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

Victor DeMiguel  Lorenzo Garlappi  Francisco J. Nogales  Raman Uppal

July 16, 2007

Abstract

In this paper, we provide a general framework for identifying portfolios that perform well out-of-sample even 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. We show that our unifying framework nests as special cases the shrinkage approaches of Jagannathan and Ma (2003) and Ledoit and Wolf (2004b), and the $1/N$ portfolio studied in DeMiguel, Garlappi, and Uppal (2007). We also use our general framework to propose several new portfolio strategies. For these new 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 (in terms of portfolio variance, Sharpe ratio, and turnover), the out-of-sample performance of the new portfolios we propose to nine strategies in the existing literature across five datasets. We find that the norm-constrained portfolios we propose have a lower variance and a higher Sharpe ratio than the portfolio strategies in Jagannathan and Ma (2003) and Ledoit and Wolf (2004b), the $1/N$ portfolio, and also other strategies in the literature such as factor portfolios and the parametric portfolios in Brandt, Santa-Clara, and Valkanov (2005).

*We wish to thank Vito Gala, Francisco Gomes, Garud Iyengar, Igor Makarov, Catalina Stefanescu, and Bruce Weber for comments and suggestions. DeMiguel and Uppal are from London Business School; Garlappi is from The University of Texas at Austin; Nogales is from the Universidad Carlos III de Madrid.
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. In this paper, we provide a general framework for determining portfolios with superior out-of-sample performance even 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 research has focused on minimum-variance portfolios, which rely solely on estimates of covariances, and thus, are less vulnerable to estimation error than mean-variance portfolios. Indeed, the superiority of minimum-variance portfolios is demonstrated by extensive empirical evidence that shows that this portfolio usually performs better out of sample than any other mean-variance portfolio—even when the Sharpe ratio or other performance measures that depend on both the portfolio mean and variance are used for evaluating performance. For example, Jagannathan and Ma (2003, p. 1652–3) report:

\]

\[ \text{For additional evidence, see Jorion (1985, 1986, 1991) and DeMiguel, Garlappi, and Uppal (2007).} \]
“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 portfolio 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. 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 structure on the estimator of the covariance matrix; for instance, Chan, Karceski, and Lakonishok (1999) propose a variety of factor models, which reduce the number of parameters to be estimated, and therefore, mitigate the impact of estimation error. A third approach is suggested by Green and Hollifield (1992). They propose a two-step method if returns are generated by a single factor model. First, diversify over the set of high beta stocks and the set of low beta stocks separately; then, short the high beta portfolio and go long the low beta portfolio in order to reduce the systematic risk. A fourth approach has been proposed by Ledoit and Wolf (2004b), who use as an estimator a weighted average of the sample covariance matrix and the identity matrix. This approach can be interpreted as a method that shrinks the sample covariance matrix toward the identity matrix. A fifth approach, which is often used in practice, is to impose shortsale constraints on the portfolio weights (see Frost and Savarino (1988) and 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, p. 1654)
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.” Finally, DeMiguel, Garlappi, and Uppal (2007) demonstrate that even constraining shortsales may not mitigate completely the error in estimating the covariance matrix, and thus, an investor may be best 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 the idea in Brandt (1999) and Britten-Jones (1999), we treat the weights rather than the moments of assets 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. The general framework we develop is then based on solving the traditional minimum-variance problem subject to this constraint on the norm of the portfolio-weight vector; that is, rather than shrinking the moments of asset returns, we require that the weights themselves be bounded. We base our analysis on the minimum-variance portfolio rather than the mean-variance portfolio because, as mentioned above, it is difficult to estimate expected returns precisely, and so portfolios that ignore sample mean returns often outperform portfolios relying on estimated means.\(^3\)

Our paper contributes to the literature on optimal portfolio choice in the presence of estimation error in several ways. One, we show that our framework nests as special cases the shrinkage approaches of Jagannathan and Ma (2003) and Ledoit and Wolf (2004b). 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 1, then one retrieves the shortsale-constrained minimum-variance portfolio considered by Jagannathan and Ma (2003). If, on the other hand, one imposes the constraint that the sum of the squares of the portfolio weights (2-norm) be smaller than a given threshold, we prove that then one recovers the class of portfolios considered by Ledoit and

\(^3\)The norm constraints can be imposed also on the traditional mean-variance-portfolio problem, and our analysis extends in a straightforward manner to this case. As mentioned above, we focus on the minimum-variance problem because it performs better empirically due to the large error associated with estimating mean returns. For completeness, however, in our empirical work we consider as benchmarks both the classical Markowitz mean-variance portfolio and also the Bayesian mean-variance portfolio, as implemented in Jorion (1985).
Wolf (2004b). Finally, we show that if one imposes the constraint that the squared 2-norm of the portfolio-weight vector be smaller than $1/N$, then one gets the $1/N$ portfolio studied in DeMiguel, Garlappi, and Uppal (2007).

Two, we use this general 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 1, then we obtain a new class of shrinkage portfolios where we limit the total amount of shortselling in the portfolio, rather than limiting the shorting asset-by-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 holdings allowed in personal margin accounts. Moreover, these portfolios have recently become quite popular among practitioners—see the articles in *The Economist* (Buttonwood (2007)) and *The New York Times* (Hershey Jr. (2007)) that describe “130–30” portfolios where investors are long 130% and short 30% of their wealth. More importantly, we use our general framework to also develop several new portfolio strategies. Specifically, we propose a different 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 may be interpreted as a discrete first-order approximation to the shrinkage portfolios proposed by Ledoit and Wolf (2004b).

Three, we show how the norm-constrained portfolios we propose and also those proposed by Jagannathan and Ma (2003) and Ledoit and Wolf (2004b) can be interpreted as those of a Bayesian investor who has a certain prior belief on portfolio weights rather than moments of asset returns. Jorion (1986) also shows that the shrinkage estimators of mean returns he considers can be obtained by assuming the investor has a certain prior belief on the means of asset returns. Thus, the main difference between his approach and ours is that we assume a prior belief on the portfolio weights, as opposed to a prior belief on the means of the asset returns.
Four, our approach to minimum-variance portfolio selection is related also 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 portfolios weights in an unconstrained mean- or minimum-variance problem can be thought of as coefficients of an OLS 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, the case where the 1-norm of the portfolio vector is constrained 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)), the case where the 2-norm of the portfolio vector is constrained to be less than a certain threshold corresponds to the statistical technique known as “ridge regression” (Hoerl and Kennard (1970)), and the “partial minimum-variance portfolio” that we study, which is a discrete first-order approximation to the 2-norm-constrained minimum-variance portfolio, corresponds to the technique developed in chemometrics that is 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. By allowing a more general constrained structure in the construction of the portfolio weights and linking this to regression techniques, our paper provides a unified framework for understanding the existing methods proposed in the literature, and our analysis suggests new strategies that perform well out of sample.

Five, the generalized framework allows one to calibrate the model using historical data in order to improve its out-of-sample performance. To demonstrate this, we show how the nonparametric technique known as cross validation can be used to estimate the optimal amount of shrinkage (the optimal value of the threshold on the portfolio norm) that minimizes the estimated out-of-sample variance; see Efron and Gong (1983) and Campbell, Lo, and MacKinley (1997, Section 12.3.2) for discussions of cross validation. We also show how the time-series properties of portfolio returns (as opposed to individual security returns) that have been documented by Campbell, Lo, and MacKinley (1997) can be used to set the level of the constraint on the portfolio norm in order to improve the portfolio Sharpe ratio.
Finally, we compare empirically the out-of-sample performance of the three norm-constrained portfolios we propose to nine strategies in the existing literature for five different datasets. The portfolios we evaluate are listed in Table 1 and the datasets we consider are listed in Table 2. We compare performance along three dimensions: (i) out-of-sample portfolio variance, (ii) out-of-sample portfolio Sharpe ratio, and (iii) portfolio turnover or trading volume. We find that the new portfolios we propose have a lower variance and a higher Sharpe ratio than those studied in Jagannathan and Ma (2003), Ledoit and Wolf (2004b), the $1/N$ portfolio evaluated in DeMiguel, Garlappi, and Uppal (2007), and also other strategies proposed in the literature, including factor portfolios and the parametric portfolios in Brandt, Santa-Clara, and Valkanov (2005), which use firm-specific characteristics of the cross-section of stock returns.

Our work is related closely to Lauprete (2001), who also considers the 1- and 2-norm-constrained portfolios. Our additional contribution is first to demonstrate that the 1- and 2-norm-constrained portfolios include as particular cases the shortsale-constrained minimum-variance portfolios considered by Jagannathan and Ma (2003), the shrinkage portfolios proposed by Ledoit and Wolf (2004b), and the $1/N$ portfolio. Second, we propose the partial minimum-variance portfolios and show that they are a discrete first-order approximation to the 2-norm-constrained portfolios. Third, we show how our general framework can be used to calibrate the portfolio strategies. Fourth, we provide comprehensive empirical results comparing the 1- and 2-norm-constrained portfolios, as well as the partial minimum-variance portfolios, to other portfolios from the literature.

The remainder of this paper is organized as follows. Section 2 reviews the approaches considered in Jagannathan and Ma (2003) and Ledoit and Wolf (2004b), which shrink some or all of the elements of the sample covariance matrix. In Section 3, we propose our general approach, which shrinks the portfolio weights directly, and we show that it nests as special cases the approaches in Jagannathan and Ma (2003), Ledoit and Wolf (2004b), and the $1/N$ portfolio. We also give Bayesian and moment-shrinking interpretations of the proposed portfolios. In Section 4, we provide a brief

---

4We consider the Sharpe ratio as a performance criterion because although none of the portfolios considered uses an estimate of the mean return, as argued above these portfolios tend to outperform those that take also mean returns into account.

5Because the parametric portfolios in Brandt, Santa-Clara, and Valkanov (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 a very important benchmark.

6See also, Lauprete, Samarov, and Welsch (2002) and Welsch and Zhou (2007).
discussion of the performance of the different portfolios on empirical data, with a more detailed
discussion given in Appendix C. Section 5 concludes. Details of how to compute the partial
minimum-variance portfolios are provided in Appendix A. Our main results are highlighted in
propositions and proofs for all the propositions are collected in Appendix B.

2 Existing Approaches: Shrinking the Sample Covariance Matrix

This section is divided into three parts. First, in Section 2.1, we describe the problem of identifying
the minimum-variance portfolio in the absence of shortsale constraints. Then, we review two
existing approaches for reducing the impact of the error in estimating the covariance matrix of
asset returns. In Section 2.2, we describe the approach analyzed in Jagannathan and Ma (2003), in
which shortsale constraints are imposed on the minimum-variance portfolio problem. In Section 2.3,
we describe the method developed by Ledoit and Wolf (2004b). Both approaches can be interpreted
as methods that shrink some or all of the elements of the sample covariance matrix to reduce the
impact of estimation error. In Section 3 that follows, we show that these portfolios can also be
interpreted as special cases of the general framework we develop, which shrinks directly the portfolio
weights rather than the elements of the covariance matrix.

2.1 Base Case: The Shorts ale-Unconstrained Minimum-Variance Portfolio

In the absence of shortsale constraints, the minimum-variance portfolio is the solution to the fol-
lowing optimization problem:

$$\min_{w} w^{\top} \hat{\Sigma} w,$$

s.t. $w^{\top} e = 1,$

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

Jagannathan and Ma (2003) study the effect of imposing shortsale constraints on the minimum-variance portfolio; that is, they consider the solution to the shortsale-constrained minimum-variance problem,

\[
\min_w w^\top \hat{\Sigma} w, \tag{3}
\]

s.t. \( w^\top e = 1, \tag{4} \)

\( w \geq 0. \tag{5} \)

We denote the solution to the shortsale-constrained minimum-variance problem by \( w_{MINC} \). Jagannathan and Ma show that the solution to the shortsale-constrained problem coincides with the solution to the unconstrained problem in (1)–(2) if the sample covariance matrix in (1) is replaced by the matrix

\[
\hat{\Sigma}_{JM} = \hat{\Sigma} - \lambda e^\top - e\lambda^\top, \tag{6}
\]

in which \( \lambda \in \mathbb{R}^N \) is the vector of Lagrange multipliers for the shortsale constraint \( w \geq 0 \) at the solution to the constrained problem (3)–(5). 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 \( \lambda_i \), the magnitude of the Lagrange multiplier associated with its shortsale constraint.

Figure 1 depicts the shortsale-constrained minimum-variance portfolio for the case with three risky assets. The three axes in the reference frame give the portfolio weights, \( w_1, w_2, \) and \( w_3 \), for the three risky assets. Two triangles are depicted in the figure. The larger triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The smaller triangle (colored) represents the set of portfolios whose weights are nonnegative and sum up to one; that is, the set of shortsale-constrained portfolios. The ellipses centered around the minimum-variance portfolio, \( w_{MINU} \), depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The shortsale-constrained minimum-variance portfolio is at the point where the colored triangle is tangent to the iso-variance curves. The figure also shows the
location of the $1/N$ portfolio, which can be interpreted as the portfolio that ignores both the mean
returns and the covariances of returns.

### 2.3 Shrinking the Sample-Covariance Matrix Toward the Identity Matrix

To reduce the effect of estimation error, Ledoit and Wolf (2004b) propose replacing the sample
covariance matrix with a convex combination of the sample covariance matrix and the identity
matrix. Concretely, they propose solving problem (1)–(2), in which the matrix $\hat{\Sigma}$ is replaced by
the following matrix

$$\hat{\Sigma}_{LW} = \hat{\Sigma} + \nu I,$$

in which $\nu \in \mathbb{R}$ is a positive constant, and $I \in \mathbb{R}^{N \times N}$ is the identity matrix. Ledoit and Wolf
also show how one can estimate the value of $\nu$ 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 identity matrix.\(^7\)

Figure 2 depicts the Ledoit-Wolf portfolios for the case with three risky assets. The set of
shortsale-constrained portfolios and the iso-variance curves are depicted as in Figure 1. The Ledoit
and Wolf portfolios form a curve that joins the $1/N$ portfolio, which is the Ledoit and Wolf portfolio
obtained if $\nu = \infty$ in equation (7), with the minimum-variance portfolio equation $w_{MINU}$, which
is the Ledoit and Wolf portfolio obtained if $\nu = 0$ in equation (7).

### 3 A Generalized Approach: Constraining the Portfolio Norms

In this section, we propose a general class of portfolios that results from solving the traditional
minimum-variance problem but subject to the additional constraint that the norm of the portfolio-
weight vector is smaller than a certain threshold $\delta$. These portfolios can be viewed as resulting from
shrinking the portfolio weights of the shortsale-unconstrained minimum-variance portfolio instead
of shrinking the moments of asset returns.

\(^7\)Ledoit and Wolf (2004b) actually propose solving the minimum-variance problem in (1)–(2) with the sample
covariance matrix $\hat{\Sigma}_{LW} = \frac{1}{1+\nu} \hat{\Sigma} + \frac{\nu}{1+\nu} I$. Note that the portfolio weights that solve this problem are the same as
the ones obtained from using the matrix $\hat{\Sigma}_{LW} = \hat{\Sigma} + \nu I$. We focus on this second matrix because it is easier for our
analytical purposes.
We start by describing the general class of portfolios in Section 3.1. Then, we consider three particular cases, which are motivated by the three methods used in the statistics literature to reduce estimation error in regression analysis: “least absolute shrinkage and selection operator” (lasso), “ridge regression”, and “partial least squares”. First, in Section 3.2, we consider the case where the 1-norm of the portfolio vector is constrained to be less than a certain threshold. Second, in Section 3.3, we consider the case where the 2-norm of the portfolio vector is constrained to be less than a certain threshold. Third, in Section 3.4, we study a portfolio that is a discrete first-order approximation to the 2-norm-constrained minimum-variance portfolio. We call this the “partial minimum-variance portfolio.” Finally, we provide two different interpretations of the norm-constrained portfolios: in Section 3.5 we provide a Bayesian interpretation and in Section 3.6 we give a moment-shrinkage interpretation.

