AP
7/13/06
3:30 PM

$1/N$

Victor DeMiguel
London Business School

Lorenzo Garlappi
University of Texas at Austin

Raman Uppal
London Business School

First draft: March 2005
This draft: June 2006*

*This paper was earlier circulated under the title, "How Inefficient is the $1/N$ Asset-Allocation Strategy?" We wish to thank John Campbell and Luis Viceira for their suggestions and for making available their data and computer code and Roberto Wessels for making available data on the ten sector portfolios of the S&P500 index. We also gratefully acknowledge comments from Pierluigi Balduzzi, John Birge, Michael Brennan, Ian Cooper, Bernard Dumas, Bruno Gerard, Francisco Gomes, Eric Jacquier, Chris Malloy, Francisco Nogales, Ľuboš Pástor, Anna Pavlova, Loriana Pelizzon, Nizar Touzi, Sheridan Titman, Rossen Valkanov, Yihong Xia, Tan Wang, Zhenyu Wang, and seminar participants at BI Norwegian School of Management, HEC Lausanne, HEC Montréal, London Business School, Stockholm School of Economics, University of Mannheim, University of Texas at Austin, University of Venice, University of Vienna, the Second McGill Conference on Global Asset Management, the 2004 International Symposium on Asset Allocation and Pension Management at Copenhagen Business School, the 2005 Conference on Developments in Quantitative Finance at the Isaac Newton Institute for Mathematical Sciences at Cambridge University, the 2005 Workshop on Optimization in Finance at University of Coimbra, the 2005 meeting of the Western Finance Association, the 2005 annual meeting of INFORMS, and the 2005 conference of The Institute for Quantitative Investment Research (Inquire UK).

Please address all correspondence to: Lorenzo Garlappi, McCombs School of Business, University of Texas at Austin, Austin TX, 78712; Email: `lorenzo.garlappi@mccombs.utexas.edu`.

$1/N$

Abstract

In this paper, we evaluate the out-of-sample performance of the portfolio policy from the sample-based mean-variance portfolio model and the various extensions of this model, designed to reduce the impact of estimation error relative to the benchmark strategy of investing a fraction $1/N$ of wealth in each of the N assets available. Of the fourteen models of optimal portfolio choice that we evaluate across seven empirical datasets, we find that none is consistently better than the $1/N$ rule in terms of Sharpe ratio, certainty-equivalent return, or turnover. This finding indicates that, out of sample, the gain from optimal diversification is more than offset by estimation error. To gauge the severity of estimation error, we derive analytically the length of the estimation window needed for the sample-based mean-variance strategy to outperform the $1/N$ benchmark; for parameters calibrated to U.S. stock market data, we find that, for a portfolio with only 25 assets, the estimation window needed is more than 3,000 months, and for a portfolio with 50 assets, it is more than 6,000 months, although in practice these parameters are estimated using 120 months of data. Using simulated data, we further document that even the various extensions to the sample-based mean-variance model designed to deal with estimation error reduce only moderately the estimation window needed to outperform the naive $1/N$ benchmark. This suggests that there are still many "miles to go" before the gains promised by optimal portfolio choice can actually be realized out of sample.

Keywords: Portfolio choice, asset allocation, investment management.
JEL Classification: G11.

Contents

1 Introduction **1**

2 Description of the asset allocation models considered **4**
2.1 Naive portfolio ("ew" or "$1/N$") 5
2.2 Sample-based mean-variance portfolio ("mv") 5
2.3 Bayesian approach to estimation error ("bs," "dm") 6
2.4 Portfolios with moment restrictions ("min," "vw," "mp") 7
2.5 Portfolios with shortsale constraints ("mv-c", "bs-c", "min-c","g-min-c") 8
2.6 Optimal combination of portfolios ("mv-min," "ew-min") 9

3 Methodology for evaluating performance **10**

4 Results from the seven empirical datasets considered **12**
4.1 Sharpe ratios 13
4.2 Certainty equivalent (CEQ) returns 16
4.3 Portfolio turnover 17
4.4 Summary of findings from the empirical datasets 17

5 Results from studying analytically the estimation error **18**

6 Results from simulated data **22**
6.1 Details about how the simulated data is generated 22
6.2 Discussion of results from simulated data 23

7 Results for other specifications: Robustness checks **26**
7.1 Different lengths of the estimation window 27
7.2 Increasing estimation window, rather than rolling window 27
7.3 Different holding periods 27
7.4 Portfolios that include the risk-free asset 28
7.5 Portfolios that exclude the factors as investable assets 28
7.6 Benchmark is $1/N$-buy-and-hold rather than $1/N$-with rebalancing 29
7.7 Different levels of risk aversion 29
7.8 Different levels of confidence in the asset-pricing model 29
7.9 Portfolios that use the cross-sectional characteristics of stocks 30

8 Conclusions **31**

A Details for some of the models of optimal portfolio choice **33**
A.1 Bayes-Stein shrinkage portfolio 33
A.2 Bayesian "Data-and-Model" portfolio 33
A.3 Portfolio implied by asset pricing models with unobservable factors 35
A.4 Optimal "Three-Fund" portfolio 35
A.5 Combination of minimum-variance strategy and the naive $1/N$ strategy 36

B Description of the seven empirical datasets **37**
B.1 Sector portfolios ("S&PSectors") 37
B.2 Industry portfolios ("Industry") 37
B.3 International equity indexes ("International") 37
B.4 MKT, SMB and HML portfolios ("MKT/SMB/HML") 37
B.5 Size- and Book-to-Market-sorted portfolio ("FF-1-factor," "FF-3-factor," "FF-4-factor") 38

C Proof for Proposition 1 **38**

Tables **40**

References **49**

1 Introduction

In about the fourth century, Rabbi Issac bar Aha proposed the following rule for asset allocation:[1] "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand." After a "brief" lull in the literature on asset allocation, there have been considerable advances starting with the pathbreaking work of Markowitz (1952), who derived the *optimal* rule for allocating wealth across risky assets in a static setting when investors care only about the mean and variance of a portfolio's return.[2] Because the implementation of these portfolios with moments estimated via their sample analogues is notorious for producing extreme weights, that fluctuate substantially over time, and perform poorly out of sample, considerable effort has been devoted to the issue of handling estimation error with the goal of improving the performance of the Markowitz (1952) model.[3]

A prominent role in this vast literature is played by the *Bayesian approach* to estimation error, with its multiple implementations ranging from the purely statistical approach relying on diffuse-priors (Barry (1974) and Bawa, Brown, and Klein (1979)), to "shrinkage estimators" (Jobson and Korkie (1980) and Jorion (1985, 1986)), to the more recent approaches that rely on an asset pricing model for establishing a prior (Pástor (2000) and Pástor and Stambaugh (2000)). Equally rich is the set of *non-Bayesian* approaches to estimation error, which include "robust" portfolio allocation rules (see, for instance, Goldfarb and Iyengar (2003) and Garlappi, Uppal, and Wang (2006)); portfolio rules designed to optimally diversify across market *and* estimation risk (Kan and Zhou (2005)); portfolios that exploit the moment restrictions imposed by the factor structure of returns (MacKinlay and Pastor (2000)); portfolios that ignore expected returns altogether and focus only on the error in estimating the covariance matrix (Best and Grauer (1992), Ledoit (1996), Chan, Karceski, and Lakonishok (1999), and Ledoit and Wolf (2003)); and, finally, portfolio rules that impose shortselling constraints (Frost and Savarino (1988), Chopra (1993), and Jagannathan and Ma (2003)).[4]

Our objective in this paper is to understand the conditions under which optimal portfolio models can be expected to perform well even in the presence of estimation risk. To do this, we evaluate the *out-of-sample* performance of the sample-based mean-variance portfolio rule and the various extensions of this model, to reduce the effect of estimation error relative to the performance of the *naive* asset-allocation rule. We define the naive asset-allocation rule to be one in which a fraction $1/N$ of wealth is allocated to each of the N assets available for investment

[1]Babylonian Talmud: Tractate Baba Mezi'a, folio 42a.

[2]Samuelson (1969) and Merton (1969) show that the static portfolio rule is optimal even in a multiperiod setting if the investment opportunity set is constant. Merton (1971) identifies the optimal rule when the investment opportunity set is stochastic. In order to limit the length of this paper, we no longer discuss the performance of dynamic portfolio strategies; for that, readers are referred to the 2005 version of the paper.

