sigma^2 \nu / (T - 1)$. Moreover, it follows from optimization theory (Nocedal and Wright 1999, Chap. 17) that there exists a threshold parameter δ such that the solution to problem (22)–(23) coincides with the 1-norm-constrained minimum-variance portfolio. Finally, because for any fixed value of σ the portfolio that maximizes the posterior distribution is a 1-norm-constrained minimum-variance portfolio, we know that the portfolio that maximizes the posterior distribution must also be a 1-norm-constrained minimum-variance portfolio. \square

PROOF OF PROPOSITION 5. Following the same line of reasoning as in the proof of Proposition 4, we have by Bayes theorem and simple algebra that

$$\begin{aligned} \pi(w, \sigma^2 | \{r_t\}_{t=1}^T) \\ \propto \frac{\pi(\sigma^2)}{\sigma^N} \exp\left(-\frac{\sum_{t=1}^T (w^T r_t - w^T \hat{\mu})^2}{2\sigma^2} - \frac{1}{2} w^T A w\right). \end{aligned} \quad (24)$$

Hence for any value of σ the portfolio that maximizes the posterior distribution of the minimum-variance portfolio

¹⁸ Note that $L(w, \sigma^2 | \{r_t\}_{t=1}^T)$ coincides with the joint density function for the distribution of the sequence of independent portfolio returns $\{w^T r_t\}_{t=1}^T$ conditional on the values of σ and w ; that is, $\pi(\{w^T r_t\}_{t=1}^T | w, \sigma^2)$. This conditional distribution is customarily termed in Bayesian statistics as the likelihood function.

weights subject to the condition that the portfolio weights add up to one is the solution to the following optimization problem:

$$\min_w w^\top \hat{\Sigma} w + \rho w^\top A w, \quad (25)$$

$$\text{s.t. } w^\top e = 1, \quad (26)$$

where $w^\top \hat{\Sigma} w = \sum_{t=1}^T (w^\top r_t - w^\top \hat{\mu})^2 / (T - 1)$ and $\rho = \sigma^2 / (T - 1)$. Moreover, it follows from optimization theory (Fiocco and McCormick 1968, Chap. 4) that there exists a threshold parameter δ such that the solution to problem (25)–(26) coincides with the A -norm-constrained minimum-variance portfolio for the threshold δ . Finally, because for any fixed value of σ the portfolio that maximizes the posterior distribution is an A -norm-constrained minimum-variance portfolio, we know that the portfolio that maximizes the posterior distribution must also be an A -norm-constrained minimum-variance portfolio. \square

PROOF OF PROPOSITION 6. Straightforward algebra shows that the 1-norm constraint can be written as

$$\|w\|_1 = 1 - 2 \sum_{i \in \mathcal{N}(w)} w_i \leq \delta, \quad (27)$$

where $\mathcal{N}(w)$ is the set of asset indexes for which the corresponding portfolio weight is negative, $\mathcal{N}(w) = \{i: w_i < 0\}$. Hence, the 1-norm-constrained minimum-variance problem is

$$\begin{aligned} \min_w \quad & w^\top \hat{\Sigma} w, \\ \text{s.t.} \quad & w^\top e = 1, \\ & 1 - 2 \sum_{i \in \mathcal{N}(w)} w_i \leq \delta. \end{aligned}$$

At the solution to this problem, and assuming none of the elements of w are equal to zero, there exists a Lagrange multiplier $\nu \geq 0$ such that the solution to this problem coincides with the solution to

$$\begin{aligned} \min_w \quad & w^\top \hat{\Sigma} w - 2\nu \sum_{i \in \mathcal{N}(w)} w_i, \\ \text{s.t.} \quad & w^\top e = 1. \end{aligned}$$

Moreover, because at the solution $w^\top e = 1$, the problem can be equivalently rewritten as

$$\begin{aligned} \min_w \quad & w^\top (\hat{\Sigma} - \nu n e^\top - \nu n^\top e) w, \\ \text{s.t.} \quad & w^\top e = 1, \end{aligned}$$