3.1 The General Norm-Constrained Minimum-Variance Portfolio

We define the \textit{norm-constrained minimum-variance problem} as the one that solves the traditional minimum-variance problem subject to the additional constraint that the norm of the portfolio-weight vector is smaller than a certain threshold $\delta$:

\[
\min_w \ w^\top \hat{\Sigma} w, \tag{8}
\]
\[
\text{s.t.} \quad w^\top e = 1, \tag{9}
\]
\[
\|w\| \leq \delta, \tag{10}
\]

in which $\|w\|$ is the norm of the portfolio-weight vector. We consider the 1-norm and the 2-norm, which are defined as:

\[
\|w\|_p = \left( \sum_{i=1}^{N} |w_i|^p \right)^{1/p}, \tag{11}
\]

for $p = 1$ and 2, respectively. We denote the solution to the general norm-constrained problem by $w_{NC}$.

\textsuperscript{8}Observe that we take as the starting point of our analysis the minimum-variance portfolio problem rather than the mean-variance problem, even though we will evaluate performance in terms of the Sharpe ratio, which includes both the mean and variance of portfolio returns. As discussed in the introduction, the reason for starting with the minimum-variance portfolio is that it is difficult to estimate mean returns with much precision, and Jagannathan and Ma (2003) and DeMiguel, Garlappi, and Uppal (2007) find empirically that the out-of-sample performance is better if
Note that the traditional short-sale-unconstrained minimum-variance portfolio, \( w_{MINU} \), is the solution to the norm-constrained problem with \( \delta = \infty \). Consequently, if \( \delta < \| w_{MINU} \| \), then the norm of the portfolio that solves problem (8)–(10) must be strictly smaller than that of the unconstrained minimum-variance portfolio, \( w_{MINU} \). To highlight this result we state it in the following proposition.

**Proposition 1** For each \( \delta < \| w_{MINU} \| \), the solution to the norm-constrained problem, \( w_{NC} \), satisfies the following inequality:

\[
\| w_{NC} \| < \| w_{MINU} \|. \tag{12}
\]

Proposition 1 states that the norm-constrained minimum variance portfolio, \( w_{NC} \), is a shrinkage estimator of the short-sale-unconstrained minimum-variance portfolio, \( w_{MINU} \). 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 deterministic target has a negative and a positive effect. The negative effect is that the shrinkage introduces bias into the resulting estimator. The positive effect is that shrinking the estimator toward a deterministic target reduces the variance of the estimator. The challenge is to choose the amount of shrinkage that optimizes the tradeoff between bias and variance. We explain in Section C.1 how this can be done for the norm-constrained policies.

### 3.2 First Particular Case: The 1-Norm-Constrained Portfolios

In this section, we consider the class of norm-constrained portfolios obtained by constraining the 1-norm of the portfolio-weight vector. The 1-norm is defined as the sum of the absolute values of one ignores estimates of mean returns and considers only the covariances between returns. If one wanted to consider the mean-variance portfolio problem with norm constraints, one would only need to replace the objective function in the problem defined in Equations (8)–(10). Specifically, one needs to replace (8) with: \( \max_w w^T \hat{\mu} - \frac{\gamma}{2} w^T \hat{\Sigma} w \), where \( \gamma \) is the investor’s risk aversion parameter and \( \mu \) is the \( \mathbb{R}^N \) vector of expected returns. With this change, our analysis would then apply to the mean-variance portfolio problem. We have implemented the resulting norm-constrained mean-variance portfolios and our empirical results (not reported) show that these portfolios do indeed achieve lower out-of-sample Sharpe ratios than the norm-constrained minimum-variance portfolios.

---

9 An estimator is unbiased if its expected value coincides with the parameter being estimated.
the portfolio weights $\|w\|_1 = \sum_{i=1}^{N} |w_i|$. The resulting 1-norm-constrained portfolio problem is:

$$
\min_{w} \quad w^\top \hat{\Sigma} w, \\
\text{s.t.} \quad w^\top e = 1,
$$

(13)

$$
\sum_{i=1}^{N} |w_i| \leq \delta,
$$

(15)

and we denote its solution by $w_{NC1}$.

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 by Jagannathan and Ma (2003).

**Proposition 2** The solution to the 1-norm-constrained problem (13)–(15) with $\delta = 1$ coincides with the solution to the shortsale-constrained problem (3)–(5).

A couple of comments are in order. First, note that Proposition 2 states that the shortsale-constrained minimum-variance portfolio, which is studied in Jagannathan and Ma (2003), is a special case of the 1-norm-constrained minimum-variance portfolio for the threshold value of $\delta = 1$. Then, by Proposition 1, we have that the shortsale-constrained portfolio can be interpreted as a portfolio resulting from shrinking the portfolio weights of the unconstrained minimum-variance portfolio. This interpretation is different from that given by Jagannathan and Ma (2003), who show that the shortsale-constrained portfolio can be interpreted as resulting from shrinking some of the coefficients of the sample covariance matrix. The interpretation we provide is useful because for an investor it is typically easier to think in terms of portfolio weights than the elements of the sample covariance matrix.

Second, for threshold values of $\delta$ in (15) that are strictly larger than 1, our approach generates a class of portfolios that generalize the shortsale-constrained minimum-variance portfolio. To see this, note that the 1-norm can be rewritten as

$$
\|w\|_1 = \sum_{i=1}^{N} |w_i| = \sum_{i \in \mathcal{P}(w)} w_i - \sum_{i \in \mathcal{N}(w)} w_i,
$$

(16)
in which \( w_i \) is the portfolio weight on the \( i^{th} \) asset, \( \mathcal{P}(w) \) is the set of asset indexes for which the corresponding portfolio weight is greater or equal than zero, \( \mathcal{P}(w) = \{i : w_i \geq 0\} \), and \( \mathcal{N}(w) \) is the set of asset indexes for which the corresponding portfolio weight is negative, \( \mathcal{N}(w) = \{i : w_i < 0\} \). Moreover, if the portfolio weights sum up to one (\( w^\top e = 1 \)), we have that

\[
\sum_{i \in \mathcal{P}(w)} w_i = 1 - \sum_{i \in \mathcal{N}(w)} w_i. \tag{17}
\]

From (16) and (17) we obtain

\[
\|w\|_1 = 1 - 2 \sum_{i \in \mathcal{N}(w)} w_i. \tag{18}
\]

Equation (18) gives insight into the structure of this new class of portfolios. In particular, observe from (18) that imposing the constraint \( \|w\|_1 < \delta \) for some \( \delta > 1 \) is equivalent to assigning a **shortsale budget**; that is, a budget for the total amount of shortselling allowed in the portfolio. Indeed, using (18) we can rewrite the constraint in (15) on the 1-norm of the portfolio-weight vector as follows:

\[
- \sum_{i \in \mathcal{N}(w)} w_i < \frac{\delta - 1}{2}, \tag{19}
\]

in which the term \(- \sum_{i \in \mathcal{N}(w)} w_i \) is the total proportion of wealth that is sold short and \((\delta - 1)/2 \) is the shortsale budget. This shortsale budget can then be freely distributed among all of the assets. For instance, one could decide to short sell only one asset, and assign to this particular asset the weight of \(- (\delta - 1)/2 \), or one could distribute the shortsale budget equally among a few of the assets.

The \( w_{NC1} \) portfolio is different also from the “generalized shortsale-constrained portfolio” considered in DeMiguel, Garlappi, and Uppal (2007), which is the minimum-variance portfolio subject to the constraint \( w \geq -\xi e \) for some \( \xi > 0 \), which requires that the amount of shortselling allowed for each assets is exactly the same. In contrast, the 1-norm constraint allows the investor more flexibility about how to distribute the allowed shortsale budget across all of the assets.

Figure 3 depicts the 1-norm-constrained portfolios for the case with three risky assets. The three axes in the reference frame give the portfolio weights for the three risky assets. Two triangles are depicted in the figure. The larger triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The smaller triangle (colored)
represents the set of portfolios whose weights are nonnegative and sum up to one; that is, the set of shortsale-constrained portfolios. The ellipses centered around the minimum-variance portfolio, \( w_{MINU} \), depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The hexagons represent the iso-1-norm curves; that is, the sets of portfolios with equal 1-norm, which are obtained from the intersection of the 2-dimensional plane formed by all portfolios whose weights sum to one with the cube describing the region of weights satisfying the 1-norm constraint. For each value of the threshold \( \delta = \{ \delta_1, \delta_2, \ldots, \delta_5 \} \), the 1-norm-constrained minimum-variance portfolio is the point where the corresponding iso-1-norm hexagon is tangent to the iso-variance curve. The 1-norm-constrained portfolios corresponding to values of the threshold parameter \( \delta \) ranging from 1 to \( \|w_{MINU}\|_1 \) describe a curve that joins the shortsale-constrained minimum-variance portfolio, \( w_{MINC} \), with the shortsale-unconstrained minimum-variance portfolio, \( w_{MINU} \).

### 3.3 Second Particular Case: The 2-Norm-Constrained Portfolios

We now consider the class of portfolios obtained by constraining the 2-norm of the vector of portfolio weights. The 2-norm is the traditional Euclidean norm in \( \mathbb{R}^N \), obtained from (11) by setting \( p = 2 \), or, in matrix form, \( \|w\|_2 = (w^\top w)^{1/2} \). Note, however, that the solution to the minimum-variance problem with a constraint on the 2-norm \( \|w\|_2 < \delta \) is the same as the solution to the problem with a constraint on the squared 2-norm \( \|w\|_2^2 < \delta^2 \). Because the squared 2-norm is easier to analyze than the 2-norm, we define the 2-norm-constrained minimum-variance portfolio problem as follows:

\[
\min_w \quad w^\top \hat{\Sigma} w, \tag{20}
\]

\[
\text{s.t.} \quad w^\top e = 1, \tag{21}
\]

\[
\sum_{i=1}^{N} w_i^2 \leq \delta, \tag{22}
\]

and we denote the solution to this problem by \( w_{NC2} \).
To gain intuition about the 2-norm constrained minimum-variance portfolios, note that the constraint in (22) can be reformulated equivalently as follows:\(^{10}\)

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

(23)

The reformulated constraint (23) demonstrates that imposing the 2-norm constraint on the portfolio weights in (22) 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 from (20), (21) and (23) that the 1/N portfolio is a special case of the 2-norm-constrained portfolio if we set \(\delta = 1/N\).

The following proposition shows that the 2-norm constrained portfolio belongs to the class of unconstrained portfolios that would be obtained by using shrinkage techniques on the covariance matrix as in Ledoit and Wolf (2004b).

**Proposition 3** Provided \(\hat{\Sigma}\) is nonsingular, for each \(\delta \geq 1/N\) there exists a \(\nu \geq 0\) such that the solution to problem (20)–(22) is the solution to the minimum-variance problem in (1)–(2) with the sample covariance matrix, \(\hat{\Sigma}\), replaced by \(\hat{\Sigma}_{LW} = \hat{\Sigma} + \nu I\).

Proposition 3 shows that the Ledoit and Wolf (2004b) portfolio can be interpreted as one that shrinks the portfolio weights. Again, this is a different interpretation than the one given in Ledoit and Wolf (2004b), in which the portfolio is viewed as one obtained from shrinking the sample covariance matrix toward the identity matrix. Ledoit and Wolf (2004a) consider other targets toward which to shrink the sample covariance matrix; for instance, the single-factor matrix and the constant correlation matrix. The 2-norm-constrained portfolios described above can be extended to generalize also these other approaches discussed in Ledoit and Wolf. To do this, one simply needs to replace the constraint \(w^T w \leq \delta^2\) by the constraint \(w^T Fw \leq \delta^2\), where \(F\) is the target matrix.

For expositional simplicity, we focus on the case where the target matrix is the identity.

Figure 4 depicts the 2-norm-constrained portfolios for the case with three risky assets. The iso-variance curves are depicted as in Figure 3. The circumferences centered on the equally-weighted

---

\(^{10}\)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 \(\sum_{i=1}^{N} 2w_i/N = 2/N\).
portfolio, $1/N$, are the iso-2-norm curves; that is, the set of portfolios with the same 2-norm. We depict four iso-2-norm curves corresponding to four different threshold levels: $\delta_1$, $\delta_2$, $\delta_3$, and $\delta_4$. For a given threshold $\delta$, the 2-norm-constrained portfolio is the point where the corresponding iso-2-norm curve is tangent to an iso-variance curve.\footnote{To see why the solution must be at a tangency point of the two level sets, note that at a point where the two level sets cross each other, we can always find a point with smaller sample variance by moving along the level set for the portfolio norm.} For values of $\delta$ ranging from $1/N$ to $\|w_{MINU}\|_2^2$, the 2-norm-constrained portfolios describe a smooth curve that joins the $1/N$ portfolio (for $\delta = 1/N$) with the minimum-variance portfolio (for $\delta \geq \|w_{MINU}\|_2^2$); this will be formally proven in Proposition 7.

We now compare the 1-norm and the 2-norm portfolios in Figures 3 and 4, respectively. From Figure 3, observe that because the 1-norm level sets are shaped as hexagons, the 1-norm-constrained minimum-variance portfolios tend to be located at the vertices of these hexagons—in Figure 3 it can be observed that this is the case for $\delta_1$, $\delta_2$, and $\delta_3$. This implies that the 1-norm-constrained portfolios are likely to assign a zero weight to at least some of the assets. This phenomenon is well known for the case of the shortsale-constrained minimum-variance portfolio, which tends to assign a zero weight to a large subset of the total number of assets available for investment; for a discussion of this see, for instance, Jagannathan and Ma (2003). But Figure 3 shows that, although this phenomenon also takes place for the rest of the 1-norm-constrained portfolios with $\delta > 1$, these portfolios will tend in general to invest in a higher number of assets than the shortsale-constrained minimum-variance portfolio. To see this, note that the vertices of the triangle (level set for $\delta = 1$) correspond to portfolios that assign a zero weight to two out of the three assets, whereas the vertices of each hexagon (level sets for $\delta > 1$) correspond to points where the weight on only one of the three assets is zero. Hence, we may expect the 1-norm-constrained portfolios with $\delta > 1$ to invest in a larger number of assets than the shortsale-constrained minimum-variance portfolio.

In contrast, the 2-norm-constrained portfolios will typically assign a non-zero (positive or negative) weight to all assets. Thus, in situations where the investor prefers a diversified portfolio that has some weight in all assets, the investor should use the 2-norm-constrained portfolios, while if the investor is interested in portfolios with investment in a smaller number of assets, then she should focus on the 1-norm-constrained portfolios.
This is consistent with the interpretation of the 1- and 2-norm portfolios given in the previous sections. Specifically, in Section 3.3 we showed 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. Also, in Section 3.2 we showed 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. Summarizing, 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 the minimum-variance problem to reduce estimation error. Investors who, on the other hand, believe the total amount of shortselling of the optimal portfolio should not exceed a given budget, would want to use a 1-norm constraint.