[3]For a discussion of the problems in implementing mean-variance optimal portfolios, see Hodges and Brealey (1978), Michaud (1989), Best and Grauer (1991), and Litterman (2003). For a general survey of the literature on portfolio selection, see Campbell and Viceira (2002) and Brandt (2004).

[4]Michaud (1998) has advocated the use of resampling methods. Scherer (2002) and Harvey, Liechty, Liechty, and Müller (2003) discuss the various limitations of this approach.

at each rebalancing date.[5] There are two reasons for using the naive asset-allocation rule as a benchmark. First, it does not rely either on estimation of the moments of asset returns or on optimization, and so is easy to implement. Second, despite the sophisticated theoretical models developed in the last fifty years and the advances in methods for estimating the parameters of these models, investors continue to use such simple allocation rules for allocating their wealth across assets.[6] We wish to emphasize, however, that the purpose of this study is *not* to advocate the use of the $1/N$ heuristic as an asset allocation strategy, but merely to use it as a benchmark to assess the success or failure of various portfolio rules proposed in the literature.

We compare the out-of-sample performance of fourteen different portfolio models relative to that of the $1/N$ policy across seven empirical datasets of monthly returns,[7] using the following three performance criteria: (i) the out-of-sample Sharpe ratio; (ii) the certainty-equivalent return (CEQ) for the expected utility of a mean-variance investor; and (iii) the turnover (trading volume) for each portfolio strategy. The fourteen models whose performance we study are listed in Table 1 and discussed in Section 2. The seven empirical datasets across which we evaluate the performance of the different portfolio models are listed in Table 2 and described in Appendix B.

Our first contribution is to show that, of the fourteen models of optimal portfolio choice that we evaluate across seven empirical datasets, none is consistently better than the naive $1/N$ benchmark in terms of Sharpe ratio, certainty-equivalent return, or turnover. Although this was shown in the literature with regard to some of the older models,[8] we demonstrate that this is true (a) for a wide range of models that include several developed more recently, (b) using three different performance metrics, and (c) across several datasets. In particular, we find that the Bayesian models of Jorion (1985, 1986), Pástor (2000), and Pástor and Stambaugh (2000), the "three-fund" model in Kan and Zhou (2005), and the "robust" portfolio in Garlappi, Uppal, and Wang (2006), do not outperform the naive $1/N$ benchmark. The approach recommended in MacKinlay and Pastor (2000), while successful in lowering turnover, is not very successful in improving the out-of-sample Sharpe ratio. In general, we find that the *unconstrained* policies that try to incorporate estimation error perform much worse than any of the strategies that constrain shortsales, and also perform much worse than the $1/N$ strategy. *Imposing constraints*

[5] We also consider the "buy-and-hold" case in which the investor allocates $1/N$ at the initial date and then holds this portfolio until the terminal date. The results for this case are similar to those for the case with rebalancing; because the results of the buy-and-hold strategy are sensitive to the starting date, we do not report these results any longer.

[6] For instance, Benartzi and Thaler (2001) document that investors allocate their wealth across assets using the naive $1/N$ rule. Huberman and Jiang (2006) find that participants tend to invest in only a small number of the funds offered to them, and that they tend to allocate their contributions evenly across the funds that they use, with this tendency weakening with the number of funds used.

[7] The 2005 version of our paper considered two additional datasets, both with quarterly returns, for which the results are similar to the ones reported here. The first is data from January 1954 to September 1996 used in Campbell and Viceira (2001), which includes returns on the U.S. equity market portfolio, a 3-month bond, and a 10-year bond. The second is data from September 1952 to September 1999 used in Campbell and Viceira (1999) and Campbell, Chan, and Viceira (2003), which includes returns on the U.S. equity market portfolio, the 90-day Treasury bill, and the 5-year bond.

[8] Bloomfield, Leftwich, and Long (1977) show that sample-based mean-variance optimal portfolios do not outperform an equally-weighted portfolio, and Jorion (1991) finds that the equally-weighted and value-weighted indices have an out-of-sample performance similar to that of the minimum-variance portfolio and the tangency portfolio obtained with Bayesian shrinkage methods.

on the sample-based mean-variance and Bayesian portfolio strategies leads to only a modest improvement in Sharpe ratios and CEQ returns, though a substantial reduction in turnover. Of all the models we study, the minimum-variance portfolio with constraints studied in Jagannathan and Ma (2003) performs best in terms of Sharpe ratio. But even this model cannot deliver a Sharpe ratio or CEQ return that is statistically superior to that delivered by the $1/N$ strategy in any of the seven empirical datasets, and its turnover is typically higher than that of the $1/N$ policy.

To understand better the reasons for the poor performance of the optimal portfolio strategies relative to the $1/N$ benchmark, our second contribution is to derive an *analytic* expression for the *critical length* of the estimation window that is needed for the sample-based mean-variance strategy in order to achieve a higher CEQ return than that for the $1/N$ strategy. This critical estimation-window length is a function of the number of assets, the ex-ante Sharpe ratio of the mean-variance portfolio, and the Sharpe ratio of the $1/N$ policy. Based on parameters calibrated to U.S. stock-market data, we find that the critical length of the estimation window is 3,000 months for a portfolio with only 25 assets, and more than 6,000 months for a portfolio with 50 assets. The severity of estimation error is startling if we consider that, in practice, these portfolio models are typically estimated using only 60 or 120 months of data. Our analytical results suggest that the error in estimating expected returns contributes much more to the poor performance of the sample-based mean-variance strategy than the error in estimating covariances.

Because the above analytic results are available only for the sample-based mean-variance strategy, we use simulated data to examine the various extensions to the sample-based mean-variance model that have been developed explicitly to deal with estimation error. Our third contribution is to show that these models, too, need very long estimation windows before they can be expected to outperform the $1/N$ policy; that is, the estimation error is so severe that these models have only modest success in reducing the effects of estimation error. From our simulation results we conclude that portfolio strategies from the optimizing models are expected to outperform the $1/N$ benchmark if: (i) the estimation window is long, (ii) the ex-ante (true) Sharpe ratio of the mean-variance efficient portfolio is substantially higher than that of the $1/N$ portfolio, and (iii) the number of assets is small. The first two conditions are intuitive. The reason for the last condition is that a smaller number of assets implies fewer parameters to be estimated, and therefore, less room for estimation error; and, all else being equal, a smaller number of assets makes naive diversification less effective relative to optimal diversification.

The intuition for the poor performance of optimizing models relative to $1/N$ is that even small errors in estimating the moments of asset returns can lead to large differences in the portfolio weights. As a result, "allocation mistakes" caused by using the $1/N$ weights can turn out to be *smaller* than mistakes caused by using the weights from an optimizing model with inputs that have been estimated with error. Consider the following extreme two-asset example. Suppose that the true per annum mean and volatility for both assets are the same, 8% and 20% respectively, and that the correlation is 0.99. In this case, because the two assets are identical, the optimal mean-variance weights in the two assets would be 50%. If, on the other hand, the mean on the first asset was not known and is estimated to be 9% instead of 8%, then the mean-variance model would recommend a weight of 635% in the first asset and −535% in the second. That is, the *optimization* tries to exploit even the smallest difference in the two assets by taking extreme long and short positions *without* taking into account the fact that these differences in returns may be the result of estimation error. As we describe in Section 4, the weights from mean-variance optimization when using actual data are much more extreme than the weights in the example above. Although the "error-maximizing" property of the mean-variance portfolio has been described in the literature (see Michaud (1989) and Best and Grauer (1991)), our contribution is to show that, because the effect of estimation error on the weights is so large, even the models in the literature that are designed to reduce the effect of estimation error achieve only modest success.

We draw two conclusions from the above findings. First, our study suggests that, although there has been considerable progress in the design of optimal portfolios, more effort needs to be devoted to improving the estimation of the moments, and especially the expected returns, of investable assets. For this, methods that complement traditional classical and Bayesian statistical techniques by exploiting empirical regularities that are present for a particular set of assets (see, for example, Brandt, Santa-Clara, and Valkanov (2005)) can represent a promising direction to pursue. Second, given the inherent simplicity and the relatively low cost of implementing the $1/N$ naive diversification rule, such a strategy should serve as a natural benchmark to assess the performance of more sophisticated asset-allocation rules. This is an important hurdle, both for academic research proposing new models of portfolio selection and for "active" portfolio-management strategies offered by the investment industry.