which completes the proof for the proposition. \square

References

- Best, M. J., R. R. Grauer. 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *Rev. Financial Stud.* **4** 315–342.
- Brandt, M. W. 1999. Estimating portfolio and consumption choice: A conditional Euler equations approach. *J. Finance* **54** 1609–1645.
- Brandt, M. W., P. Santa-Clara, R. Valkanov. 2005. Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. Working paper, University of California, Los Angeles.
- Britten-Jones, M. 1999. The sampling error in estimates of mean-variance efficient portfolio weights. *J. Finance* **54** 655–671.
- Broadie, M. 1993. Computing efficient frontiers using estimated parameters. *Ann. Oper. Res.* **45** 21–58.
- Buttonwood. 2007. Hedge funds' identity crisis. *The Economist* (June 28), http://www.economist.com/finance/displaystory.cfm?story_id=9416476.
- Campbell, J. Y., A. W. Lo, A. G. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
- Chan, L. K. C., J. Karceski, J. Lakonishok. 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *Rev. Financial Stud.* **12** 937–974.
- Chopra, V. K. 1993. Improving optimization. *J. Investing* **8** 51–59.
- Chopra, V. K., W. T. Ziemba. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *J. Portfolio Management* **19** 6–11.
- DeMiguel, V., L. Garlappi, R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financial Stud.* **22** 1915–1953.
- Efron, B., G. Gong. 1983. A leisurely look at the bootstrap, the jack-knife, and cross-validation. *Amer. Statistician* **1** 36–48.
- Elton, E. J., M. J. Gruber. 1973. Estimating the dependence structure of share prices. *J. Finance* **28** 1203–1232.
- Fama, E. F., K. R. French. 1992. The cross-section of expected stock returns. *J. Finance* **47** 427–465.
- Fiocco, A. V., G. P. McCormick. 1968. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York.
- Frank, I. E., J. H. Friedman. 1993. A statistical view of some chemometrics regression tools. *Technometrics* **35** 109–135.
- Frost, P. A., J. E. Savarino. 1986. An empirical Bayes approach to efficient portfolio selection. *J. Financial Quant. Anal.* **21** 293–305.
- Frost, P. A., J. E. Savarino. 1988. For better performance constrain portfolio weights. *J. Portfolio Management* **15** 29–34.
- Green, R., B. Hollifield. 1992. When will mean-variance efficient portfolios be well diversified? *J. Finance* **47** 1785–1809.
- Hastie, T., R. Tibshirani, F. Jerome. 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- Hershey, R. D., Jr. 2007. This fund concept blurs old lines. *The New York Times* (July 8).
- Hoerl, A. E., R. W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **8** 27–51.
- Jagannathan, R., T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance* **58** 1651–1684.
- James, W., C. Stein. 1961. Estimation with quadratic loss. *Proc. 4th Berkeley Sympos. Probab. Statist.* **1**. University of California Press, Berkeley, 452–460.
- Jorion, P. 1985. International portfolio diversification with estimation risk. *J. Bus.* **58** 259–278.
- Jorion, P. 1986. Bayes-Stein estimation for portfolio analysis. *J. Financial Quant. Anal.* **21** 279–292.
- Jorion, P. 1991. Bayesian and CAPM estimators of the means: Implications for portfolio selection. *J. Banking Finance* **15** 717–727.
- Lauprete, G. J. 2001. Portfolio risk minimization under departures from normality. Ph.D. thesis, Sloan School of Management, Massachusetts Institute of Technology, Cambridge.
- Lauprete, G. J., A. M. Samarov, R. E. Welsch. 2002. Robust portfolio optimization. *Metrika* **25** 139–149.

- Ledoit, O., M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance* **10** 603–621.
- Ledoit, O., M. Wolf. 2004. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* **88** 365–411.
- Ledoit, O., M. Wolf. 2008. Robust performance hypothesis testing with the sharpe ratio. *J. Empirical Finance* **15** 850–859.
- Litterman, R. 2003. *Modern Investment Management: An Equilibrium Approach*. John Wiley & Sons, Hoboken, NJ.
- Markowitz, H. M. 1952. Mean-variance analysis in portfolio choice and capital markets. *J. Finance* **7** 77–91.
- Merton, R. C. 1980. On estimating the expected return on the market: An exploratory investigation. *J. Financial Econom.* **8** 323–361.
- Michaud, R. O. 1989. The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts J.* **45** 31–42.
- Nocedal, J., S. J. Wright. 1999. *Numerical Optimization*. Springer-Verlag, New York.
- Park, T., G. Casella. 2008. The Bayesian lasso. *J. Amer. Statist. Assoc.* **103** 681–686.
- Phatak, A., F. De Hoog. 2003. Exploiting the connection between PLS, lanczos, and conjugate gradients: Alternative proofs of some properties of PLS. *J. Chemometrics* **16** 361–367.
- Politis, D. N., J. P. Romano. 1994. The stationary bootstrap. *J. Amer. Statist. Assoc.* **89** 1303–1313.
- Steihaug, T. 1983. The conjugate gradient method and trust regions in large scale optimization. *SIAM J. Numer. Anal.* **20** 626–637.
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. B* **58** 267–288.
- Tu, J., G. Zhou. 2009. Incorporating economic objectives into Bayesian priors: Portfolio choice under parameter uncertainty. *J. Financial Quant. Anal.* Forthcoming.
- Welsch, R. E., X. Zhou. 2007. Application of robust statistics to asset allocation models. *Revstat.* **5** 97–114.
- Wold, H. 1975. Soft modeling by latent variables; the nonlinear iterative partial least squares approach. J. Gani, ed. *Perspectives in Probability and Statistics, Papers in Honour of M. S. Bartlett*. Academic Press, London, 520–540.