3.4 Third Particular Case: The Partial Minimum-Variance Portfolios

In this section, we propose a class of portfolios that are obtained by applying the classical conjugate-gradient method (Nocedal and Wright (1999, Chapter 5)) to solve the minimum-variance problem. The conjugate gradient is an iterative method that, starting from an initial guess reaches the optimal solution, which is the shortsale-unconstrained minimum-variance portfolio, in $N - 1$ successive steps, where $N$ is the number of assets (in Appendix A we provide a detailed description of this algorithm). This method generates a sequence of portfolios, where the first portfolio is some initial guess (taken to be the $1/N$ portfolio in our implementation) and the terminal portfolio is the shortsale-unconstrained minimum-variance portfolio. We term each of the intermediate portfolios a partial minimum-variance portfolio.

Even though the partial minimum-variance portfolios are not obtained by imposing explicitly a constraint on the norm of the minimum-variance portfolio, we show that the norm of the par-
tial minimum-variance portfolios is indeed smaller than the norm of the shortsale-unconstrained minimum-variance portfolios. Moreover, we also show that the partial minimum-variance portfolios can be viewed as a discrete first-order approximation to the 2-norm-constrained portfolios. For this reason, we consider the partial minimum-variance portfolios as a particular case of the norm-constrained portfolios.

The partial minimum-variance portfolios are the counterpart in portfolio selection to the statistical technique of partial least squares for regression analysis, which has been shown to perform very well under certain conditions (see Wold (1975); Frank and Friedman (1993), and Friedman and Popescu (2004)). Our empirical analysis in Section 4 shows that the partial minimum-variance portfolios perform very well also on financial data and that they often outperform not just the portfolios in the existing literature but also the 1- and 2-norm-constrained portfolios described above.

The partial minimum-variance portfolios form a set of \( N-1 \) portfolios that join the \( 1/N \) portfolio and the shortsale-unconstrained minimum-variance portfolio. Specifically, the first of these \( N-1 \) portfolios, which we term the “first partial minimum-variance portfolio,” is a weighted average of the \( 1/N \) portfolio and the so-called “first conjugate portfolio.” The term “partial” refers to the fact that the first partial minimum-variance portfolio minimizes the sample variance within the subset of portfolios formed by combinations of the \( 1/N \) portfolio and the first conjugate portfolio. This first conjugate portfolio is the zero-cost portfolio (i.e., a portfolio whose weights sum up to zero) that induces the maximum marginal decrease in the sample variance when combined with the \( 1/N \) portfolio.

The second partial minimum-variance portfolio is then that combination of the first partial minimum-variance portfolio and the second conjugate portfolio that minimizes the sample variance, where the second conjugate portfolio is the zero-cost portfolio that, when combined with the first partial minimum-variance portfolio, induces the maximum marginal decrease in the sample portfolio variance, subject to the condition that it is conjugate with respect to the first conjugate portfolio; that is, subject to the condition that the first two conjugate portfolios are uncorrelated with respect to the sample covariance matrix.
By iterating this process \( N - 1 \) times we generate a discrete set of \( N - 1 \) portfolios (including the shortsale-unconstrained minimum-variance portfolio) that join the \( 1/N \) and the shortsale-unconstrained minimum-variance portfolios. Details of how to compute the partial minimum-variance portfolios are provided in Appendix A.

### 3.4.1 Analytical Expression for the First Partial Minimum-Variance Portfolio

In this section, we provide an analytical expression for the first partial minimum-variance portfolio and use this to give some intuition for its properties.

To do so, we first note that the gradient of the sample portfolio variance, \( w^\top \hat{\Sigma} w \), with respect to the portfolio weight, \( \nabla_w \), is

\[
\nabla_w \left( w^\top \hat{\Sigma} w \right) = 2 \hat{\Sigma} w.
\]

Moreover, this gradient evaluated at the starting \( 1/N \) portfolio is

\[
\nabla_w \left( w^\top \hat{\Sigma} w \right) \bigg|_{e/N} = 2 \hat{\Sigma} e/N.
\]

Hence, \( -\hat{\Sigma} e/N \) is the portfolio that when combined with the \( 1/N \) portfolio yields the largest marginal decrease in the portfolio sample variance.\(^{12}\) But note that this portfolio is not a zero-cost portfolio; that is, its weights do not add up to zero in general. The properties of the gradient then imply that the first conjugate portfolio is simply the zero-cost portfolio that is closest (in 2-norm) to the portfolio \( -\hat{\Sigma} e/N \). The properties of Euclidean spaces and some algebra then imply that the first conjugate portfolio, \( w_{CG1} \), is

\[
w_{CG1} = - \left( I - ee^\top/N \right) \hat{\Sigma} e/N,
\]

because the matrix \( (I - ee^\top/N) \) is the projection matrix that projects any portfolio into the set of portfolios with zero-cost.

The first partial minimum-variance portfolio is the combination of the \( 1/N \) portfolio and the first conjugate portfolio that minimizes the sample portfolio return variance; that is, the first partial

---

\(^{12}\)Note that any scalar multiple of the vector \( -2\hat{\Sigma} e/N \) yields the largest marginal decrease in sample variance when combined with \( 1/N \). For notational convenience, we use \( -\hat{\Sigma} e/N \) instead of \( -2\hat{\Sigma} e/N \).
minimum-variance portfolio is

\[ w_{PAR1} = e/N + \alpha_0 w_{CG1}, \tag{27} \]

in which \( \alpha_0 \) is chosen to minimize the sample variance of the first partial minimum-variance portfolio. We now explain the intuition underlying this portfolio. First, in Proposition 4, we show that the \( i^{th} \) component of the vector \( \hat{\Sigma}_{N} \) is the sample covariance of the return on the \( 1/N \) portfolio with the return on the \( i^{th} \) risky asset.

**Proposition 4** The \( i^{th} \) component of the vector \( \hat{\Sigma}_{N} \) is the sample covariance between the \( 1/N \) portfolio and the return on the \( i^{th} \) risky asset.

Then, in Proposition 5, we show that the \( i^{th} \) component of the first conjugate portfolio is the negative of the deviation of this covariance from the average of the covariances between the \( 1/N \) portfolio and each of the individual assets.

**Proposition 5** Let \( \sigma_{ei} \) be the covariance between the \( 1/N \) portfolio return and the \( i^{th} \) risky asset return. Then, the \( i^{th} \) component of the first conjugate portfolio, \( (w_{CG1})_i \), is

\[ (w_{CG1})_i = -\left( \sigma_{ei} - \frac{1}{N} \sum_{j=1}^{N} \sigma_{ej} \right). \tag{28} \]

This result makes sense: the first conjugate portfolio is a zero-cost portfolio that assigns a positive weight to those assets whose covariance with the return of the \( 1/N \) portfolio is below average. As a result, given that \( \alpha_0 > 0 \), the first partial minimum-variance portfolio assigns a weight larger than \( 1/N \) to those assets whose covariances with the \( 1/N \) portfolio are below average. This improves the diversification of the portfolio, and thus, decreases the portfolio variance. Also, note that the minimum-variance portfolio is the portfolio whose covariance with all other assets is equal to a constant (see Huang and Litzenberger (1988, Chapter 3, Proposition 12)). Thus, according to the interpretation above, the first partial minimum-variance portfolio should be “closer” to the shortsale-unconstrained minimum-variance portfolio than the \( 1/N \) portfolio.
3.4.2 Relation to the 2-Norm-Constrained Portfolios

The following proposition shows that although the partial minimum-variance portfolios are not obtained by imposing explicitly a constraint on the 2-norm of the portfolio, they lead to portfolios whose 2-norm is indeed smaller than that of the shortsale-unconstrained minimum-variance portfolio.

**Proposition 6** The 2-norm of the $k$th partial minimum-variance portfolio is smaller or equal than the 2-norm of the shortsale-unconstrained minimum-variance portfolio for $k \leq N - 1$.

To demonstrate that the partial minimum-variance portfolios provide a discrete first-order approximation to the 2-norm-constrained portfolios, we show in the following proposition that the first conjugate-gradient portfolio is tangential to the continuously differentiable curve described by the 2-norm-constrained portfolios at the $1/N$ portfolio.

**Proposition 7** Provided the matrix $\hat{\Sigma}$ is nonsingular, the 2-norm-constrained portfolios form a continuously differentiable curve for $\delta$ ranging from $1/N$ to $\|w_{MINU}\|_2^2$, and the first conjugate portfolio, $w_{CG1} = -(I - \frac{e e^T}{N})\hat{\Sigma} \frac{e}{N}$, is tangential to this curve at $\delta = 1/N$.

Figure 5 depicts the partial minimum-variance portfolios together for the case with three risky assets. The set of 2-norm-constrained portfolios is depicted as in Figure 4. Figure 5 also depicts the first partial minimum-variance portfolio, $w_{PAR}$, which is the combination of the $1/N$ portfolio and the first conjugate portfolio $w_{CG1}$ that minimizes the portfolio variance: $w_{PAR} = e/N + \alpha_0 w_{CG1}$. Note that the first conjugate portfolio, $w_{CG1}$, is tangent to the curve of 2-norm-constrained portfolios at the equally-weighted, $1/N$ portfolio—this was formally shown in Proposition 7 and it demonstrates that the first conjugate portfolio, $w_{CG1}$, is a first-order approximation to the curve of 2-norm-constrained portfolios. The second conjugate portfolio, $w_{CG2}$, is not orthogonal to the first conjugate portfolio, but the product $w_{CG1} \hat{\Sigma} w_{CG2}$ is zero. Finally, because this is a case with only three risky assets, the second partial minimum-variance portfolio already coincides with the shortsale-unconstrained minimum-variance portfolio, $w_{MINU}$. 

23
3.5 A Bayesian Interpretation of the Norm-Constrained Portfolios

Tibshirani (1996, Section 5) gives a Bayesian interpretation for the regression-analysis techniques of “least absolute shrinkage and selection operator” (lasso) and “ridge” regressions. Here we adapt his analysis to give a Bayesian interpretation of the 1- and 2-norm-constrained minimum-variance portfolios.

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

Proposition 8 Assume that the investor’s prior belief for each of the shortsale-unconstrained minimum-variance portfolio weights is independently and identically distributed as a double Exponential distribution with probability density function:

$$\pi(w_i) = \frac{\nu}{2} e^{-\nu |w_i|}, \quad (29)$$

and the variance of the minimum-variance portfolio return, denoted by $\sigma^2$, has an independent prior, $\pi(\sigma^2)$. Then, there exists a threshold parameter $\delta$ such that the corresponding 1-norm-constrained minimum-variance portfolio maximizes the posterior likelihood of the investor.

The next proposition shows that the 2-norm-constrained portfolio is the Bayesian portfolio of an investor whose prior belief is that the portfolio weights are independently and identically distributed as a Normal distribution.

Proposition 9 Assume that the investor’s prior belief for each of the shortsale-unconstrained minimum-variance portfolio weights is independently and identically distributed as a Normal distribution with probability density function:

$$\pi(w) = \sqrt{\nu} \frac{1}{\pi} e^{-\nu w^2}, \quad (30)$$

and the minimum-variance portfolio return variance, denoted by $\sigma^2$, has an independent prior, $\pi(\sigma^2)$. Then, there exists a threshold parameter $\delta$ such that the corresponding 2-norm-constrained minimum-variance portfolio maximizes the posterior likelihood of the investor.
Intuitively, the explanation for these propositions is that when the investor’s prior belief is that the portfolio weights follow a Double-Exponential distribution, then the prior likelihood of a portfolio $w$ is proportional to its 1-norm $\|w\|_1$. On the other hand, when the investor’s prior belief is that portfolio weights follow a Normal distribution, then the prior likelihood of a portfolio $w$ is proportional to its 2-norm $\|w\|_2$.

### 3.6 A Moment-Shrinkage Interpretation of the Norm-Constrained Portfolios

The 1- and 2-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 10** Let the solution to the 1-norm-constrained minimum-variance problem (13)–(15) be such that $(w_{NC1})_i \neq 0$ for $i = 1, \ldots, N$. 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

$$
\hat{\Sigma}_{NC1} = \hat{\Sigma} - \nu ne^\top - \nu en^\top,
$$

in which $\nu \in \mathbb{R}$ 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 10 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 (31) shows that the 1-norm-constrained portfolios can be seen as the result of shrinking by the constant amount $\nu$ the covariances of those assets that are being sold short with all the other assets. Note that the amount of shrinkage $\nu$ 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 (6) 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 $\lambda_i$, and these Lagrange multipliers may take different values for the different assets.
The following proposition gives a moment-shrinkage interpretation for the 2-norm-constrained portfolios. Concretely, it shows that the 2-norm-constrained portfolios can be obtained by shrinking all elements of the sample covariance matrix towards the elements of the identity matrix.

**Proposition 11** Provided $\hat{\Sigma}$ is nonsingular, for each $\delta > 1/N$ there exists a $\nu \geq 0$ such that the solution to the 2-norm-constrained minimum-variance problem (20)–(22) coincides with the solution to the shortsale-unconstrained minimum-variance problem (1)–(2) if the sample covariance matrix, $\hat{\Sigma}$, is replaced by the matrix

$$
\Sigma_{NC2} = \left( \frac{1}{1 + \nu} \right) \hat{\Sigma} + \left( \frac{\nu}{1 + \nu} \right) I.
$$

(32)

4 Out-of-Sample Evaluation of the Proposed Portfolios

We have carried out an extensive comparison of the out-of-sample empirical performance of the 1-norm-constrained, 2-norm-constrained, and partial minimum-variance portfolios to nine portfolios from the existing literature across five different datasets using three performance metrics. In this section, we give a summary of the main insights that emerge from this empirical analysis, and a more detailed discussion is given in Appendix C.\(^{13}\)

Note that to use the 1- and 2-norm constrained minimum-variance portfolios one needs to choose the value of the threshold parameter $\delta$, 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 $\delta$ and $k$ could be specified exogenously. But, in our general framework, these can also be calibrated to achieve a particular objective and/or to exploit a particular feature of the returns data. We use two different criteria to calibrate the norm-constrained portfolios: (i) minimizing the *portfolio variance* via cross validation, and (ii) maximizing the last period *portfolio return* in order

\(^{13}\)In particular, the appendix contains the following sections. Section C.1 describes two methods that can be used to calibrate the norm-constrained portfolios. The various portfolios considered in our experiments are described in Section C.2, the datasets across which performance is evaluated are described in Section C.3, and the methodology used to compare performance is explained in Section C.4. Finally, the detailed results of this comparison are reported in Section C.5.
to exploit positive autocorrelation in portfolio returns, as opposed to autocorrelation in the return of individual securities.\textsuperscript{14}

We compare the performance of the three norm-constrained portfolios, which are listed in Panel A of Table 1, to that of nine portfolios from the existing literature, which are listed in Panel B of Table 1. The first two portfolios in Panel B are simple benchmarks that require neither estimation nor optimization. Namely, the $1/N$ portfolio and the value-weighted market portfolio. We also consider two portfolios that rely on estimates of mean returns. These are the traditional mean-variance portfolio and the Bayesian mean-variance portfolio, which is selected using the approach in Jorion (1985, 1986). We consider these portfolios only for completeness because there is extensive empirical evidence showing that portfolios that rely on estimates of mean returns are usually outperformed by portfolios that ignore estimates of expected returns—even when the Sharpe ratio or other performance measures that rely on both the mean and variance are used for the comparison (see Jorion (1985, 1986, 1991), Jagannathan and Ma (2003), DeMiguel, Garlappi, and Uppal (2007), and DeMiguel and Nogales (2007)). The next three portfolios are variants of the minimum-variance portfolio. Concretely, we consider the traditional shortsale-unconstrained minimum-variance portfolio, the shortsale-constrained minimum-variance portfolio that is analyzed in Jagannathan and Ma (2003), and the minimum-variance portfolio with shrinkage of the covariance matrix as in Ledoit and Wolf (2004b). We consider also a 1-factor model with the market being the factor. Finally, we consider also the parametric portfolios proposed in Brandt, Santa-Clara, and Valkanov (2005).