The rest of the paper is organized as follows. In Section 2, we describe the various models of optimal asset allocation whose performance we evaluate. In Section 3, we explain our methodology for comparing the performance of these models to that of $1/N$; the results of this comparison for seven empirical datasets are given in Section 4. Section 5 contains the analytic results on the critical length of the estimation window needed for the sample-based mean-variance policy to outperform the $1/N$ benchmark, and in Section 6 we present a similar analysis for other models of portfolio choice using simulated data. The various experiments that we undertake to verify the robustness of our findings are described in Section 7. We conclude in Section 8. Details of how to implement some of the portfolio-selection models we study are presented in Appendix A, the empirical datasets we use are described in Appendix B, and the proof for the main analytic result is provided in Appendix C. The tables for the experiments undertaken to verify the robustness of our results are collected in a separate appendix titled "Tables with Results for Robustness Checks", which can be downloaded from our website.

2 Description of the asset allocation models considered

In this section, we discuss the various models from the portfolio-choice literature that we consider in our study. Because these models are familiar to most readers, we provide only a brief description of each, and instead focus on explaining how the different models are related to each other. The list of models we analyze is summarized in Table 1, and the details on how to implement these models are given in Appendix A.

We start by defining some notation. We use R_t to denote the N-vector of *excess* returns (over the risk free asset) on the N risky assets available for investment at date t. The N-dimensional vector μ_t is used to denote the *expected* returns on the risky asset in excess of the risk-free rate, and Σ_t to denote the corresponding $N \times N$ variance-covariance matrix of returns, with their sample counterparts given by $\hat{\mu}_t$ and $\hat{\Sigma}_t$, respectively. Let M denote the length over which these moments are estimated, and T the total length of the data series. We use $\mathbf{1}_N$ to define an N-dimensional vector of ones, and I_N to indicate the $N \times N$ identify matrix. Finally, x_t is the vector of portfolio weights invested in the N risky assets, with $1-\mathbf{1}_N^\top \mathrm{x}_t$ invested in the risk-free asset, and the vector of *relative* weights in the portfolio with only-risky assets given by

$$\mathrm{w}_t = \frac{\mathrm{x}_t}{\mathbf{1}_N^\top \mathrm{x}_t}. \tag{1}$$

To facilitate the comparison across different strategies, we consider an investor whose preferences are fully described by the mean and variance of a chosen portfolio, x_t. At each time t, the decision-maker selects x_t, to maximize expected utility:[9]

$$\max_{\mathrm{x}_t} \mathrm{x}_t^\top \mu_t - \frac{\gamma}{2} \mathrm{x}_t^\top \Sigma_t \mathrm{x}_t, \tag{2}$$

in which γ can be interpreted as the investor's risk aversion. The solution of the above optimization is $\mathrm{x}_t = (1/\gamma)\Sigma_t^{-1}\mu$. The vector of *relative* portfolio weights invested in the N risky assets at time t is

$$\mathrm{w}_t = \frac{\Sigma_t^{-1}\mu_t}{\mathbf{1}_N \Sigma_t^{-1}\mu_t}. \tag{3}$$

Almost all the models we consider deliver portfolio weights that can be expressed as in Equation (3), with the main difference being in how one obtains μ_t and Σ_t.

2.1 Naive portfolio ("ew" or "1/N")

The naive strategy that we consider involves holding a portfolio weight $\mathrm{w}_t^{\mathrm{ew}} = 1/N$ in each of the N risky assets. This strategy does not involve any optimization or estimation and completely ignores the data. For comparison with the weights in (3), one can also think of the $1/N$ portfolio as a strategy that does estimate the moments μ_t and Σ_t, but imposes the restriction that $\mu_t \propto \Sigma_t \mathbf{1}_N$ for all t, which implies that expected returns are proportional to total risk rather than systematic risk.

2.2 Sample-based mean-variance portfolio ("mv")

In the mean-variance model of Markowitz (1952), the investor optimizes the trade-off between the mean and variance of portfolio returns. To implement this model, we follow the classic

[9] The constraint that the weights sum to 1 is incorporated implicitly by expressing the optimization problem in terms of returns in excess of the risk-free rate.

"plug-in" approach; that is, we solve problem (2), with the mean and covariance matrix of asset returns replaced by their sample counterparts, $\hat{\mu}$ and $\hat{\Sigma}$, respectively. We shall refer to this strategy as the "sample-based mean-variance portfolio." Note that this portfolio strategy completely ignores the presence of estimation error.

2.3 Bayesian approach to estimation error ("bs," "dm")

Under the Bayesian approach, the estimates of μ and Σ are computed using the *predictive distribution* of asset returns. This distribution is obtained by integrating the *conditional likelihood*, $f(R|\mu, \Sigma)$, over μ and Σ with respect to a certain *subjective prior*, $p(\mu, \Sigma)$. In the literature, the Bayesian approach to estimation error has been implemented in different ways. Below, we describe three common implementations that we consider in our study.

2.3.1 Bayesian diffuse-prior portfolio

Barry (1974), Brown (1979), and Klein and Bawa (1976) show that if the prior is chosen to be diffuse, that is, $p(\mu, \Sigma) \propto |\Sigma|^{-(N+1)/2}$, and the conditional likelihood is Normal, then the predictive distribution is a Student-t with mean $\hat{\mu}$ and variance $\hat{\Sigma}\,(1 + 1/M)$. Hence, while still using the historical mean to estimate expected returns, this approach inflates the covariance matrix by a factor of $(1 + 1/M)$. For a sufficiently long estimation window M (as in our study, where $M = 120$ months), the effect of this correction is negligible, and the performance of the Bayesian diffuse-prior portfolio is virtually indistinguishable from that of the sample-based mean-variance portfolio. For this reason we do not report the results for this Bayesian strategy.

2.3.2 Bayes-Stein shrinkage portfolio ("bs")

The Bayes-Stein shrinkage portfolio is an application of the idea of shrinkage estimation pioneered by Stein (1955) and James and Stein (1961), and is designed to handle the error in estimating expected returns by using estimators of the form $\mu^{\mathrm{S}} = (1 - \phi)\hat{\mu} + \phi\overline{\mu}$, where $0 < \phi < 1$. These estimators "shrink" the sample mean toward a common "grand mean," $\overline{\mu}$. In our analysis, we use the estimator proposed by Jorion (1985, 1986), who takes the grand mean, $\overline{\mu}$, to be the mean of the minimum-variance portfolio, $\mu^{\min}$. In addition to shrinking the estimate of the mean, Jorion also accounts for estimation error in the covariance matrix via traditional Bayesian-estimation methods.[10]

2.3.3 Bayesian portfolio based on belief in an asset pricing model ("dm")

Under the Bayesian "Data-and-Model" (dm) approach developed in Pástor (2000) and Pástor and Stambaugh (2000), the shrinkage target depends on the investor's prior belief in a particular asset-pricing model, and the degree of shrinkage is determined by the variability of the prior

[10]See also Jobson and Korkie (1980), Frost and Savarino (1986), and Dumas and Jacquillat (1990) for other applications of shrinkage estimation in the context of portfolio selection.

belief relative to the information contained in the data. These portfolios are a further refinement of shrinkage portfolios because they address the arbitrariness of the choice of a shrinkage target, $\overline{\mu}$, and of the shrinkage factor, ϕ, by using the investor's belief about the validity of an asset pricing model. In our empirical analysis, we consider the case in which the investor believes either in the Capital Asset Pricing Model or the Arbitrage Pricing Theory (with either 3 or 4 factors) with a subjective probability of $\omega = 0.50$.[11]

2.4 Portfolios with moment restrictions ("min," "vw," "mp")

In this subsection, we consider portfolio strategies that impose restrictions on the estimation of the moments of asset returns.

2.4.1 Minimum-variance portfolio ("min")

Under this strategy, we choose the portfolio of risky assets that minimizes the variance of returns; that is,

$$\min_{\mathrm{w}_t} \ \mathrm{w}_t^\top \Sigma_t \mathrm{w}_t, \quad \text{s.t.} \quad \mathbf{1}_N^\top \mathrm{w}_t = 1. \tag{4}$$

To implement this policy we use only the estimate of the covariance matrix of asset returns (in our case, we use the sample covariance matrix) and completely ignore estimates of the expected returns.[12] Also, although this strategy does not fall into the general structure of mean-variance expected utility, its weights can be thought of as a limiting case of (3), if a mean-variance investor either ignores expected returns or, equivalently, restricts expected returns so that they are identical across all assets, that is, $\mu_t \propto \mathbf{1}_N$.

2.4.2 Value-weighted portfolio implied by the Market Model ("vw")

The optimal strategy in a CAPM world is the value-weighted market portfolio. So, for each of the datasets we identify a benchmark "market" portfolio and report the Sharpe ratio and CEQ for holding this portfolio. The turnover of this strategy is zero.

2.4.3 Portfolio implied by asset pricing models with unobservable factors ("mp")

MacKinlay and Pastor (2000) show that, if returns are explained by an exact factor structure but some factors are not observed, then the resulting mispricing is contained in the covariance matrix of the residuals. They use this insight to construct an estimator of expected returns that is more stable and reliable than estimators obtained using traditional methods. MacKinlay and

[11] In the separate document titled, "Tables with Results for Robustness Checks," we consider the cases of ω equal to 0.25, 0.75 and 1.00. The results for these values of ω are similar to those in the manuscript for $\omega = 0.50$.

[12] Note that expected returns *do* appear in the likelihood function needed to estimate Σ_t. However, under the assumption of Normal asset return, it is possible to show (see Morrison (1990)) that for any estimator of the covariance matrix, the MLE estimator of the mean is always the sample mean. This allows one to remove the dependence on expected returns μ_t for constructing the MLE estimator of Σ_t.

Pastor show that, in this case, the covariance matrix of returns takes the following form:[13]

$$\Sigma = \nu\mu\mu^\top + \sigma^2 I, \tag{5}$$

in which ν and σ^2 are positive scalars. They use the maximum-likelihood estimates of ν, σ^2, and μ to derive the corresponding estimates of the mean and covariance matrix of asset returns. The optimal portfolio weights are obtained by substituting these estimates into Equation (2).

2.5 Portfolios with shortsale constraints ("mv-c", "bs-c", "min-c","g-min-c")

We also consider a number of strategies that constrain shortselling. The sample-based mean-variance-constrained, "mv-c," Bayes-Stein-constrained, "bs-c," and minimum-variance-constrained, "min-c," policies are obtained by imposing an additional non-negativity constraint on the portfolio weights in the corresponding optimization problems.

To interpret the effect of shortsale constraints, observe that imposing the constraint $\mathrm{x}_i \geq 0$, $i = 1, \ldots, N$ in the basic mean-variance optimization (2) yields the following Lagrangian,

$$\mathcal{L} = \mathrm{x}_t^\top \mu_t - \frac{\gamma}{2}\mathrm{x}_t^\top \Sigma_t \mathrm{x}_t + \mathrm{x}_t^\top \lambda_t, \tag{6}$$

in which λ_t is the $N \times 1$ vector of Lagrange multipliers for the shortselling constraints. Rearranging the above expression, we can see that the constrained mean-variance portfolio weights are equivalent to the unconstrained weights but with the adjusted mean vector: $\tilde{\mu}_t = \mu_t + \lambda_t$. To see why this is a form of shrinkage on the expected returns, note that the shortselling constraint on asset i is likely to be binding when expected returns are negative. When the constraint for asset i binds, $\lambda_{t,i} > 0$ and the expected return is increased from $\mu_{t,i}$ to $\tilde{\mu}_{t,i} = \mu_{t,i} + \lambda_{t,i}$. Hence, imposing a shortsale constraint on the sample-based mean-variance problem is equivalent to "shrinking" the expected return toward zero.

Similarly, Jagannathan and Ma (2003) show that imposing a shortsale constraint on the minimum-variance portfolio is similar in effect to shrinking the elements of the variance-covariance matrix. Jagannathan and Ma (2003, p. 1654) find that, with a constraint on shortsales, "the sample covariance matrix performs almost as well as those constructed using factor models, shrinkage estimators or daily returns." Because of this finding, we do not evaluate the performance of other models—such as Best and Grauer (1992), Ledoit (1996), Chan, Karceski, and Lakonishok (1999), and Ledoit and Wolf (2003)—that have been developed to deal explicitly with the problems associated with estimating the covariance matrix.[14]

We also consider a new strategy that has not been considered in the existing literature. Motivated by the desire to examine whether the out-of-sample performance can be improved by ignoring expected returns (which are difficult to estimate) but still taking into account the correlations between returns, we consider a strategy that is a combination of the $1/N$ policy

[13] MacKinlay and Pastor (2000) express the restriction in terms of the covariance matrix of *residuals* instead of returns. However, this does not affect the determination of the optimal portfolios.

[14] See Sections III.B and III.C of Jagannathan and Ma (2003) for an extensive discussion of the performance of other models used for estimating the sample covariance matrix.

and the constrained-minimum-variance strategy. This strategy, denoted by "g-min-c," can be interpreted as a simple generalization of the shortsale-constrained minimum-variance portfolio. It is obtained by imposing the additional constraint on the minimum-variance problem (4): $\mathrm{w} \geq a\mathbf{1}_N$, with $a \in [0, 1/N]$. Observe that the shortsale-constrained minimum-variance portfolio corresponds to the case in which $a = 0$, while setting $a = 1/N$ yields the $1/N$ portfolio. In the empirical section, we study the case in which $a = \frac{1}{2}\frac{1}{N}$, arbitrarily chosen as the middle ground between the constrained-minimum-variance portfolio and the $1/N$ portfolio.

2.6 Optimal combination of portfolios ("mv-min," "ew-min")

We also consider portfolios that are themselves combinations of other portfolios, such as the mean-variance portfolio, the minimum-variance portfolio, and the equally-weighted portfolio. The mixture portfolios are constructed by applying the idea of shrinkage *directly* to the portfolio weights. That is, instead of first estimating the moments and then constructing portfolios with these moments, one can directly construct (non-normalized) portfolios of the form:

$$\mathrm{x}^{\mathrm{S}} = c\,\mathrm{x}^{c} + d\,\mathrm{x}^{d}, \quad \text{s.t.} \quad \mathbf{1}_N^{\top}\mathrm{x}^{\mathrm{S}} = 1, \tag{7}$$

in which x^c and x^d are two reference portfolios chosen by the investor. Working directly with portfolio weights is intuitively appealing because it makes it easier to select a specific target toward which one is shrinking a given portfolio.[15] The three mixture portfolios that we consider are described below.

2.6.1 The Kan and Zhou (2005) three-fund portfolio ("mv-min")

In order to improve on the models that use Bayes-Stein shrinkage estimators, Kan and Zhou (2005) propose a "Three-Fund" portfolio rule, in which the role of the third fund is to minimize "estimation risk." The intuition underlying their model is that because estimation risk cannot be diversified away by holding only a combination of the tangency portfolio and the risk-free asset, an investor will benefit from also holding some other risky-asset portfolio, that is, a third fund. Kan and Zhou search for this optimal three-fund portfolio rule in the class of portfolios that can be expressed as a combination of the sample-based mean-variance portfolio and the minimum-variance portfolio. The non-normalized weights of this mixture portfolio are

$$\hat{\mathrm{x}}_t^{\text{mv-min}} = \frac{1}{\gamma}(c\,\hat{\Sigma}_t^{-1}\hat{\mu}_t + d\,\hat{\Sigma}_t^{-1}\mathbf{1}_N), \tag{8}$$

in which c and d are chosen optimally to maximize the expected utility of a mean-variance investor.[16] The weights in the risky assets used in our implementation are given by normalizing (8), that is, $\hat{\mathrm{w}}_t^{\text{mv-min}} = \mathrm{x}_t^{\text{mv-min}}/\mathbf{1}_N^{\top}\mathrm{x}_t^{\text{mv-min}}$.

[15] See Brandt (2004) for a comprehensive analysis of the benefits of directly modeling portfolio weights in solving portfolio choice problems.

[16] The expression for the optimal portfolio weight, with the optimal expressions for c and d substituted in, is reported in equation (A20) of Appendix A.4.