Our results are reported in three tables. Table 3 gives the out-of-sample portfolio variance for the different strategies, together with the P-value for the difference between the variance of each portfolio and that of the shortsale-unconstrained minimum-variance portfolio. Table 4 gives the out-of-sample Sharpe ratio for the various portfolio strategies, along with the P-value for the differ-

\textsuperscript{14}Our motivation for the portfolio autocorrelation criterion for calibration is the work by Campbell, Lo, and MacKinley (1997), who report that: “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.” In particular, Campbell, Lo, and MacKinley (1997, Panel C of Table 2.4) shows that the return in the last month explains 17% of the variability on the return of the $1/N$ portfolio. Note that the $1/N$ portfolio is just one of the extremes of the set of portfolios we produce with the norm-constrained or partial minimum-variance portfolios, and thus, this feature of the data explains why calibrating the norm-constrained portfolios to maximize the portfolio return in the last period may improve the performance of the portfolio out of sample.
ence between the Sharpe ratio of each portfolio and that of the shortsale-unconstrained minimum-variance strategy. Table 5 reports the turnover of the various strategies.

From Panel A of Table 3, we see that the out-of-sample variance for the three norm-constrained portfolios calibrated using cross-validation over the return variance (NC1V, NC2V, PARV) is similar to each other across the five datasets. And, not surprisingly, the out-of-sample variance is lower for these policies than for those that are calibrated using the criterion of maximizing the return of the portfolio in the previous period (NC1R, NC2R, PARR). Comparing the variances of the portfolios in Panels A and B of Table 3, we see that typically the norm-constrained portfolios have lower out-of-sample variances than the portfolios from the existing literature. For example, the 1-norm-constrained portfolios (NC1V) generally have lower variance than the shortsale-constrained minimum-variance portfolios (MINC), which they nest. Similarly, the 2-norm-constrained portfolio (NC2V) and its discrete first-order approximation (PARV), outperform the Ledoit and Wolf (2004b) portfolio (MINL), which they nest.

Table 4 reports the Sharpe ratio of the various strategies. Panel A of this table shows that of the three norm-constrained strategies, the partial minimum-variance portfolio calibrated by maximizing the portfolio return in the last period (PARR) usually has a higher Sharpe ratio than the 1- and 2-norm-constrained portfolios, NC1R and NC2R. Comparing the Sharpe ratios of the norm-constrained portfolios in Panel A to the portfolios from the existing literature listed in Panel B, we see that the three norm-constrained portfolios have higher Sharpe ratios than both the equally- weighted (1/N) and the value-weighted (VW) portfolios for all datasets, and the difference is substantial in most cases. The difference in performance is even more striking when the norm-constrained policies are compared to the traditional mean-variance (MEAN) and the Bayesian-mean-variance strategies (BAYE). The norm-constrained policies typically outperform also the minimum-variance constrained portfolio, MINC, and the Ledoit and Wolf (2004b) portfolio, MINL. The norm-constrained portfolio PARR has higher Sharpe ratios compared to also the portfolios based on the 1-factor market model (FAC1) for all the datasets. Finally, even though the norm-constrained portfolios do not use firm-specific characteristics, they are able to achieve Sharpe ratios
that are at least as good as those for the parametric portfolios (BSV) developed in Brandt, Santa-Clara, and Valkanov (2005).

We now discuss the results on turnover. From Table 5, we see that the norm-constrained portfolios calibrated using cross-validation over the portfolio variance have much lower turnover compared to the portfolios calibrated by maximizing the portfolio return over the last month. Not surprisingly, the best portfolios in terms of turnover are the $1/N$ and value-weighted portfolios. The turnover of these portfolios is followed by that of the shortsale-constrained minimum-variance portfolio (MINC). The turnover of the 1- and 2-norm constrained portfolios and the partial minimum-variance portfolio calibrated with cross-validation over variance is higher than that of the minimum-variance portfolio with shortsale constraints. The shortsale-unconstrained minimum-variance portfolio, the Ledoit and Wolf (2004b) 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, Santa-Clara, and Valkanov (2005) have similar turnovers, which are much higher than those of the rest of the portfolios.

5 Conclusion

We conclude by summarizing the main contributions of our paper. One, we provide a general framework for determining portfolios in the presence of estimation error. This framework is based on shrinking directly the portfolio weights 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. Two, we show that our general framework nests as special cases the shrinkage approaches of Jagannathan and Ma (2003), Ledoit and Wolf (2004b), and the $1/N$ portfolio studied in DeMiguel, Garlappi, and Uppal (2007). Three, we show that all these portfolios can be interpreted as those of a Bayesian investor who has a certain prior belief on portfolio weights rather than moments of asset returns. Four, we use our general framework to extend portfolio strategies in the existing literature and to develop several new
portfolio strategies that have not been considered in the literature before. Five, we illustrate how
the norm-constrained portfolio strategies can be calibrated in order to improve their performance.
Finally, we compare empirically using three performance metrics—portfolio variance, Sharpe ratio,
and turnover—the out-of-sample performance of the new polices we have proposed to nine strategies
in the existing literature across five datasets. We find that the new portfolios we propose can
outperform the ones studied in Jagannathan and Ma (2003), Ledoit and Wolf (2004b), and the
$1/N$ portfolio evaluated in DeMiguel, Garlappi, and Uppal (2007), and that their performance
is similar to that of other strategies proposed in the literature such as Brandt, Santa-Clara, and
Valkanov (2005) even though the strategies we develop do not rely on firm-specific characteristics.

We hope that developing a generalized framework that unifies and extends existing methods for
dealing with estimation error when choosing portfolios will spur researchers in finance to propose
new methods for determining portfolios that perform well out of sample. And, we hope that showing
that this generalized framework is related to problems in optimization, statistics and chemometrics,
will spur researchers in areas outside finance to work on the important problem of portfolio selection
in the presence of estimation error.
References


Sloan School of Management, MIT.
Lauprete, G., A. Samarov, and R. Welsch, 2002, “Robust Portfolio Optimization,” Metrika, 25,
139–149.
Ledoit, O., and M. Wolf, 2004a, “Honey, I Shrank the Sample Covariance Matrix,” Journal of
Portfolio Management, 30, 110–119.
Ledoit, O., and M. Wolf, 2004b, “A Well-Conditioned Estimator for Large-Dimensional Covariance
of Finance, 7, 77–91.
Memmel, C., 2003, “Performance Hypothesis Testing with the Sharpe Ratio,” Finance Letters, 1,
21–23.
Analysts Journal, 45, 31–42.
Phatak, A., and F. De Hoog, 2003, “Exploiting the Connection Between PLS, Lanczos, and Con-
jugate Gradients: Alternative Proofs of Some Properties of PLS,” Journal of Chemometrics,
16, 361–367.
Steinhaug, T., 1983, “The Conjugate Gradient Method and Trust Regions in Large Scale Optimiza-
Tibshirani, R., 1996, “Regression Shrinkage and Selection via the Lasso,” Journal of the Royal
Statistical Society, B, 58, 267–288.
University.
of Financial Studies, 18, 673–705.
Forthcoming in Revstat.
University of Zurich.
A Appendix: Details of the Partial Minimum-Variance Portfolios

In this appendix we give a detailed discussion of the method used to compute the partial minimum-variance portfolios. Because the conjugate-gradient method can be used to solve only unconstrained optimization problems, we first show in Section A.1 how to eliminate the constraint that the portfolio weights sum to one \((w^\top e = 1)\). Then, in Section A.2, we give a detailed statement of the algorithm used to compute partial minimum-variance portfolios.

A.1 Expressing the Minimum-Variance Problem Without the Adding-Up Constraint

The conjugate-gradient method can be applied only to unconstrained problems, and thus, to apply it to the minimum-variance problem we first need to eliminate the constraint \(w^\top e = 1\). To do so, we will use of the following proposition that shows that any portfolio whose weights sum up to one can be rewritten as the \(1/N\) portfolio plus a portfolio whose weights sum up to zero, herein a zero-cost portfolio.

**Proposition A1** Let \(w \in \mathbb{R}^N\) and \(w^\top e = 1\), then the following holds:

1. The portfolio \(w\) may be expressed as

   \[ w = \frac{e}{N} + w_0, \]  
   \( \text{in which } w_0 \text{ is a zero-cost portfolio; that is, } w_0^\top e = 0. \)

2. There exists a matrix \(Z \in \mathbb{R}^{N \times (N-1)}\) whose columns form an orthonormal basis for the subspace of zero cost portfolios (that is, an orthonormal basis for the subspace of all portfolios satisfying \(w^\top e = 0\)) and this matrix satisfies \(e^\top Z = 0\) and \(Z^\top Z = I\).

3. There exists a vector \(w_Z \in \mathbb{R}^{N-1}\) such that

   \[ w = \frac{e}{N} + Zw_Z. \]  

(A2)
Proof: Part 1: Let \( w_0 \) be the portfolio defined as \( w_0 = w - e/N \). Then note that if the weights of the portfolio \( w \) sum up to one, \( w^\top e = 1 \), then the weights of the portfolio \( w_0 \) must sum up to zero because \( e^\top w_0 = e^\top w - e^\top e/N = 0 \).

Part 2: The existence of an orthonormal basis for the subspace of zero-cost portfolios follows from general vector space theory (Nocedal and Wright (1999)). Let \( Z \in \mathbb{R}^{N\times(N-1)} \) be the matrix whose columns are the vectors composing the orthonormal basis for the zero-cost portfolio subspace, then because all zero-cost portfolios are orthogonal to \( e \), we must have \( e^\top Z = 0 \). Finally, because the columns of \( Z \) are orthogonal to each other and their 2-norm is equal to one, we have that \( Z^\top Z = I \).

Part 3: The result follows from Part 1 because \( w_0 \) is a zero-cost portfolio and \( Z \) is a basis of the zero-cost portfolio subspace. ■

We define portfolio space to be the \( \mathbb{R}^N \) space, zero-cost portfolio subspace the subspace in \( \mathbb{R}^N \) that is formed by all linear combinations of the columns of \( Z \), and reduced space the \( N-1 \) dimensional space \( \mathbb{R}^{N-1} \), which \( w_Z \) inhabits. We illustrate these concepts in Figure 6, which shows these spaces for the case with two risky assets. The portfolio space is simply the \( \mathbb{R}^2 \) space in which the plot lives, the zero-cost subspace is the straight line passing through the origin and labeled as \( w_0 \), and the reduced space in this case is the set of real numbers \( \mathbb{R} \). The \( 1/N \) portfolio is the vector \( w = (1/2, 1/2)^\top \). The matrix \( Z \) has only one column equal to the vector \( (1/\sqrt{2}, -1/\sqrt{2})^\top \). The portfolio that assigns a weight of zero to the first asset and one to the second asset can then be written as

\[
(0, 1)^\top = \frac{e}{N} + Z w_Z = (1/2, 1/2)^\top - (1/\sqrt{2}, -1/\sqrt{2})^\top \times (1/\sqrt(2)).
\] (A3)

The following proposition shows that expression (A2) may be used to eliminate the constraint from the minimum-variance problem.

**Proposition A2** The sample minimum-variance portfolio can be written as \( w_{MINU} = \frac{e}{N} + Z w_Z \), where \( w_{MINU} \) is the solution to the unconstrained minimum-variance problem:

\[
\min_{w_Z} \left( \frac{e}{N} + Z w_Z \right)^\top \Sigma \left( \frac{e}{N} + Z w_Z \right).
\] (A4)
Proof: Substituting (A1) into (1)-(2) yields the result.

A.2 Applying the Conjugate-Gradient Method

To compute the partial minimum-variance portfolios, we apply the conjugate-gradient method to solve the first-order optimality conditions for the unconstrained minimum-variance problem (A4):

\[ Z^\top \hat{\Sigma} Z w = -Z^\top \hat{\Sigma} \frac{e}{N}. \]

Given a portfolio \( w = \frac{e}{N} + Z w_Z \), the residual of the first-order conditions is defined as

\[ \epsilon = -Z^\top \hat{\Sigma} \frac{e}{N} - Z^\top \hat{\Sigma} Z w_Z. \]  

(A5)

Note that the residual is zero only for the minimum-variance portfolio, \( w_{MINU} \), and thus the residual serves as a measure of the distance to the solution. For reasons given above, we use the \( 1/N \) portfolio as the starting point for the conjugate-gradient method. The method commences by finding the zero-cost portfolio \( w_{CG0} \) that induces the maximum marginal decrease in the sample variance when combined with the \( 1/N \) portfolio. We term this portfolio the first conjugate portfolio \( w_{CG0} \), and it can be shown (Nocedal and Wright (1999)) that it is equal to the residual at the \( 1/N \) portfolio, which is our starting point:

\[ w_{CG0} = \epsilon^0 = -Z^\top \hat{\Sigma} \frac{e}{N}. \]  

(A6)

Then the conjugate-gradient method finds the combination of the \( 1/N \) portfolio and the first conjugate-gradient portfolio that minimizes the sample portfolio variance. We term the resulting portfolio \( (w_{PAR1}) \) the first partial minimum-variance portfolio, where the term “partial” refers to the fact that the portfolio minimizes the sample variance within the subset of portfolios formed by combinations of the \( 1/N \) portfolio and the first conjugate portfolio.

One may feel tempted to follow the same strategy in the successive iterations, and hence find the zero-cost portfolio that would induce the maximum marginal decrease in sample variance when combined with \( (w_{PAR1}) \), but it can be shown (Nocedal and Wright (1999)) that it is more efficient to find the zero-cost portfolio \( w_{CG1} \) that induces the maximum marginal decrease in sample variance,
subject to the condition that it is *conjugate* with respect to \( w_{CG0} \); that is, subject to the condition

\[ w_{CG1} \tilde{\Sigma} w_{CG0} = 0, \]

which simply implies that these two portfolios must be uncorrelated with respect to the sample covariance matrix. Then the method finds the optimal combination of \( w_{PAR1} \) and \( w_{CG1} \), which is termed the second partial minimum-variance portfolio. If this process is iterated \( N - 1 \) times, we recover the minimum-variance portfolio. If instead, we only perform this operation a number \( K < N - 1 \) times, we obtain a portfolio that lies between the benchmark portfolio and the minimum-variance portfolio.