2.6.2 Robust portfolio for uncertainty-averse investors with multiple priors

Next, we consider the portfolio strategy proposed in Garlappi, Uppal, and Wang (2006) for an investor who is not just risk averse but also uncertainty averse, in the sense of Knight (1921). This portfolio incorporates explicitly both the investor's uncertainty about the true parameters and the investor's desire to choose a portfolio that is robust to estimation error. This portfolio is constructed via a max-min optimization, in which the investor selects the "worst-case" portfolio based on the set of possible priors for the unknown parameters. Garlappi, Uppal, and Wang show that if returns on the N assets are estimated jointly, then the "robust" portfolio is equivalent to a weighted average of the mean-variance portfolio and the minimum-variance portfolio, where the weights depend on the amount of parameter uncertainty and the investor's degree of aversion to uncertainty. In the limiting case of infinite uncertainty aversion, the optimal portfolio converges to the minimum-variance portfolio. By construction, therefore, the performance of such a portfolio lies between the performances of the sample-based mean-variance portfolio and the minimum-variance portfolio. Because we report the performance for these two portfolios, (both of which perform quite poorly, we do not report separately the performance of robust portfolio strategies.

2.6.3 Mixture of equally-weighted and minimum-variance portfolios ("ew-min")

Finally, we consider a new portfolio strategy that has not been studied in the existing literature. This strategy is a combination of the naive $1/N$ portfolio and the minimum-variance portfolios ("ew-min"), rather than the mean-variance portfolio and the minimum-variance portfolio considered in Kan and Zhou (2005) and Garlappi, Uppal, and Wang (2006). Again, our motivation for considering this portfolio is that, because expected returns are more difficult to estimate than covariances, one may want to ignore estimates of mean returns and not the estimates of covariances. And so, one may wish to combine the $1/N$ portfolio with the minimum-variance portfolio. Specifically, the portfolio we consider is

$$\hat{\mathrm{w}}^{\text{ew-min}} = c\,\frac{1}{N}\mathbf{1}_N + d\,\hat{\Sigma}^{-1}\mathbf{1}_N, \quad \text{s.t.} \;\; \mathbf{1}_N^\top \hat{\mathrm{w}}^{\text{ew-min}} = 1, \tag{9}$$

in which c and d are chosen to maximize the expected utility of a mean-variance investor.[17]

3 Methodology for evaluating performance

Our goal is to study the performance of each of the above models across a variety of datasets that have been considered in the literature on asset allocation. The full list of datasets considered is summarized in Table 2 and described in Appendix B.

Our analysis relies on a "rolling-sample" approach. Specifically, given a T-month long dataset of asset returns, we choose an estimation window of length $M = 60$ or $M = 120$

[17]The expressions for the optimal c and d are given in Appendix A.5.

months.[18] In each month t, starting from $t = M$, we use the data in the previous M months to estimate the parameters needed to implement a particular strategy. These estimated parameters are then used to determine the relative portfolio weights in the portfolio of only-risky assets. We then use these weights to compute the return in month $t + 1$.[19] This process is continued by adding the return for the next period in the dataset and dropping the earliest return, until the end of the dataset is reached.[20] The outcome of this rolling-window approach is a series of $T - M$ monthly *out-of-sample* returns generated by each of the portfolio strategies listed in Table 1, for each of the empirical datasets in Table 2.

Given the time series of monthly out-of-sample returns generated by each strategy and in each dataset, we compute the following quantities. One, we measure the *out-of-sample Sharpe ratio* of strategy k, defined as the sample mean of out-of-sample excess returns (over the risk-free asset), $\hat{\mu}_k$, divided by their sample standard deviation, $\hat{\sigma}_k$:

$$\widehat{\mathrm{SR}}_k = \frac{\hat{\mu}_k}{\hat{\sigma}_k}. \tag{10}$$

To test whether the Sharpe ratios of two strategies are statistically distinguishable, we also compute the P-value of the difference, using the approach suggested by Jobson and Korkie (1981) after making the correction pointed out in Memmel (2003).[21]

In order to assess the effect of estimation error on performance, we also compute the *in-sample Sharpe ratio* for each strategy. This is computed by using the *entire* time-series of excess returns, that is, with the estimation window $M = T$. Formally, the in-sample Sharpe ratio of strategy k is

$$\widehat{\mathrm{SR}}_k^{\mathrm{IS}} = \frac{\mathrm{Mean}_k}{\mathrm{Std}_k} = \frac{\hat{\mu}_k^{\mathrm{IS}\top}\,\hat{\mathrm{w}}_k}{\sqrt{\hat{\mathrm{w}}_k^\top\,\hat{\Sigma}_k^{\mathrm{IS}}\,\hat{\mathrm{w}}_k}}, \tag{11}$$

[18]The results for the case of $M = 60$ are not different from the case of $M = 120$ and hence, in the interest of conserving space, are not reported but are discussed in Section 7.1.

[19]As a robustness check, we also consider the "total" portfolios of both risky and risk-free assets. The qualitative results for this portfolio are similar to those for the portfolio with only risky assets, as discussed in Section 7.4. Because the performance of a portfolio that includes the risk-free asset depends both on the asset-allocation decision and on the timing decision, we decided to report the case of portfolios with only risky assets in order to focus on the effect of asset allocation on portfolio performance.

[20]We also consider the case in which we do *not* drop the earlier observations, and so the estimation window is increasing over time. The empirical results in this case are similar to those obtained with a rolling window, because, as we show in Sections 5 and 6, small changes in the length of the estimation window have almost no effect on the results. Because it is difficult to study analytically the case with an increasing estimation window, as we wish to do in Section 5, we report results only for the case of a rolling window. The results for the case with an estimation window that is increasing over time are discussed in Section 7.2.

[21]Specifically, given two portfolios i and n, with $\hat{\mu}_i$, $\hat{\mu}_n$, $\hat{\sigma}_i$, $\hat{\sigma}_n$, $\hat{\sigma}_{i,n}$ as their estimated means, variances, and covariances over a sample of size $T - M$, the test of the hypothesis $\mathrm{H}_0 : \hat{\mu}_i/\hat{\sigma}_i - \hat{\mu}_n/\hat{\sigma}_n = 0$ is obtained via the test statistic $\hat{z}_{\mathrm{JK}}$, which is asymptotically distributed as a standard normal:

$$\hat{z}_{\mathrm{JK}} = \frac{\hat{\sigma}_n\hat{\mu}_i - \hat{\sigma}_i\hat{\mu}_n}{\sqrt{\hat{\vartheta}}},$$

in which

$$\hat{\vartheta} = \frac{1}{T - M}\left(2\hat{\sigma}_i^2\hat{\sigma}_n^2 - 2\hat{\sigma}_i\hat{\sigma}_n\hat{\sigma}_{i,n} + \frac{1}{2}\hat{\mu}_i^2\hat{\sigma}_n^2 + \frac{1}{2}\hat{\mu}_n^2\hat{\sigma}_i^2 - \frac{\hat{\mu}_i\hat{\mu}_n}{\hat{\sigma}_i\hat{\sigma}_n}\hat{\sigma}_{i,n}^2\right).$$

in which $\hat{\mu}_k^{\text{IS}}$ and $\hat{\Sigma}_k^{\text{IS}}$ are the in-sample mean and variance estimates and $\hat{\text{w}}_k$ is the portfolio obtained with these estimates.

Two, we calculate the *certainty-equivalent return* (CEQ), defined as the risk-free rate that an investor is willing to accept rather than adopting a risky portfolio strategy. Formally, we compute the CEQ of strategy k as

$$\widehat{\text{CEQ}}_k = \hat{\mu}_k - \frac{\gamma}{2}\hat{\sigma}_k^2, \tag{12}$$

in which $\hat{\mu}_k$ and $\hat{\sigma}_k^2$ are the mean and variance of out-of-sample excess returns for strategy k, and γ is risk aversion.[22] The results we report are for the case of $\gamma = 1$; results for other values of γ are discussed in Section 7.7. To test whether the CEQ returns from two strategies are statistically different, we also compute the P-value of the difference, relying on the asymptotic properties of functional forms of the estimators of means and variance.[23]

Three, to get a sense of the amount of trading required to implement each portfolio strategy, we compute the portfolio *turnover*, defined as the average sum of the absolute value of the trades across the N available assets:

$$\text{Turnover} = \frac{1}{T-M}\sum_{t=1}^{T-M}\sum_{j=1}^{N}\left(\left|\hat{\text{w}}_{k,j,t+1} - \hat{\text{w}}_{k,j,t^+}\right|\right), \tag{13}$$

in which $\hat{\text{w}}_{k,j,t}$ is the portfolio weight in asset j at time t under strategy k, $\hat{\text{w}}_{k_j,t^+}$ the portfolio weight *before* rebalancing at $t+1$, and $\hat{\text{w}}_{k,j,t+1}$ the desired portfolio weight at time $t+1$, after rebalancing. For example, in the case of the $1/N$ strategy, $\text{w}_{k,j,t} = \text{w}_{k,j,t+1} = 1/N$, but w_{k,j,t^+} may be different due to changes in asset prices between t and $t+1$. The turnover quantity defined above can be interpreted as the average percentage of wealth traded in each period.