We now give a brief mathematical description of the algorithm followed to compute the partial minimum-variance portfolios. For a rigorous treatment of the conjugate-gradient method see Nocedal and Wright (1999, Chapter 5). Let \( \epsilon^k = -Z^\top \tilde{\Sigma} e/N - Z^\top \tilde{\Sigma} Z w_k^k \) denote the residual at the \( k \)th partial minimum-variance portfolio \( w_{PARk} = e/N + Z w_k^k \). Select the number of conjugate portfolios \( K \leq N \), let \( w_{MINU}^0 = e/N \) and \( k = 0 \). The \((k + 1)\)st partial minimum-variance portfolio is defined as \( w_{PARk+1} = w_{PARk} + \alpha_k w_{CGk} \), in which \( w_{CGk} = Z(\epsilon^k + \beta^k w_{CGk-1}) \) is the \( k \)th conjugate portfolio, for \( k = 0 \) we set \( \beta^k = 0 \) and for \( k > 0 \) we set

\[
\beta^k = \frac{\epsilon^k}{(\epsilon^{k-1})^\top \epsilon^{k-1}},
\]

and \( \alpha_k = \frac{(w_{CGk})^\top \epsilon^k}{(w_{CGk})^\top \tilde{\Sigma} \Delta w_{CGk}} \) is the \( k \)th conjugate portfolio step-size.
B Appendix: Proofs for All the Propositions

Proof for Proposition 1

Note that the solution to the norm-constrained problem must satisfy \( \|w_{NC}\| \leq \delta \). Hence, if \( \delta < \|w_{MINU}\| \), we must have \( \|w_{NC}\| < \|w_{MINU}\| \).

Proof for Proposition 2

Note that the 1-norm constrained problem (13)–(15) with \( \delta = 1 \) can be obtained from the shortsale-constrained problem by replacing the shortsale constraints \( w \geq 0 \) with the 1-norm-constraint \( \|w\|_1 \leq 1 \). Hence, to prove the result it suffices to show that these two constraints are equivalent. To see this, note that the 1-norm can be rewritten as in (16). Moreover, if the portfolio weights sum up to one \( (w^T e = 1) \), we have (17). From (16) and (17), we obtain (18), in which by definition \( \sum_{i \in \mathcal{N}(w)} w_i < 0 \), except if the index set \( \mathcal{N}(w) \) is empty. Therefore, if the portfolio weights sum up to one, the constraint \( \|w\|_1 \leq 1 \) is satisfied if and only if \( \mathcal{N}(w) \) is empty, which occurs if and only if \( w \geq 0 \).

Proof for Proposition 3

If \( \delta \geq 1/N \) then there exists a feasible point for problem (20)–(22). If, in addition, \( \hat{\Sigma} \) is nonsingular, then there is a unique global minimizer to problem (20)–(22) and this unique minimizer is characterized by the first-order optimality conditions, which imply that there exist Lagrange multipliers \( \phi \) and \( \nu \) such that

\[
2\hat{\Sigma}w + 2\nu Iw - \psi e = 0, \quad \text{(B1)}
\]

\[
w^T e = 1, \quad \text{(B2)}
\]

\[
\sum_{i=1}^{N} w_i^2 \leq \delta, \quad \text{(B3)}
\]

\[
\nu \geq 0. \quad \text{(B4)}
\]
These conditions imply that the first-order optimality conditions hold for the minimum-variance problem with $\hat{\Sigma}_{LW} = \hat{\Sigma} + \nu I$. Moreover, because $\hat{\Sigma}$ is nonsingular and positive definite and $\nu \geq 0$, the matrix $\hat{\Sigma}_{LW}$ is also positive definite and the first-order optimality conditions for the minimum-variance problem with $\hat{\Sigma}_{LW}$ are then also sufficient for optimality.

Proof for Proposition 4

The sample covariance between two portfolios $w_1$ and $w_2$ is $w_1^\top \hat{\Sigma} w_2$. The $i^{th}$ asset is simply the portfolio that assigns a weight of one to the $i^{th}$ asset and zero to all the other assets; $w_i = (0, \ldots, 0, 1, 0, \ldots, 0)^\top$. Then the product $w_i^\top \hat{\Sigma} e / N$ is the covariance between the $i^{th}$ asset and the $1/N$ portfolio. Finally note that $w_i^\top \hat{\Sigma} e / N$ is simply the $i^{th}$ component of the vector $\hat{\Sigma} e / N$.

Proof for Proposition 5

From (26) we know that the first conjugate portfolio is

$$w_{CG1} = -(I - e e^\top / N) \hat{\Sigma} e / N = -(I - e e^\top / N) \sigma_e,$$

in which $\sigma_e = (\sigma_{e1}, \sigma_{e2}, \ldots, \sigma_{eN})^\top$. This, together with the fact that all components of the $i^{th}$ row of $(I - e e^\top / N)$ are equal to $-1/N$ except the $i^{th}$ component which is equal to $1 - 1/N$, prove the result.

Proof for Proposition 6

Phatak and De Hoog (2003) and Steihaug (1983) use properties of the conjugate-gradient method to show that when started from a starting point equal to zero ($w^0_Z = 0$), the 2-norm of the iterates generated by the method is strictly nondecreasing. This implies that $\|w^k_Z\|_2$ is strictly nondecreasing. Also, note that $w_{PARk} = e / N + Z w^k_Z$. Moreover, because $e$ is orthogonal to the columns of $Z$, we have that $\|w_{PARk}\|_2^2 = \|e / N\|_2^2 + \|Z w^k_Z\|_2^2$. Because $\|w^k_Z\|_2$ is strictly nondecreasing and by definition the columns of $Z$ are orthogonal and of 2-norm one, we then have that $\|w_{PARk}\|_2$ is strictly nondecreasing. This together with the fact that $w_{PARN} = w_{MINU}$ yields the result.
Proof for Proposition 7

Using the unconstrained minimum-variance problem (A4), we can rewrite the 2-norm-constrained minimum-variance problem as

$$\min_{w_Z} \left( \frac{e}{N} + Zw_Z \right)^\top \hat{\Sigma} \left( \frac{e}{N} + Zw_Z \right)$$ \hspace{1cm} (B6)

s.t. \( \|e/N\|_2^2 + \|w_Z\|_2^2 \leq \delta. \) \hspace{1cm} (B7)

Moreover, because \( \|e/N\|_2^2 = 1/N \) we can then rewrite this problem as

$$\min_{w_Z} \left( \frac{e}{N} + Zw_Z \right)^\top \hat{\Sigma} \left( \frac{e}{N} + Zw_Z \right)$$ \hspace{1cm} (B8)

s.t. \( \|w_Z\|_2^2 \leq \tilde{\delta}, \) \hspace{1cm} (B9)

in which \( \tilde{\delta} = \delta - 1/N. \) Also, provided the matrix \( \hat{\Sigma} \) is nonsingular, then we also have that \( Z^\top \hat{\Sigma} Z \) is nonsingular, and hence the solution to problem (B8)–(B9) is unique and is the unique solution to the first-order optimality conditions for problem (B8)–(B9). Then following an argument similar to that in the proof of Proposition 3, we have that, provided the matrix \( \hat{\Sigma} \) is nonsingular, for each \( \delta \) there exists a \( \nu \geq 0 \) such that the solution to the unique 2-norm-constrained problem coincides with the unique solution to the following problem:

$$\min_{w_Z} \left( \frac{e}{N} + Zw_Z \right)^\top \hat{\Sigma} \left( \frac{e}{N} + Zw_Z \right) + \nu w_Z^\top w_Z,$$ \hspace{1cm} (B10)

in which \( w_Z^\top w_Z = \|w_Z\|_2^2. \) Moreover, the uniqueness of the minimizer to this problem ensures that the curve of 2-norm-constrained portfolios for different values of \( \delta \) or \( \nu \) is continuously differentiable (see Fiacco and McCormick (1968, Theorem 14)). The first-order conditions give the unique solution to problem: (B10)

$$w_Z = H^{-1}g,$$ \hspace{1cm} (B11)

in which \( H \) is the Hessian matrix of the objective function in problem (B10),

$$H = Z^\top \hat{\Sigma} Z + \nu I,$$ \hspace{1cm} (B12)

and \( g \) is the gradient of the objective function in problem (B10),

$$g = \frac{e}{N} + Zw_Z.$$ \hspace{1cm} (B13)
We are interested in the tangent to the curve of solutions to the 2-norm-constrained problem. This is given by the differential of expression (B11) with respect to \( \nu \). We are interested in this differential at the 1/N portfolio; that is, for \( w_Z = 0 \). We then have that

\[
w_Z = -(Z^\top \hat{\Sigma} Z + \nu I)^{-1} \frac{e}{N},
\]

and thus we only need to compute

\[
\frac{d}{d\nu} (Z^\top \hat{\Sigma} Z + \nu I)^{-1} = \lim_{d\nu \to 0} \frac{(Z^\top \hat{\Sigma} Z + \nu I + d\nu I)^{-1} - (Z^\top \hat{\Sigma} Z + \nu I)^{-1}}{d\nu} = -I,
\]

in which the last equality follows from the fact that a first-order approximation to \((Z^\top \hat{\Sigma} Z + \nu I + d\nu I)^{-1}\) is given by \((Z^\top \hat{\Sigma} Z + \nu I)^{-1} - d\nu I\). Hence,

\[
\frac{dw_Z}{d\nu} = Z \hat{\Sigma} \frac{e}{N},
\]

and thus,

\[
\frac{d w}{d\nu} = ZZ \hat{\Sigma} \frac{e}{N},
\]

which is the negative of the first conjugate portfolio.

Proof for Proposition 8

Because the prior distributions for each of the portfolio weights are independent, the prior distribution for a portfolio \( w \) is simply

\[
\pi(w) = \prod_{i=1}^{N} \frac{\nu}{2} e^{-\nu |w_i|} = \frac{\nu^N}{2^N} e^{-\nu \|w\|_1}.
\]

Then the posterior portfolio weight distribution conditional on the sample returns, \( \{r_t\}_{t=1}^{T} \) can be expressed as follows

\[
\pi(w, \sigma^2 | \{r_t\}_{t=1}^{T}) \propto \frac{\pi(\sigma^2)}{\sigma^{(N-1)}} \exp\left(-\frac{\sum_{t=1}^{T} (w^\top r_t - w^\top \hat{\mu})^2}{2\sigma^2} - \nu \|w\|_1\right),
\]

where \( \pi(\sigma^2) \) is the prior distribution of \( \sigma^2 \).
where the symbol $\propto$ means that the distribution is proportional to the right-hand-side term. Hence, the portfolio that maximizes the posterior distribution subject to the condition that the portfolio weights add up to one is the solution to the following optimization problem:

$$\min_w w^\top \hat{\Sigma} w + \rho \|w\|_1, \quad (B20)$$

$$\text{s.t. } w^\top e = 1, \quad (B21)$$

where $w^\top \hat{\Sigma} w = \sum_{t=1}^T (w^\top r_t - w^\top \hat{\mu})^2 / (T-1)$ and $\rho = 2\sigma^2\nu / (T-1)$. Moreover, it follows from optimization theory (Nocedal and Wright, 1999, Chapter 17) that there exists a threshold parameter $\delta$ such that the solution to problem (B20)–(B21) coincides with the 1-norm-constrained minimum-variance portfolio for the threshold $\delta$.

\[\blacksquare\]

**Proof for Proposition 9**

Because the prior distributions for each of the portfolio weights are independent, the prior distribution for a portfolio $w$ is simply

$$\pi(w) = \prod_{i=1}^N \sqrt{\frac{\nu}{\pi}} e^{-\nu(w_i)^2} \propto e^{-\nu \|w\|_2^2}. \quad (B22)$$

Then the posterior portfolio weight distribution conditional on the sample returns, $\{r_t\}_{t=1}^T$ can be expressed as follows

$$\pi(w, \sigma^2 | \{r_t\}_{t=1}^T) \propto \frac{\pi(\sigma^2)}{\sigma^{(N-1)}} \exp\left( -\frac{\sum_{t=1}^T (w^\top r_t - w^\top \hat{\mu})^2}{2\sigma^2} - \nu \|w\|_2^2 \right), \quad (B23)$$

where the symbol $\propto$ means that the distribution is proportional to the right-hand-side term. Hence, the portfolio that maximizes the posterior distribution subject to the condition that the portfolio weights add up to one is the solution to the following optimization problem:

$$\min_w w^\top \hat{\Sigma} w + \rho \|w\|_2^2, \quad (B24)$$

$$\text{s.t. } w^\top e = 1, \quad (B25)$$

where $w^\top \hat{\Sigma} w = \sum_{t=1}^T (w^\top r_t - w^\top \hat{\mu})^2 / (T-1)$ and $\rho = 2\sigma^2\nu / (T-1)$. Moreover, it follows from optimization theory (Nocedal and Wright, 1999, Chapter 17) that there exists a threshold parameter
\( \delta \) such that the solution to problem (B24)--(B25) coincides with the 2-norm-constrained minimum-variance portfolio for the threshold \( \delta \).

\[ \text{Proof for Proposition 10} \]

The 1-norm-constrained minimum-variance problem is
\[
\min_w \ w^\top \hat{\Sigma} w, \\
\text{s.t.} \quad w^\top e = 1, \\
\sum_{i=1}^{N} |w_i| \leq \delta.
\]

Using (18), we can rewrite this problem as
\[
\min_w \ w^\top \hat{\Sigma} w, \\
\text{s.t.} \quad w^\top e = 1, \\
1 - 2 \sum_{i \in \mathcal{N}(w)} w_i \leq \delta,
\]
in which \( \mathcal{N}(w) \) is the set of asset indexes for which the corresponding portfolio weight is negative, \( \mathcal{N}(w) = \{i : w_i < 0\} \). 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
\[
\min_w \ w^\top \hat{\Sigma} w - 2\nu \sum_{i \in \mathcal{N}(w)} w_i, \\
\text{s.t.} \quad w^\top e = 1.
\]

Moreover, because at the solution \( w^\top e = 1 \), the problem can be equivalently rewritten as
\[
\min_w \ w^\top (\hat{\Sigma} - \nu ne^\top - \nu n^\top e) w, \\
\text{s.t.} \quad w^\top e = 1,
\]
which completes the proof for the proposition.
Proof for Proposition 11

Proposition 3 shows that the 2-norm-constrained portfolios coincide with the shortsale-unconstrained portfolios obtained by replacing the sample covariance matrix by the matrix $\hat{\Sigma} + \nu I$, in which $\nu$ is the Lagrange multiplier for the 2-norm constraint. Moreover, note that multiplying this matrix by a scalar does not change the solution to the corresponding minimum-variance problem. As a result, the 2-norm-constrained portfolios can also be obtained by replacing the sample covariance matrix by the following matrix:

$$\Sigma_{NC2} = \left( \frac{1}{1+\nu} \right) \hat{\Sigma} + \left( \frac{\nu}{1+\nu} \right) I.$$
C Appendix: Details of Out-of-Sample Evaluation of Portfolios

In this appendix, we compare the out-of-sample performance of the 1-norm-constrained, 2-norm-constrained, and partial minimum-variance portfolios that we have developed in this paper to nine portfolio strategies in the existing literature. This performance is evaluated for five different datasets using three performance metrics. To ensure that this appendix is self-contained, we repeat some of the discussion in Section 4.

In Section C.1, we describe two methods that can be used to calibrate the norm-constrained portfolios. We describe the various portfolios considered in our experiments in Section C.2, the datasets across which performance is evaluated are described in Section C.3, and the methodology used to compare performance is explained in Section C.4. Finally, the results of this comparison are reported in Section C.5.