4 Results from the seven empirical datasets considered

In this section, we compare empirically the performances of the optimal asset-allocation strategies listed in Table 1, to the benchmark naive-diversification strategy. For each strategy, we compute the Sharpe ratio (both in-sample and out-of-sample), the certainty-equivalent return, the turnover, and the rank of the strategy relative to the other strategies across all the datasets

[22]To be precise, the definition in equation (12) refers to the level of expected utility of a mean-variance investor, and it can be shown that this is approximately the CEQ of an investor with quadratic utility. Notwithstanding this caveat, and following common practice, we interpret it as the certainty equivalent for strategy k.

[23]If v denotes the vector of moments $v = (\mu_i, \mu_n, \sigma_i, \sigma_n)$, $\hat{v}$ its empirical counterpart obtained from a sample of size $T-M$, and $f(v) = (\mu_i - \frac{\gamma}{2}\sigma_i^2) - (\mu_n - \frac{\gamma}{2}\sigma_n^2)$ the difference in the certainty equivalent of two strategies i and n, then the asymptotic distribution of $f(v)$ is (see Greene (2002)): $\sqrt{T}(f(\hat{v}) - f(v)) \to \mathcal{N}\left(0, \frac{\partial f}{\partial v}^{\top}\Theta\frac{\partial f}{\partial v}\right)$, in which

$$\Theta = \begin{pmatrix} \sigma_i^2 & \sigma_{i,n} & 0 & 0 \\ \sigma_{in} & \sigma_n^2 & 0 & 0 \\ 0 & 0 & 2\sigma_i^4 & 2\sigma_{i,n}^2 \\ 0 & 0 & 2\sigma_{i,n}^2 & 2\sigma_n^4 \end{pmatrix}.$$

listed in Table 2.[24] These quantities are reported in Tables 3–6; in each of these tables, the various strategies being examined are listed in the rows, while the columns refer to the different datasets.

4.1 Sharpe ratios

The first row of Table 3 gives the Sharpe ratio of the naive $1/N$ benchmark strategy for the various datasets being considered.[25] The second row of the table, "mv (in-sample)" gives the Sharpe ratio of the Markowitz mean-variance strategy *in-sample*, that is, when there is no estimation error; by construction this is the highest Sharpe ratio of all the strategies considered. Note that the magnitude of the difference between the in-sample Sharpe ratio for the mean-variance strategy and the $1/N$ strategy gives a measure of the loss from naive rather than optimal diversification when there is no estimation error. For the datasets we are considering, this difference is substantial. For example, for the first dataset considered in Table 3 ("S&PSectors"), the in-sample mean-variance portfolio has a monthly Sharpe ratio of 0.3848, while the Sharpe ratio of the $1/N$ strategy is less than half, only 0.1876. Similarly, in the last column of this table (for the "FF-4-factor" dataset), the in-sample Sharpe ratio for the mean-variance strategy is 0.5364, while that for the $1/N$ strategy is only 0.1753. Similar results hold for all the other datasets, suggesting that *in the absence of estimation error*, there are substantial gains to be made from using the optimal rather than the naive investment strategy.

To assess the magnitude of the potential gains that can actually be realized by an investor, it is necessary to analyze the *out-of-sample* performance of the strategies from the optimizing models. Given that the mean-variance strategy is optimal, the difference between the mean-variance strategy's in-sample and out-of-sample Sharpe ratios allows us to gauge the severity of estimation error. This comparison delivers striking results. From the out-of-sample Sharpe ratio reported in the row titled "mv" in Table 3, we see that for *all* the datasets the sample-based mean-variance strategy has a substantially lower Sharpe ratio out-of-sample than in-sample. Moreover, the out-of-sample Sharpe ratio for the sample-based mean-variance strategy is less that for the naive-diversification strategy for all the datasets. That is, the effect of estimation error is so large that it erodes completely the gains from optimal diversification. For instance, for the dataset "S&PSectors," the sample-based mean-variance portfolio has a Sharpe ratio of only 0.0794 compared to its in-sample value of 0.3848, and 0.1876 for the $1/N$ strategy. For the dataset "International," the in-sample Sharpe ratio for the mean-variance strategy is 0.2090, which drops to -0.0719 out of sample, while the Sharpe ratio of the $1/N$ strategy is 0.1277. Similarly, in the last column of the table, for the "FF-4-factor" dataset, the out-of-sample Sharpe ratio for the mean-variance strategy is -0.0031 compared to its in-sample value of 0.5364, while for the $1/N$ strategy the Sharpe ratio is 0.1753.

As explained in the introduction to the paper, the intuition for the poor out-of-sample performance of the sample-based mean-variance strategy is that the out-of-sample portfolio weights

[24] We do not report the results for the "FF-3-factor" dataset, because these are very similar to the results for the "FF-1-factor" dataset.

[25] Because the $1/N$ strategy does not rely on data, its in-sample and out-of-sample Sharpe ratios are the same.

vary substantially from the in-sample optimal weights. For example, for the "International" dataset, the in-sample mean-variance portfolio weight for the World index is −505% and for the U.S. index it is 290%. Out-of-sample, however, the weight on the World index ranges from −80009% to +256461% and the weight on the U.S. index ranges from −105096% to +32176%. Similarly, for the "Industry" dataset, the in-sample mean-variance optimal weights range from −322% to +145%, but out-of-sample the weights range from −3797409% to +969094%.

The comparisons of Sharpe ratios and portfolio weights above confirm the well known perils of using classical sample-based estimates of the moments of asset returns to implement Markowitz's mean-variance portfolios, and our first observation is that *out of sample*, the $1/N$ strategy outperforms the sample-based mean-variance strategy if one were to make no adjustment at all for the presence of estimation error.

But what about the out-of-sample performance of optimal-allocation strategies that explicitly account for estimation error? Our second observation is that, in general, Bayesian strategies do not seem to be very effective at dealing with estimation error. The Bayes-Stein strategy, "bs", has a lower out-of-sample Sharpe ratio than the $1/N$ strategy for all the datasets except "MKT/SMB/HML," and even in this case the difference is not statistically significant at conventional levels (the P-value is 0.25). In fact, the Sharpe ratio for the Bayes-Stein portfolios is only slightly better than that for the sample-based mean-variance portfolio. The reason why in our datasets the Bayes-Stein strategy is only a small improvement over the out-of-sample mean-variance strategy can be traced back to the fact that while the Bayes-Stein approach does shrink the portfolio weights, the resulting weights are still much closer to the out-of-sample mean-variance weights than to the in-sample optimal weights.[26] The Data-and-Model strategy, "dm," in which the investor is assumed to have a 50% confidence ($\omega = 0.5$) in the validity of an asset pricing model (CAPM, or three- or four-factor APT) when forming a Bayesian prior does improve on the Bayes-Stein approach in four of the six cases.[27] However, even the Bayesian Data-and-Model strategy has a lower Sharpe ratio than the $1/N$ strategy in all the datasets, with this difference being statistically significant for the "Industry," "FF-1-factor" and "FF-4-factor" datasets.

Our third observation is about portfolios that are based on restrictions on the moments of returns. From the row for the minimum-variance strategy titled "min", we see that ignoring the estimates of expected returns altogether but exploiting the information about correlations does lead to better performance relative to the out-of-sample mean-variance strategy. Ignoring mean returns is very successful in reducing the extreme portfolio weights: the out-of-sample portfolio weights under the minimum-variance strategy are much more reasonable than under the sample-based mean-variance strategy. For example, in the "International" dataset, the minimum-variance portfolio weight on the World index ranges from −140% to

[26] The factor that determines the shrinkage of expected returns toward the mean return on the minimum-variance portfolio is $\hat{\phi}$ (see Equation (A1)). For the datasets we are considering, $\hat{\phi}$ ranges from a low of 0.31 for the "FF-4-factor" dataset to a high of 0.66 for the "MKT/SMB/HML" dataset; thus, the Bayes-Stein strategy is still relying too much on the estimated means, $\hat{\mu}$.

[27] As we describe in Section 7.8, our qualitative conclusions are not very sensitive to the choice of ω.

+124% rather than ranging from −148195% to +116828% as it did for the mean-variance strategy. The $1/N$ strategy has a higher Sharpe ratio than the minimum-variance strategy for the datasets "S&PSectors," and "FF-4-factor," while for the "Industry," "International" and "MKT/SMB/HML" datasets the minimum-variance strategy has a higher Sharpe ratio, although the difference is not statistically significant (the P-values are greater than 0.20); only for the "FF-1-factor" dataset is the difference in Sharpe ratios statistically significant. Similarly, the value-weighted market portfolio has a lower Sharpe ratio than the $1/N$ benchmark in all the datasets.[28] The out-of-sample Sharpe ratio for the "mp" approach proposed by MacKinlay and Pastor (2000) represent a clear improvement over the sample-based mean-variance portfolio and most of the Bayesian portfolios; however, the out-of-sample Sharpe ratio of this strategy is still less than that of the $1/N$ strategy for all the datasets we consider.

Our fourth observation is that, contrary to the view commonly held among practitioners, constraints alone do not improve performance sufficiently; that is, the Sharpe ratio of the sample-based mean-variance-*constrained* strategy, "mv-c", is less than that of the benchmark $1/N$ strategy for the "S&PSectors", "Industry", "International" and "MKT/SMB/HML" datasets (with P-values of 0.09, 0.03, 0.17, and 0.02, respectively) while the opposite is true for the "FF-1-factor" and "FF-4-factor" datasets, with the difference being statistically significant only for the "FF-1-factor" dataset. Similarly, the Bayes-Stein strategy with shortsale constraints, "bs-c," has a lower Sharpe ratio than the $1/N$ strategy for the first four datasets, and outperforms the naive strategy only the "FF-1-factor" and "FF-4-factor" datasets, but again the P-value is significant only for the "FF-1-factor" dataset.

Our fifth observation is that, strategies that *combine* portfolio constraints with some form of shrinkage of the mean estimates are usually much more effective in reducing the effect of estimation error. This can be seen, for example, by analyzing the *constrained*-minimum-variance strategy, "min-c", which shrinks completely (by ignoring them) the estimate of expected returns, while at the same time, shrinking the extreme values of the covariance matrix by imposing shortsale constraints. Our results indicate that although the $1/N$ strategy has a higher Sharpe ratio than the "min-c" strategy for the "S&PSectors," and "FF-1-factor" datasets, the reverse is true for the "Industry", "International", "MKT/SMB/HML" and "FF-4-factor"datasets, although the differences are statistically significant only for the "FF-4-factor" dataset. This finding suggests that it may be best to ignore the data on expected returns, but still exploit the correlation structure between assets to reduce risk with the constraints helping to reduce the effect of the error in estimating the covariance matrix. The benefit from combining constraints and shrinkage is also evident for the generalized minimum-variance policy, "g-min-c", which

[28] Note that the small-firm effect cannot be driving this result entirely because we are not allocating wealth across individual assets; and, we should not expect the small-firm effect to be present in the "International" dataset, or to be significant when we are allocating wealth across sector or industry portfolios. In Section 6, we use simulations that are free of the small-firm effect to confirm that the superior performance of the benchmark $1/N$ strategy is not driven by the small-firm effect.

has a higher Sharpe ratio than $1/N$ in all but two datasets, "S&PSectors" and "FF-1-factor," but the difference is statistically significant for only the "FF-4-factor" dataset.[29]

Finally, we consider the two mixture portfolios, "mv-min" and "ew-min", in which the former is a combination of the mean-variance portfolio and the minimum-variance portfolio (see Kan and Zhou (2005)), and the latter is a combination of the $1/N$and minimum-variance portfolios. We find that the "mv-min" strategy has a higher Sharpe ratio than the benchmark $1/N$ strategy only for the "MKT/SMB/HML" dataset, but even here the P-value is 0.22. The second mixture portfolio, "ew-min", outperforms the $1/N$ strategy for the "Industry," "International," "MKT/SMB/HML," and "FF-1-factor" datasets, but the difference is statistically significant only for the "FF-1-factor" dataset.

4.2 Certainty equivalent (CEQ) returns

The comparison of CEQ returns in Table 4 confirms the conclusions from the analysis of Sharpe ratios: in sample, the "mv" strategy has the highest CEQ return, but out of sample none of the strategies from the optimizing models can consistently earn a CEQ return that is statistically superior to that of the $1/N$ strategy. In fact, in only two cases are the CEQ returns from optimizing models statistically superior to the CEQ return from the $1/N$ model. This happens in the "FF-1-factor" dataset, in which the combination portfolio "mv-c" has a CEQ return of 0.0090 and the "bs-c" strategy has a CEQ return of 0.0088, while the $1/N$ strategy has a CEQ of 0.0073, and the P-value of the difference is 0.03 and 0.05, respectively.

For the "MKT/SMB/HML" dataset, the strategies from some of the other optimizing models achieve a higher CEQ return than the $1/N$ strategy, although the difference is not statistically significant; these include the sample-based mean-variance strategy, "mv," the Bayes-Stein strategy, "bs," the Data-and-Model strategy, "dm," the minimum-variance strategy, "min," the value-weighted strategy, "vw," and the Kan and Zhou (2005) strategy, "mv-min." The reason why the optimizing models do well for this dataset, as we show analytically in the next section, is because of the small number of risky assets ($N = 3$) in this dataset. Thus, the number of moments to be estimated is small and so the estimation problem is less severe. Moreover, when the number of assets is small, the naive $1/N$ strategy is less likely to match the benefits from optimal diversification.

For all the datasets other than MKT/SMB/HML, the optimizing models do not perform that well in terms of CEQ returns: The sample-based mean-variance strategy, "mv," the Bayes-Stein strategy, "bs," the Data-and-Model strategy, "dm," the value-weighted portfolio, "vw," and the generalized minimum-variance strategy with constraints, "g-min-c," all have a lower CEQ than the $1/N$ benchmark. The other strategies occasionally have a CEQ that is higher than that of the $1/N$ strategy, but in none of these cases is the difference statistically significant.

[29] The benefit from combining constraints and shrinkage is also present, albeit to a lesser degree, for the constrained Bayes-Stein strategy ("bs-c"), which improves upon the performance of its unconstrained counterpart in all cases except for the "MKT/SMB/HML" dataset, in which the effect of constraints is to generate corner solutions with all wealth invested in a single asset at a particular time.

4.3 Portfolio turnover

Table 5 contains the results for portfolio turnover, our third metric of performance. In Panel A of this table, we report the actual turnover of each strategy, while in Panel B, we report the turnover of the strategies relative to the turnover of the benchmark $1/N$ policy. From this table, we see that in all cases but one, the turnover for the portfolios from the optimizing models is much higher than for the benchmark $1/N$ strategy. The only exception is the "mp" strategy for the "FF-4-factor" dataset, with the turnover being 0.98 of the turnover for the $1/N$ strategy.

Comparing the turnover across the various datasets in Table 5, it is evident that the turnover of the strategies from the optimizing models is smaller relative to the $1/N$ policy in the "MKT/SMB/HML" dataset than in the other datasets. This is not surprising given the fact that two of the three assets in this dataset, HML and SMB, are already actively managed portfolios and, as explained above, because the number of assets in this dataset is small ($N = 3$) and so the estimation problem is less severe.

Comparing the portfolio turnover for the different optimizing models, we see that the turnover for the sample-based mean-variance portfolio, "mv," is substantially greater than that for the $1/N$ strategy. The Bayes-Stein portfolio, "bs," has less turnover than the sample-based mean-variance portfolio, and the Data-and-Model Bayesian approach, "dm," is also successful in reducing turnover relative the mean-variance portfolio. The minimum-variance portfolio, "min," is even more successful in reducing turnover, and the MacKinlay and Pastor (2000) strategy is yet more successful. Also, as one would expect, the strategies with shortsale constraints have much lower turnover than their unconstrained counterparts.

4.4 Summary of findings from the empirical datasets

To summarize our empirical findings, we report in Table 6 a simple ranking of all the strategies using each of the three performance metrics and for each of the datasets. The objective of ranking these strategies is to get a broad sense of the pattern in their relative performance rather than a precise measure, and so no adjustment is made for the magnitude and statistical significance of the difference in performance.