C.1 Calibration of the Norm-Constrained Portfolios

Note that for the 1- and 2-norm constrained minimum-variance portfolios we have developed in Sections 3.2 and 3.3, one needs to choose the value of the threshold parameter $\delta$, which bounds the maximum value that the portfolio norm may take. Similarly, for the partial minimum-variance portfolios, we need to choose the order parameter $k$ that indicates how many of the $N - 1$ partial minimum-variance portfolios to use. The parameters $\delta$ and $k$ could be specified exogenously. But, in our general framework, these can also be calibrated to achieve a particular objective and/or to exploit a particular feature of the returns data. Below, we describe the two different criteria we use to calibrate the norm-constrained portfolios: (i) minimizing the portfolio variance, and (ii) maximizing the last period portfolio return in order to exploit positive autocorrelation in portfolio returns, as opposed to autocorrelation in the return of individual securities.

Our motivation for the portfolio autocorrelation criterion for calibration is the work by Campbell, Lo, and MacKinley (1997), who report that: “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.” In particular, Campbell, Lo, and MacKinley (1997, Panel C of Table 2.4) shows that the return in the last month explains 17% of the variability on the return of the $1/N$ portfolio. Note that the $1/N$ portfolio is just one of the extremes of the set of portfolios we produce with the norm-constrained or partial minimum-variance portfolios, and thus, this feature of the data explains why calibrating the norm-constrained portfolios to maximize the portfolio return in the last period may improve the performance of the portfolio out of sample.

We first describe how to calibrate the norm-constrained portfolios when the objective is to minimize the “out-of-sample” variance of the portfolio return. To do this, we use the nonparametric technique known as cross validation—see Efron and Gong (1983) and Campbell, Lo, and MacKinley (1997, Section 12.3.2). Cross validation works as follows. Given an estimation window composed of $\tau$ sample asset returns, for each $t$ ranging from 1 to $\tau$ 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 dataset 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. For the 1- and 2-norm-constrained portfolios, we compute $(w_\theta)(t)$ by solving problem (8)–(10) with $\hat{\Sigma}$ replaced by $\hat{\Sigma}(t)$ and $\theta = \delta$; and, for the partial minimum-variance portfolios, we compute $(w_\theta)(t)$ by applying the conjugate-gradient method for $\theta = k$ iterations with $\hat{\Sigma}$ replaced by $\hat{\Sigma}(t)$. Third, compute the “out-of-sample” return attained by this portfolio on the $t^{th}$ sample asset return $(r_\theta)(t) = ((w_\theta)(t)^\top r_t)$. Then the variance of the “out-of-sample” portfolio return is given by the sample variance of the $\tau$ out-of-sample returns, $(r_\theta)(t)$; that is,

$$
\hat{\sigma}_\theta^2 = \frac{\sum_{t=1}^{\tau} [(r_\theta)(t) - (\bar{r}_\theta)(t)]^2}{\tau - 1},
$$

(C26)

in which $(\bar{r}_\theta)(t) = \sum_{t=1}^{\tau} (r_\theta)(t) / \tau$. Finally, choose the parameter $\theta^*$ that minimizes this “out-of-sample” return variance,\(^\text{15}\) that is,

$$
\theta^* = \text{argmin}_{\theta} \hat{\sigma}_\theta^2.
$$

(C27)

\(^\text{15}\)From a computational perspective, this implies computing $\hat{\sigma}_\theta^2$ for different values of $\theta$ and then choosing the value of $\theta$ that minimizes the out-of-sample variance. Alternatively, one can use Newton’s method to compute the optimal $\theta$ more efficiently.
The second criterion we use for calibrating $\theta$ is based on maximizing the portfolio return over the last period. The choice of this objective is motivated by the desire to exploit the positive autocorrelation in portfolio returns documented in Campbell, Lo, and MacKinley (1997). In this case, we choose $\theta^*$ so that

$$\theta^* = \arg\max_\theta \ w_\theta^\top r_\tau,$$

in which $r_\tau$ is the asset-return vector for the last period within the estimation window, and $w_\theta$ is the norm-constrained or partial minimum-variance portfolio computed using all the data over the estimation window with the parameter $\theta$. That is, we choose the parameter $\theta = \{\delta, k\}$ to maximize the return in the last period within the estimation window. We consider the return in only the last period because this is where the autocorrelation is highest.

C.2 Description of the Portfolios Considered

We evaluate the three norm-constrained portfolios we have developed in this paper, which are listed in Panel A of Table 1, by comparing their performance to the nine portfolio strategies taken from the existing literature that are listed in Panel B of Table 1. In this section, we describe briefly the three norm-constrained portfolios and the nine portfolios that serve as benchmarks.

In Panel A of Table 1, the first portfolio is the 1-norm-constrained portfolio obtained by solving problem (13)–(15). As explained above, the threshold parameter $\delta$ is calibrated in two ways: using cross validation over the portfolio return variance and maximization of the portfolio return in the last period. The next portfolio in Panel A is the 2-norm-constrained portfolio obtained by solving problem (20)–(22); again, the threshold parameter $\delta$ is calibrated in the two ways described above. The third portfolio in Panel A is the partial minimum-variance portfolio. We report the performance of the partial minimum-variance portfolios where $k$, the number of steps of the conjugate-gradient method, is calibrated as before—by cross-validation over the portfolio-return variance and by maximization of the portfolio return in the previous period.

In Panel B of Table 1, we list the nine benchmark portfolios used for evaluating the performance of the norm-constrained portfolios. The first two portfolios in Panel B are simple benchmarks that
require neither estimation nor optimization. Namely, the $1/N$ portfolio and the value-weighted market portfolio.\footnote{We compute the value-weighted portfolio for each dataset as the portfolio that assigns a weight to each asset equal to the market capitalization of that asset divided by the market capital of all other assets in the dataset. 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.}

We also consider two portfolios that rely on estimates of mean returns. These are the traditional mean-variance portfolio and the Bayesian mean-variance portfolio, which is selected using the approach in Jorion (1985, 1986). We consider these portfolios only for completeness because there is extensive empirical evidence showing that portfolios that rely on estimates of mean returns are usually outperformed by portfolios that ignore estimates of expected returns—even when the Sharpe ratio or other performance measures that rely on both the mean and variance are used for the comparison (see Jorion (1985, 1986, 1991), Jagannathan and Ma (2003), DeMiguel, Garlappi, and Uppal (2007), and DeMiguel and Nogales (2007)).

The next three portfolios are variants of the minimum-variance portfolio. Concretely, we consider the traditional shortsale-unconstrained minimum-variance portfolio, the shortsale-constrained minimum-variance portfolio that is analyzed in Jagannathan and Ma (2003), and the minimum-variance portfolio with shrinkage of the covariance matrix as in Ledoit and Wolf (2004b).

We consider also four minimum-variance portfolios based on factor models, but we report the results for only one of these. The portfolio whose performance we report is based on a 1-factor model with the market being the factor. The other three portfolios we consider are based on factor models that use as factors the principal components associated with the largest eigenvalues of the sample covariance matrix. Specifically, we consider 1-, 2-, and 3-factor models based on the principal components associated with the largest, the two largest, and the three largest eigenvalues, respectively, of the sample covariance matrix. The results for these portfolio strategies are not reported because they are not superior to that of the portfolio based on the 1-factor market model.

Finally, we consider also the parametric portfolios proposed in Brandt, Santa-Clara, and Valkanov (2005). These portfolios make use of firm-specific characteristics, size, book-to-market ratio, and momentum of each individual asset, to choose the weight to assign to each asset. These
portfolios are computed using a risk-aversion parameter $\gamma = 5$, as in Brandt, Santa-Clara, and Valkanov.

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, Section III) find that their performance is similar to that of the minimum-variance portfolio with shortsale constraints based on monthly returns. Also, we do not consider portfolios based on constant correlation models because these portfolios are outperformed by those proposed in Ledoit and Wolf (2004b). Regarding factor models, we report results only for a one-factor model with the market portfolio as the factor. Our motivation for considering this factor model is that Chan, Karceski, and Lakonishok (1999, p. 955) 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 one-factor market model. As Chan, Karceski, and Lakonishok report: “Since the role of the remaining factors is obscured by the major factor, there appears to be little benefit from obtaining finer breakdowns of the systematic component of the volatility.” Finally, we consider imposing shortselling constraints only to the minimum-variance portfolio because Jagannathan and Ma (2003, 1653–4) find: “for the factor models and shrinkages estimators, imposing such constraints is likely to hurt.”

C.3 Description of the Empirical Datasets Considered

We compare across five datasets the performance of the portfolios we have developed to those in the existing literature. These five datasets are listed in Table 2. As a summary, the datasets include: (1) 10 industry portfolios representing the U.S. stock market (10Ind); (2) 48 industry portfolios representing the U.S. stock market (48Ind); (3) 6 Fama and French portfolios of firms sorted by size and book to market (6FF); (4) 25 Fama and French portfolios of firms sorted by size and book to market (25FF); and (5) A dataset of CRSP returns (500CRSP) constructed in a way that is similar to Jagannathan and Ma (2003), with monthly rebalancing. The 500CRSP dataset is formed as follows. In April of each year we randomly select 500 assets among all assets in the CRSP dataset 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. The various portfolio policies are rebalanced every month.

C.4 Description of the Methodology Used to Evaluate Performance

We compare the performance of the norm-constrained portfolios to the portfolios in the existing 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 dataset. For our experiments, we use an estimation window of \( \tau = 120 \) data points, which for monthly data corresponds to ten years.\(^{17}\) Two, using the return data over the estimation window, \( \tau \), we compute the various portfolios. Three, 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 dataset is reached. At the end of this process, we have generated \( T - \tau \) portfolio-weight vectors for each strategy; that is, \( w_t^k \) for \( t = \tau, \ldots, T - 1 \) and for each strategy \( k \).

Following this “rolling horizon” methodology, holding the portfolio \( w_t^k \) for one month gives the out-of-sample return at time \( t + 1 \): \( r_{t+1}^k = w_t^k \top r_{t+1} \), where \( r_{t+1} \) denotes the returns. After collecting the time series of \( T - \tau \) returns \( r_t^k \), the out-of-sample mean, variance, and Sharpe ratio of returns are,

\[
\hat{\mu}_k = \frac{1}{T - \tau} \sum_{t=\tau}^{T-1} w_t^k \top r_{t+1},
\]

\[
(\hat{\sigma}_k^2) = \frac{1}{T - \tau - 1} \sum_{t=\tau}^{T-1} \left( w_t^k \top r_{t+1} - \hat{\mu}_k \right)^2,
\]

\[
\hat{\text{SR}}_k = \frac{\hat{\mu}_k}{\hat{\sigma}_k}
\]

\(^{17}\)We have tried other estimation window lengths such as \( \tau = 60 \) and 240 but the results are similar, and thus, we report the results only for the case \( \tau = 120 \).
To measure the statistical significance of the difference in the variance and Sharpe ratio of a particular strategy from that of the shortsale-unconstrained minimum-variance strategy that serves as benchmark, we report also the P-values for the variance and Sharpe ratios of each strategy relative to this strategy. Because we are interested in finite-sample properties of the proposed techniques, instead of using the test by Jobson and Korkie (1981), with the correction by Memmel (2003), we use the bootstrapping methodology described in Efron and Tibshirani (1993).

Finally, we wish to obtain a measure of portfolio turnover. Let \( w_{j,t}^k \) denote the portfolio weight in asset \( j \) chosen at time \( t \) under strategy \( k \), \( w_{j,t+1}^k \) the portfolio weight before rebalancing but at \( t+1 \), and \( w_{j,t+1}^k \) the desired portfolio weight at time \( t+1 \) (after rebalancing). Then, turnover, which is the average percentage of wealth traded in each month, is defined as 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:

\[
\text{Turnover} = \frac{1}{T - \tau - 1} \sum_{t=\tau}^{T-1} \sum_{j=1}^{N} \left( |w_{j,t+1}^k - w_{j,t}^k| \right).
\]

(C32)

C.5 Discussion of the Out-of-Sample Performance

In this section, we compare across the five datasets listed in Table 2 the performance of the norm-constrained portfolios listed in Panel A of Table 1 to that of the strategies from the existing literature, which are listed in Panel B of Table 1. The results are reported in the following four tables. Table 3 gives the out-of-sample portfolio variance for the different strategies, together with the P-value for the difference between the variance of each portfolio and that of the shortsale-unconstrained minimum-variance portfolio. Table 4 gives the out-of-sample Sharpe ratio for the various portfolio strategies, along with the P-value for the difference between the Sharpe ratio of

---

18 Specifically, consider two portfolios \( i \) and \( n \), with \( \mu_i, \sigma_i, \sigma_n \) as their true means and variances. We wish to test the hypothesis that the Sharpe ratio of portfolio \( i \) is equal to that of portfolio \( n \), that is, \( H_0 : \mu_i/\sigma_i - \mu_n/\sigma_n = 0 \). To do this, we obtain \( B \) pairs of size \( T - \tau \) of the portfolio returns \( i \) and \( n \) by resampling with replacement. If \( \hat{F} \) denotes the empirical distribution function of the \( B \) bootstrap pairs corresponding to \( \hat{\mu}_i/\hat{\sigma}_i - \hat{\mu}_n/\hat{\sigma}_n \), then a two-sided P-value for the previous null hypothesis is given by \( \hat{p} = 2\hat{F}(0) \). In a similar way, to test the hypothesis that the variances of two portfolio returns are identical, \( H_0 : \sigma_i^2/\sigma_n^2 = 1 \). If \( \hat{F} \) denotes the empirical distribution function of the \( B \) bootstrap pairs corresponding to \( \hat{\sigma}_i^2/\hat{\sigma}_n^2 \), then a two-sided P-value for this null hypothesis is given by \( \hat{p} = 2\hat{F}(0) \). For a nice discussion of the application of other bootstrapping methods to test the significance of Sharpe ratios, see Wolf (2007).
each portfolio and that of the shortsale-unconstrained minimum-variance strategy. Table 5 reports
the turnover of the various strategies. The average values of the threshold parameters for the three
norm-constrained portfolio that we consider across the various datasets are reported in Table 6.
The first column of each of these tables lists the various portfolio strategies considered. Each of
the other five columns contains the results for each of the five datasets.

From Panel A of Table 3, we see that the out-of-sample variance for the three norm-constrained
portfolios calibrated using cross-validation over the return variance (NC1V, NC2V, PARV) is similar
to each other across the five datasets. And, not surprisingly, the out-of-sample variance is lower
for these policies than for those that are calibrated using the criterion of maximizing the return of
the portfolio in the previous period (NC1R, NC2R, PARR).

Comparing the norm-constrained portfolios in Panel A to the portfolios from the existing lit-
erature listed in Panel B of Table 3, we see that the out-of-sample portfolio variance of the 1/N
portfolio and the value-weighted (VW) market portfolio is significantly higher than that of the
norm-constrained portfolios calibrated using cross-validation over the return variance. Compar-
ing the norm-constrained portfolios to the traditional mean-variance and Bayesian mean-variance
policies, we see that the difference is even more striking; for all the datasets, the variance of the
norm-constrained strategies is at least one order of magnitude smaller than that for the mean-
variance strategies.

Comparing the norm-constrained portfolios to the minimum-variance portfolios that they nest,
we see that the 1-norm-constrained portfolio (NC1V) generally has a lower variance than the
shortsale-constrained minimum-variance portfolio (MINC). For example, for the 48Ind dataset the
variance of the NC1V policy is 0.00126, while that for the minimum-variance constrained (MINC)
policy is 0.00133 and for the minimum-variance unconstrained (MINU) policy is 0.00186. Simi-
larly, for the 500CRSP dataset, the variance of the NC1V policy is 0.00074, while that for the
minimum-variance constrained (MINC) policy is 0.00087 and for the minimum-variance uncon-
strained (MINU) policy is 0.00104. Similarly, the 2-norm-constrained portfolio, NC2V, and its dis-
crete first-order approximation, PARV, outperform the Ledoit and Wolf (2004b) portfolio, MINL,
which they nest. For example, for the 48Ind dataset, the variance of the NC2V policy is 0.00137 and
for the PARV policy it is 0.00141, while for the Ledoit and Wolf policy (MINL) it is 0.00185. Note also that the out-of-sample variance of the three norm-constrained portfolios calibrated using cross-validation over the portfolio variance (NC1V, NC2V, PARV) is smaller than or equal to the variance of the shortsale-unconstrained minimum-variance portfolio (MINU) for almost every dataset; however, this is not true for the shortsale-constrained minimum-variance portfolio (MINC), which has a higher out-of-sample variance than the shortsale-unconstrained minimum-variance portfolio for the 6FF and 25FF datasets.

The out-of-sample variance for the three norm-constrained portfolios (NC1V, NC2V, and PARV) is also smaller than that for the portfolios based on the 1-factor model (FAC1) for all datasets. For instance, for the 10Ind dataset the NC1V policy has a variance of 0.00134, while for FAC1 it is 0.00145. Similarly, for the 25FF dataset the NC1V policy has a variance of 0.00135, while for FAC1 it is 0.00240. The norm-constrained portfolios also attain a significantly lower variance than the Brandt, Santa-Clara, and Valkanov (2005) parametric portfolios (BSV), which of course are not designed to minimize variance. Overall, the results in Table 3 indicate that the norm-constrained portfolios have lower out-of-sample portfolio variance compared to the portfolio strategies in the existing literature.

Table 4 reports the Sharpe ratio of the various strategies. Panel A of this table shows that of the three norm-constrained strategies, the partial minimum-variance portfolio calibrated using cross-validation over the portfolio autocorrelation (PARR) usually has a higher Sharpe ratio than the 1- and 2-norm-constrained portfolios, NC1R and NC2R. For instance, for the 48Ind dataset, the Sharpe ratio for the PARR strategy is 0.3166, while for NC1R it is 0.2831 and for NC2R it is 0.2891. Similarly, for the 500CRSP dataset, the Sharpe ratio for the PARR strategy is 0.4768, while for NC1R it is 0.3706 and for NC2R it is 0.4672.

Comparing the Sharpe ratios of the norm-constrained portfolios in Panel A to the portfolios from the existing literature listed in Panel B, we see that the three norm-constrained portfolios have higher Sharpe ratios than both the equally-weighted (1/N) and the value-weighted (VW) portfolios for all datasets, and the difference is substantial in most cases. For example, for the 500CRSP dataset, the Sharpe ratio for the PARR strategy is 0.4768, while for 1/N it is 0.3326 and
for VW it is 0.2748. The difference in performance is even more striking when the norm-constrained policies are compared to the traditional mean-variance (MEAN) and the Bayesian-mean-variance strategies (BAYE): for the 500CRSP dataset, the Sharpe ratio for the PARR strategy is 0.4768, while for MEAN it is 0.0723 and for BAYE it is 0.4018.

The norm-constrained portfolios also have a higher Sharpe ratio than the benchmark shortsale-unconstrained portfolio (MINU). For example, for the 500CRSP dataset, the Sharpe ratio for the PARR strategy is 0.4768, while for MINU it is 0.3820. The norm-constrained policies typically outperform also the minimum-variance constrained portfolio, MINC, and the Ledoit and Wolf (2004b) portfolio, MINL. For instance, for the 500CRSP dataset, the Sharpe ratio for the PARR strategy is 0.4768, while for MINC it is 0.3985 and for MINL it is 0.4028. The norm-constrained portfolio PARR has higher Sharpe ratios compared to also the portfolios based on the 1-factor market model (FAC1) for all the datasets. For example, for the 500CRSP dataset, the Sharpe ratio for FAC1 is only 0.4166 compared to the 0.4768 for the PARR strategy.

Finally, even though the norm-constrained portfolios do not use firm-specific characteristics, they are able to achieve Sharpe ratios that are at least as good as those for the parametric portfolios (BSV) developed in Brandt, Santa-Clara, and Valkanov (2005).

We now discuss the turnover results in Table 5. From Panel A of this table, we see that the norm-constrained portfolios calibrated using cross-validation over the portfolio variance have much lower turnover compared to the portfolios calibrated by maximizing the portfolio return over the last month. That is, NC1V has lower turnover than NC1R, NC2V has lower turnover than NC2R, and PARV has lower turnover than PARR. And, within the set of portfolios calibrated by maximizing last period’s return, the 1-norm-constrained portfolio (NC1R) has a lower turnover than the 2-norm-constrained (NC2R) and partial minimum-variance portfolios (PARR).

Comparing the norm-constrained portfolios in Panel A of Table 5 to the portfolios in Panel B, it is not surprising to see that the best portfolios in terms of turnover are the $1/N$ and value-weighted portfolios. The turnover of these portfolios is followed by that of the shortsale-constrained minimum-variance portfolio (MINC); the traditional mean-variance portfolio has the highest turnover and the Bayesian-mean-variance portfolio also has relatively high turnover. The turnover of the
1- and 2-norm constrained portfolios and the partial minimum-variance portfolio calibrated with cross-validation over variance is higher than that of the minimum-variance portfolio with short-sale constraints. The shortsale-unconstrained minimum-variance portfolio, the Ledoit and Wolf (2004b) portfolio, and the portfolios based on factor models have higher turnover than MINC and the norm-constrained strategies calibrated to 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, Santa-Clara, and Valkanov (2005) have similar turnovers, which are much higher than those of the rest of the portfolios.

We conclude by comparing the threshold values of the norm-constraints obtained from the two different calibration methods considered: cross-validation to minimize the portfolio variance and maximization of the last-period portfolio return. From Table 6, we see that the threshold values obtained by using cross-validation to minimize portfolio variance are generally larger than those obtained by maximizing the portfolio return in the last period. That is, the threshold numbers for NC1V are higher than for NC1R, for NC2V are higher than for NC2R, and for PARV are higher than for PARR. The larger threshold implies that the norm-constrained portfolios calibrated to minimize the variance are closer to the shortsale-unconstrained minimum-variance portfolio, which minimizes the variance in sample. On the other hand, the lower threshold of the norm-constrained portfolios calibrated to maximize the return in the last period implies that these portfolios are closer to the \(1/N\) portfolio (for the case of the 2-norm-constrained and partial minimum-variance portfolios) and closer to the shortsale-constrained minimum-variance portfolios (for the case of the 1-norm-constrained portfolios). Both the \(1/N\) portfolio and the shortsale-constrained minimum-variance portfolio have been shown to attain a relatively high out-of-sample Sharpe ratio by DeMiguel, Garlappi, and Uppal (2007).
Table 1: List of Portfolios Considered

This table lists the various portfolio strategies we consider. Panel A of the table lists the norm-constrained portfolios developed in this paper, while Panel B lists the portfolios from the existing literature. Note that $\delta$ is the threshold parameter that limits the norm of the portfolio-weight vector, while the order parameter $k$ indicates the number of steps of the conjugate-gradient method to take, that determines which of the $N-1$ partial minimum-variance portfolios to use. The last column of the table gives the abbreviation that we use to refer to the strategy in the tables in which we compare the performance of the norm-constrained portfolios to the portfolios in the existing literature. A detailed description of these portfolios is given in Section C.2 on page 47.

<table>
<thead>
<tr>
<th>#</th>
<th>Model</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-norm-constrained minimum-variance portfolio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• With $\delta$ calibrated using cross-validation over portfolio variance</td>
<td>NC1V</td>
</tr>
<tr>
<td></td>
<td>• With $\delta$ calibrated by maximizing portfolio return in previous period</td>
<td>NC1R</td>
</tr>
<tr>
<td>2</td>
<td>2-norm-constrained minimum-variance portfolio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• With $\delta$ calibrated using cross-validation over portfolio variance</td>
<td>NC2V</td>
</tr>
<tr>
<td></td>
<td>• With $\delta$ calibrated by maximizing portfolio return in previous period</td>
<td>NC2R</td>
</tr>
<tr>
<td>3</td>
<td>Partial minimum-variance portfolios</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• With $k$ calibrated using cross-validation over portfolio variance</td>
<td>PARV</td>
</tr>
<tr>
<td></td>
<td>• With $k$ calibrated by maximizing the portfolio return in the previous period</td>
<td>PARR</td>
</tr>
<tr>
<td></td>
<td><strong>Panel A: Portfolio strategies developed in this paper</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Equally-weighted (1/$N$) portfolio</td>
<td>1/N</td>
</tr>
<tr>
<td>2</td>
<td>Value-weighted (market) portfolio</td>
<td>VW</td>
</tr>
<tr>
<td></td>
<td><strong>Simple benchmarks</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Mean-variance portfolio with shortsales unconstrained</td>
<td>MEAN</td>
</tr>
<tr>
<td>4</td>
<td>Bayesian mean-variance portfolio using the approach in Jorion (1985, 1986)</td>
<td>BAYE</td>
</tr>
<tr>
<td></td>
<td><strong>Portfolios that use mean returns</strong></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Minimum-variance portfolio with shortsales unconstrained</td>
<td>MINU</td>
</tr>
<tr>
<td>6</td>
<td>Minimum-variance portfolio with shortsales constrained</td>
<td>MINC</td>
</tr>
<tr>
<td>7</td>
<td>Minimum-variance portfolio with covariance matrix as in Ledoit and Wolf (2004b)</td>
<td>MINL</td>
</tr>
<tr>
<td></td>
<td><strong>Minimum-variance portfolios that ignore mean returns</strong></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Minimum-variance portfolio with the market as the single factor</td>
<td>FAC1</td>
</tr>
<tr>
<td>9</td>
<td>Brandt, Santa-Clara, and Valkanov (2005) strategy with a risk-aversion parameter of $\gamma = 5$ using the factors Size, Book-to-Market, and Momentum</td>
<td>BSV</td>
</tr>
<tr>
<td></td>
<td><strong>Portfolios based on a factor model and parametric portfolios</strong></td>
<td></td>
</tr>
</tbody>
</table>
This table lists the various datasets analyzed, the abbreviation used to refer to each dataset in the tables reporting the performance of the various portfolio strategies, the number of risky assets $N$ in each dataset, the time period spanned by the dataset, and the source of the data. The 500CRSP dataset is formed as follows. In April of each year we randomly select 500 assets among all assets in the CRSP dataset 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. The various portfolio policies are rebalanced every month.

<table>
<thead>
<tr>
<th>#</th>
<th>Dataset</th>
<th>Abbreviation</th>
<th>$N$</th>
<th>Time Period</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ten industry portfolios</td>
<td>10Ind</td>
<td>10</td>
<td>07/1963–12/2004</td>
<td>K. French</td>
</tr>
<tr>
<td>2</td>
<td>Forty eight industry portfolios</td>
<td>48Ind</td>
<td>48</td>
<td>07/1963–12/2004</td>
<td>K. French</td>
</tr>
<tr>
<td>3</td>
<td>6 Fama and French portfolios</td>
<td>6FF</td>
<td>6</td>
<td>07/1963–12/2004</td>
<td>K. French</td>
</tr>
<tr>
<td>5</td>
<td>500 randomized stocks from CRSP</td>
<td>500CRSP</td>
<td>500</td>
<td>04/1968–04/2005</td>
<td>CRSP</td>
</tr>
<tr>
<td></td>
<td>balanced monthly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Portfolio Variances

The quantities reported in this table are the monthly out-of-sample variances for the portfolios listed in Table 1, and the corresponding P-value that the portfolio variance for that strategy is different from that for the benchmark shortsale-unconstrained minimum-variance strategy. These quantities are computed for the datasets listed in Table 2.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>10Ind</th>
<th>48Ind</th>
<th>6FF</th>
<th>25FF</th>
<th>500CRSP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Portfolio strategies developed in this paper</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC1V</td>
<td>0.00134</td>
<td>0.00126</td>
<td>0.00156</td>
<td>0.00135</td>
<td>0.00074</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.00)</td>
<td>(0.98)</td>
<td>(0.28)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>NC1R</td>
<td>0.00138</td>
<td>0.00135</td>
<td>0.00159</td>
<td>0.00143</td>
<td>0.00080</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.00)</td>
<td>(0.70)</td>
<td>(0.84)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>NC2V</td>
<td>0.00134</td>
<td>0.00137</td>
<td>0.00156</td>
<td>0.00130</td>
<td>0.00066</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.00)</td>
<td>(0.79)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>NC2R</td>
<td>0.00149</td>
<td>0.00176</td>
<td>0.00163</td>
<td>0.00152</td>
<td>0.00087</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.25)</td>
<td>(0.18)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>PARV</td>
<td>0.00138</td>
<td>0.00141</td>
<td>0.00159</td>
<td>0.00133</td>
<td>0.00065</td>
</tr>
<tr>
<td></td>
<td>(0.71)</td>
<td>(0.00)</td>
<td>(0.21)</td>
<td>(0.07)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>PARR</td>
<td>0.00153</td>
<td>0.00163</td>
<td>0.00161</td>
<td>0.00146</td>
<td>0.00085</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.40)</td>
<td>(0.74)</td>
<td>(0.01)</td>
</tr>
<tr>
<td><strong>Panel B: Portfolio strategies from existing literature</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.00179</td>
<td>0.00221</td>
<td>0.00230</td>
<td>0.00249</td>
<td>0.00169</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.15)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>VW</td>
<td>0.00158</td>
<td>0.00190</td>
<td>0.00191</td>
<td>0.00186</td>
<td>0.00157</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.90)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>MEAN</td>
<td>0.01090</td>
<td>0.38107</td>
<td>0.00353</td>
<td>0.00942</td>
<td>0.00626</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>BAYE</td>
<td>0.00264</td>
<td>0.06793</td>
<td>0.00221</td>
<td>0.00400</td>
<td>0.00066</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>MINU</td>
<td>0.00138</td>
<td>0.00186</td>
<td>0.00156</td>
<td>0.00143</td>
<td>0.00104</td>
</tr>
<tr>
<td></td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>MINC</td>
<td>0.00134</td>
<td>0.00133</td>
<td>0.00186</td>
<td>0.00176</td>
<td>0.00087</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>MINL</td>
<td>0.00138</td>
<td>0.00185</td>
<td>0.00156</td>
<td>0.00143</td>
<td>0.00066</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.31)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>FAC1</td>
<td>0.00145</td>
<td>0.00159</td>
<td>0.00201</td>
<td>0.00240</td>
<td>0.00075</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.04)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>BSV</td>
<td>0.00602</td>
<td>0.00392</td>
<td>0.00306</td>
<td>0.00344</td>
<td>0.00574</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>
Table 4: Portfolio Sharpe Ratios

The quantities reported in this table are the monthly out-of-sample Sharpe ratios for the portfolios listed in Table 1, and the corresponding P-value that the Sharpe ratio for each of these strategies is different from that for the benchmark shortsale-unconstrained minimum-variance strategy. These quantities are computed for the datasets listed in Table 2.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>10Ind</th>
<th>48Ind</th>
<th>6FF</th>
<th>25FF</th>
<th>500CRSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Portfolio strategies developed in this paper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC1V</td>
<td>0.2854</td>
<td>0.2886</td>
<td>0.3385</td>
<td>0.3649</td>
<td>0.4013</td>
</tr>
<tr>
<td></td>
<td>(0.91 )</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>NC1R</td>
<td>0.2890</td>
<td>0.2831</td>
<td>0.3374</td>
<td>0.3553</td>
<td>0.3706</td>
</tr>
<tr>
<td></td>
<td>(0.78 )</td>
<td>(0.02)</td>
<td>(0.17)</td>
<td>(0.03)</td>
<td>(0.85)</td>
</tr>
<tr>
<td>NC2V</td>
<td>0.2919</td>
<td>0.2855</td>
<td>0.3527</td>
<td>0.4089</td>
<td>0.3994</td>
</tr>
<tr>
<td></td>
<td>(0.48 )</td>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.47)</td>
<td>(0.68)</td>
</tr>
<tr>
<td>NC2R</td>
<td>0.3193</td>
<td>0.2891</td>
<td>0.3922</td>
<td>0.4278</td>
<td>0.4672</td>
</tr>
<tr>
<td></td>
<td>(0.21 )</td>
<td>(0.02)</td>
<td>(0.31)</td>
<td>(0.73)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>PARV</td>
<td>0.2841</td>
<td>0.2823</td>
<td>0.3478</td>
<td>0.4077</td>
<td>0.3937</td>
</tr>
<tr>
<td></td>
<td>(0.75 )</td>
<td>(0.01)</td>
<td>(0.07)</td>
<td>(0.50)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>PARR</td>
<td>0.3293</td>
<td>0.3166</td>
<td>0.3912</td>
<td>0.4403</td>
<td>0.4768</td>
</tr>
<tr>
<td></td>
<td>(0.10 )</td>
<td>(0.00)</td>
<td>(0.29)</td>
<td>(0.48)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Panel B: Portfolio strategies from existing literature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.2541</td>
<td>0.2508</td>
<td>0.2563</td>
<td>0.2565</td>
<td>0.3326</td>
</tr>
<tr>
<td></td>
<td>(0.42 )</td>
<td>(0.50)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>VW</td>
<td>0.2619</td>
<td>0.2698</td>
<td>0.2437</td>
<td>0.2558</td>
<td>0.2748</td>
</tr>
<tr>
<td></td>
<td>(0.49 )</td>
<td>(0.24)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>MEAN</td>
<td>0.0499</td>
<td>-0.0334</td>
<td>0.3214</td>
<td>0.2253</td>
<td>0.0723</td>
</tr>
<tr>
<td></td>
<td>(0.00 )</td>
<td>(0.01)</td>
<td>(0.37)</td>
<td>(0.01)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>BAYES</td>
<td>0.1685</td>
<td>-0.0121</td>
<td>0.3666</td>
<td>0.3151</td>
<td>0.4018</td>
</tr>
<tr>
<td></td>
<td>(0.00 )</td>
<td>(0.04)</td>
<td>(0.99)</td>
<td>(0.12)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>MINU</td>
<td>0.2865</td>
<td>0.2222</td>
<td>0.3640</td>
<td>0.4199</td>
<td>0.3820</td>
</tr>
<tr>
<td></td>
<td>(1.00 )</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>MINC</td>
<td>0.2852</td>
<td>0.2914</td>
<td>0.2629</td>
<td>0.2720</td>
<td>0.3985</td>
</tr>
<tr>
<td></td>
<td>(0.94 )</td>
<td>(0.04)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>MINL</td>
<td>0.2865</td>
<td>0.2224</td>
<td>0.3640</td>
<td>0.4200</td>
<td>0.4028</td>
</tr>
<tr>
<td></td>
<td>(0.23 )</td>
<td>(0.36)</td>
<td>(0.31)</td>
<td>(0.76)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>FAC1</td>
<td>0.3060</td>
<td>0.2674</td>
<td>0.2485</td>
<td>0.2486</td>
<td>0.4166</td>
</tr>
<tr>
<td></td>
<td>(0.31 )</td>
<td>(0.18)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>BSV</td>
<td>0.1157</td>
<td>0.3314</td>
<td>0.3908</td>
<td>0.4047</td>
<td>0.2674</td>
</tr>
<tr>
<td></td>
<td>(0.00 )</td>
<td>(0.06)</td>
<td>(0.57)</td>
<td>(0.81)</td>
<td>(0.38)</td>
</tr>
</tbody>
</table>
Table 5: Portfolio Turnovers

The quantities reported in this table are the monthly turnovers for the portfolios listed in Table 1. This turnover is the average percentage of wealth traded in each period, and as defined in Equation (C32) 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. These quantities are computed for the datasets listed in Table 2.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>10Ind</th>
<th>48Ind</th>
<th>6FF</th>
<th>25FF</th>
<th>500CRSP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Portfolio strategies developed in this paper</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC1V</td>
<td>0.1494</td>
<td>0.2680</td>
<td>0.1729</td>
<td>0.2407</td>
<td>0.6141</td>
</tr>
<tr>
<td>NC1R</td>
<td>0.6013</td>
<td>0.8232</td>
<td>1.0064</td>
<td>0.9767</td>
<td>0.9753</td>
</tr>
<tr>
<td>NC2V</td>
<td>0.1448</td>
<td>0.3266</td>
<td>0.1946</td>
<td>0.4570</td>
<td>0.5808</td>
</tr>
<tr>
<td>NC2R</td>
<td>1.0177</td>
<td>2.7556</td>
<td>1.6594</td>
<td>3.6275</td>
<td>1.0443</td>
</tr>
<tr>
<td>PARV</td>
<td>0.1689</td>
<td>0.3838</td>
<td>0.2600</td>
<td>0.4628</td>
<td>0.5743</td>
</tr>
<tr>
<td>PARR</td>
<td>1.0414</td>
<td>2.4846</td>
<td>1.6407</td>
<td>3.5657</td>
<td>1.0984</td>
</tr>
<tr>
<td><strong>Panel B: Portfolio strategies from existing literature</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/N</td>
<td>0.0232</td>
<td>0.0311</td>
<td>0.0155</td>
<td>0.0174</td>
<td>0.0595</td>
</tr>
<tr>
<td>VW</td>
<td>0.0382</td>
<td>0.0540</td>
<td>0.0213</td>
<td>0.0310</td>
<td>0.2316</td>
</tr>
<tr>
<td>MEAN</td>
<td>1.0135</td>
<td>105.6126</td>
<td>0.7987</td>
<td>4.2495</td>
<td>3.0014</td>
</tr>
<tr>
<td>BAYES</td>
<td>0.3565</td>
<td>6.6314</td>
<td>0.5388</td>
<td>2.1264</td>
<td>0.6191</td>
</tr>
<tr>
<td>MINU</td>
<td>0.1656</td>
<td>0.8286</td>
<td>0.2223</td>
<td>0.7953</td>
<td>0.7769</td>
</tr>
<tr>
<td>MINC</td>
<td>0.0552</td>
<td>0.0741</td>
<td>0.0461</td>
<td>0.0841</td>
<td>0.4222</td>
</tr>
<tr>
<td>MINL</td>
<td>0.1656</td>
<td>0.8207</td>
<td>0.2222</td>
<td>0.7935</td>
<td>0.6111</td>
</tr>
<tr>
<td>FAC1</td>
<td>0.0935</td>
<td>0.2047</td>
<td>0.1152</td>
<td>0.2398</td>
<td>0.3650</td>
</tr>
<tr>
<td>BSV</td>
<td>0.4685</td>
<td>0.9066</td>
<td>0.5381</td>
<td>0.5564</td>
<td>2.1926</td>
</tr>
</tbody>
</table>

60
Table 6: Average values for the threshold parameters

The quantities reported in this table are the average values for the threshold parameter, $\delta$ (for the 1- and 2-norm constrained portfolios) and the number of conjugate-gradient steps, $k$ (for the case of the partial minimum-variance portfolio). These quantities are reported for the two different calibration criteria: one, using cross-validation to minimize the portfolio variance (NC1V, NC2V, PARV); and two, maximizing the portfolio return in the last period (NC1R, NC2R, PARR). The time-series average is computed over the “rolling-windows”. Each column of the table reports the average for each of the datasets listed in Table 2.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>10Ind</th>
<th>48Ind</th>
<th>6FF</th>
<th>25FF</th>
<th>500CRSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC1V ($\delta$)</td>
<td>1.66215</td>
<td>1.68183</td>
<td>2.34914</td>
<td>2.39926</td>
<td>1.87370</td>
</tr>
<tr>
<td>NC1R ($\delta$)</td>
<td>1.55207</td>
<td>1.67957</td>
<td>1.79909</td>
<td>1.80242</td>
<td>1.46340</td>
</tr>
<tr>
<td>NC2V ($\delta$)</td>
<td>0.77320</td>
<td>0.57623</td>
<td>1.53733</td>
<td>1.23985</td>
<td>0.13167</td>
</tr>
<tr>
<td>NC2R ($\delta$)</td>
<td>0.60905</td>
<td>0.53564</td>
<td>1.10692</td>
<td>0.96842</td>
<td>0.07978</td>
</tr>
<tr>
<td>PARV ($k$)</td>
<td>3.49383</td>
<td>3.28244</td>
<td>3.01481</td>
<td>3.48148</td>
<td>3.35802</td>
</tr>
<tr>
<td>PARR ($k$)</td>
<td>1.83210</td>
<td>2.56997</td>
<td>1.65926</td>
<td>2.74321</td>
<td>1.69444</td>
</tr>
</tbody>
</table>
Figure 1: The Shortsale-Constrained Portfolio for 3 Assets

This figure depicts the shortsale-constrained minimum-variance portfolio for the case with three risky assets. The three axes in the reference frame give the portfolio weights for the three risky assets. Two triangles are depicted in the figure. The larger triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The smaller triangle (colored) represents the set of portfolios whose weights are nonnegative and sum up to one; that is, the set of shortsale-constrained portfolios. The ellipses centered around the minimum-variance portfolio, $w_{MINU}$, depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The shortsale-constrained minimum-variance portfolio is at the point where the colored triangle is tangent to the iso-variance curves. The figure also shows the location of the $1/N$ portfolio.
Figure 2: The Ledoit-Wolf Portfolio for 3 Assets

This figure depicts the Ledoit-Wolf portfolios for the case with three risky assets. The three axes in the reference frame give the portfolio weights for the three risky assets. Two triangles are depicted in the figure. The larger triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The smaller triangle (colored) represents the set of portfolios whose weights are nonnegative and sum up to one; that is, the set of shortsale-constrained portfolios. The ellipses centered around the minimum-variance portfolio, $w_{MINU}$, depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The Ledoit-Wolf portfolios form a curve that joins the $1/N$ portfolio, which is the Ledoit-Wolf portfolio obtained by setting $\nu = \infty$ in equation (7), with the minimum-variance portfolio equation $w_{MINU}$, which is the Ledoit-Wolf portfolio obtained by setting $\nu = 0$ in equation (7).
Figure 3: The 1-Norm-Constrained Portfolios for 3 Assets

This figure depicts the 1-norm-constrained portfolios for the case with three risky assets. The three axes in the reference frame give the portfolio weights for the three risky assets. Two triangles are depicted in the figure. The larger triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The smaller triangle (colored) represents the set of portfolios whose weights are nonnegative and sum up to one; that is, the set of shortsale-constrained portfolios. The ellipses centered around the minimum-variance portfolio, $w_{MINU}$, depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The hexagons represent the iso-1-norm curves; that is, the sets of portfolios with equal 1-norm. For each value of the threshold $\delta = \delta_1, \delta_2, \ldots, \delta_5$, the 1-norm-constrained minimum-variance portfolio is the point where the corresponding iso-1-norm curve is tangent to the iso-variance curve. The 1-norm-constrained portfolios corresponding to values of the threshold parameter $\delta$ ranging from 1 to $\|w_{MINU}\|_1$ describe a curve that joins the shortsale-constrained minimum-variance portfolio $w_{MINC}$ with the shortsale-unconstrained minimum-variance portfolio $w_{MINU}$. 
Figure 4: The 2-Norm-Constrained Portfolios for 3 Assets

This figure depicts the 2-norm-constrained portfolios for the case with three risky assets. The large triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The ellipses centered around the minimum-variance portfolio, $w_{MINU}$, depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The circumferences centered on the equally-weighted portfolio, $1/N$, are the iso-2-norm curves; that is, the set of portfolios with the same 2-norm. For a given threshold $\delta$, the 2-norm-constrained portfolio is the point where the corresponding iso-2-norm curve is tangent to an iso-variance curve. For values of $\delta$ ranging from 0 to $\|w_{MINU}\|_2^2$, the 2-norm-constrained portfolios describe a smooth curve that joins the $1/N$ portfolio (for $\delta = 0$) with the minimum-variance portfolio (for $\delta \geq \|w_{MINU}\|_2^2$).
Figure 5: The Partial Minimum-Variance Portfolios for 3 Assets

This figure depicts the partial minimum-variance portfolios and the 2-norm-constrained portfolios for the case with three risky assets. The three axes in the reference frame give the portfolio weights for the three risky assets. The large triangle depicts the intersection of the plane formed by all portfolios whose weights sum up to one with the reference frame. The ellipses centered around the minimum-variance portfolio, $w_{MINU}$, depict the iso-variance curves; that is, the curves formed by portfolios with equal variance. The circumferences centered on the equally-weighted portfolio, $1/N$, are the iso-2-norm curves; that is, the set of portfolios with the same 2-norm. For a given threshold $\delta$, the 2-norm-constrained portfolio is the point where the corresponding iso-2-norm ellipse is tangent to an iso-variance curve. For values of $\delta$ ranging from 0 to $\|w_{MINU}\|_2^2$, the 2-norm-constrained portfolios describe a smooth curve that joins the $1/N$ portfolio (for $\delta = 0$) with the minimum-variance portfolio (for $\delta \geq \|w_{MINU}\|_2^2$). The figure also depicts the first partial minimum-variance portfolio, $w_{PAR1}$, which is the combination of the $1/N$ portfolio and the first conjugate portfolio $w_{CG1}$ that minimizes the portfolio variance: $w_{PAR1} = e/N + \alpha_0 w_{CG1}$. Note that the first conjugate portfolio, $w_{CG1}$, is tangent to the curve of 2-norm-constrained portfolios at the equally-weighted, $1/N$ portfolio. The second conjugate portfolio, $w_{CG2}$, is not orthogonal to the first conjugate portfolio, but the product $w_{CG1} \hat{\Sigma} w_{CG2}$ is zero. Finally, because this is a case with only three risky assets, the second partial minimum-variance portfolio already coincides with the shortsale-unconstrained minimum-variance portfolio, $w_{MINU}$. 

![Diagram of portfolio weights and variances](image-url)
This figure shows the portfolio space for the case with two risky assets. The zero-cost subspace is the straight line passing through the origin and labelled as \( w_0 \), and the reduced space in this case is the set of real numbers \( \mathbb{R} \). The \( 1/N \) portfolio is the vector \( w = (1/2, 1/2)^\top \), the matrix \( Z \) has only one column equal to the vector \( (1/\sqrt{2}, -1/\sqrt{2})^\top \). The portfolio that assigns a weight of one to the first asset and zero to the second asset can then be written as \((0, 1)^\top = e/N + Z w = (1/2, 1/2)^\top - (1/\sqrt{2}, -1/\sqrt{2})^\top \times (1/\sqrt{2})\).