From Panel A of Table 6 we see that, while by construction the in-sample Sharpe ratio is highest for the mean-variance strategy, the out-of-sample Sharpe ratio is typically the lowest for this strategy because it makes no adjustment at all for estimation risk. The Bayes-Stein policy, "bs," and the Data-and-Model approach, "dm," rank just ahead of the sample-based mean-variance strategy; they improve performance only marginally. The strategies that impose moment restrictions, minimum-variance ("min"), value-weighted market portfolio ("vw"), and the MacKinlay and Pastor (2000) strategy ("mp"), rank ahead of the Bayesian models. Imposing constraints improves performance moderately in the case of the Bayes-Stein ("bs-c") strategy, and significantly for the minimum-variance, "min-c," and the generalized version of the minimum-variance, "g-min-c" strategies. In fact, after the in-sample mean-variance strategy, the second-ranking strategy in terms of Sharpe ratio is the generalized minimum-variance with constraints strategy ("g-min-c") followed by the mixture policy ("ew-min") and the minimum-variance constrained strategy ("min-c").

Panel B of Table 6 ranks the strategies in terms of CEQ returns. Again, the best performing strategy is the in-sample mean-variance model. But, if one does not know the true parameters, then it is the $1/N$ strategy that delivers the highest CEQ returns, followed by the "ew-min," "min," and "g-min-c" policies. In general, the ranking among the optimal strategies based on CEQ returns is roughly the same as it was for Sharpe ratios. The main difference is that now the minimum-variance strategy with constraints, "min-c," is ranked 9th, while based on Sharpe ratio it was ranked 4th.

Panel C of Table 6 ranks the strategies in terms of their turnover. Under the assumption that the investor can hold the market portfolio, the strategy of holding just this portfolio (vw) has zero turnover by construction, and so is ranked first. The $1/N$ strategy ranks second in terms of turnover. Of the optimizing models, the generalized and the standard minimum-variance-constrained strategies, "g-min-c" and "min-c," are the best in terms of turnover.

From the above discussion, we conclude that of the strategies from the optimizing models, there is no single strategy that always dominates the others in terms of Sharpe ratio. In general, the $1/N$ strategy has Sharpe ratios that are higher (or statistically indistinguishable) relative to the constrained policies, which, in turn, have Sharpe ratios that are higher than those for the unconstrained policies. In terms of CEQ, no strategy from the optimal models is consistently better than the benchmark $1/N$ strategy. And in terms of turnover, only the "vw" strategy, in which the investor holds the market portfolio and does not trade at all, is better than the $1/N$ strategy.

5 Results from studying analytically the estimation error

In this section, we study analytically some of the determinants of the empirical results identified above. Our objective is to understand why the strategies from the various optimizing models do not perform better relative to the $1/N$ strategy. Our focus is on identifying the relation between the expected performance (measured in terms of the CEQ of returns) of the strategies from the various optimizing models and that of the $1/N$ strategy, as a function of: (i) the number of assets, N; (ii) the length of the estimation window, M; (iii) the ex ante Sharpe ratio of the mean-variance strategy; and, (iv) the Sharpe ratio of the $1/N$ strategy. Unfortunately, closed-form results are available only for the sample-based mean-variance strategy and not for the strategies from the other optimizing models; therefore, in the next section, we will use simulations to perform a complete assessment of the effect of estimation error on all the portfolio policies.

As in Kan and Zhou (2005), we treat the portfolio weights as an *estimator*, that is, as a function of the data. The optimal portfolio can therefore be determined by directly solving the problem of finding the weights that maximize expected utility, instead of first estimating the moments on which these weights depend, and then constructing the corresponding portfolio rules. Applying this insight, we derive a measure of the *expected loss* incurred in using a particular portfolio strategy that is based on estimated rather than true moments.

Let us consider an investor who chooses a vector of portfolio weights, x, to maximize the following mean-variance utility (see equation (2)):

$$U(\mathrm{x}) = \mathrm{x}^\top \mu - \frac{\gamma}{2}\mathrm{x}^\top \Sigma \mathrm{x}. \tag{14}$$

The optimal weight is $\mathrm{x}^* = \frac{1}{\gamma}\Sigma^{-1}\mu$, and the corresponding optimized utility is:

$$U(\mathrm{x}^*) = \frac{1}{2\gamma}\mu\Sigma^{-1}\mu \equiv \frac{1}{2\gamma}S_*^2, \tag{15}$$

in which $S_*^2 = \mu\Sigma^{-1}\mu$ is the squared Sharpe ratio of the *ex ante* tangency portfolio of risky assets. Because μ and Σ are not known, the optimal portfolio weight is also unknown, and is estimated as a function of the available data:

$$\hat{\mathrm{x}} = f(R_1, R_2, \ldots, R_M). \tag{16}$$

We define the *expected loss* from using a particular estimator of the weight $\hat{\mathrm{x}}$ as

$$L(\mathrm{x}^*, \hat{\mathrm{x}}) = U(\mathrm{x}^*) - E[U(\hat{\mathrm{x}})], \tag{17}$$

in which the expectation $E[U(\hat{\mathrm{x}})]$ represents the average utility realized by an investor who "plays" the strategy $\hat{\mathrm{x}}$ infinitely many times.

When using the sample-based mean-variance portfolio policy, $\hat{\mathrm{x}}^{\mathrm{mv}}$, μ and Σ are estimated from their sample counterparts, $\hat{\mu} = \frac{1}{M}\sum_{t=1}^{M} R_t$ and $\hat{\Sigma} = \frac{1}{M}\sum_{t=1}^{M}(R_t - \hat{\mu})(R_t - \hat{\mu})^\top$, and the expression for the optimal portfolio weight is $\hat{\mathrm{x}} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}$. Under the assumption that the distribution of returns is jointly Normal, $\hat{\mu}$ and $\hat{\Sigma}$ are independent and are distributed as follows: $\hat{\mu} \sim \mathcal{N}(\mu, \Sigma/M)$ and $M\,\hat{\Sigma} \sim \mathcal{W}_N(M-1, \Sigma)$, in which $\mathcal{W}_N(M-1, \Sigma)$ denotes a Wishart distribution with $M-1$ degrees of freedom and covariance matrix Σ.

Following an approach similar to that in Kan and Zhou (2005), we derive the expected loss from using the $1/N$ rule. By comparing the expected loss, L_{mv}, from using the sample-based mean-variance policy to the expected loss, L_{ew}, from using the $1/N$ strategy, we can analyze the conditions under which the $1/N$ rule is expected to deliver a lower/higher expected loss than the mean-variance policy. To facilitate the comparison between these policies, we define the *critical value* M_{mv}^* of the sample-based mean-variance strategy as the smallest number of estimation periods necessary for the mean-variance portfolio to outperform, on average, the $1/N$ rule. Formally,

$$M_{\mathrm{mv}}^* \equiv \inf\{M : L_{\mathrm{mv}}(\mathrm{x}^*, \hat{\mathrm{x}}) < L_{\mathrm{ew}}(\mathrm{x}^*, \mathrm{w}^{\mathrm{ew}})\}. \tag{18}$$

Just as in Kan and Zhou (2005), we distinguish between three possible cases: (1) the vector of expected returns is not known, but the covariance matrix of returns is known; (2) the vector of expected returns is known, but the covariance matrix of returns is not; and, (3) both the vector of expected returns and the covariance matrix are unknown and need to be estimated. The following proposition derives the conditions for the $1/N$ rule to outperform the sample-based mean-variance rule for these three cases.

Proposition 1 *Let $S_*^2 = \mu\Sigma^{-1}\mu$ be the squared Sharpe ratio of the tangency (mean-variance) portfolio of risky assets and $S_{ew}^2 = (\mathbf{1}_N^\top\mu)^2/\mathbf{1}_N^\top\Sigma\mathbf{1}_N$ the squared Sharpe ratio of the $1/N$ portfolio. Then:*

1. *If μ is unknown and Σ is known, the 1/N strategy has a lower expected loss than the sample-based mean-variance strategy if*

$$S_*^2 - S_{ew}^2 - \frac{N}{M} < 0. \tag{19}$$

2. *If μ is known and Σ is unknown, the $1/N$ strategy has a lower expected loss than the sample-based mean-variance strategy if*

$$kS_*^2 - S_{ew}^2 < 0, \tag{20}$$

in which:

$$k = \left(\frac{M}{M-N-2}\right)\left(2 - \frac{M(M-2)}{(M-N-1)(M-N-4)}\right) < 1. \tag{21}$$

3. *If both μ and Σ are unknown, the $1/N$ strategy has a lower expected loss than the sample-based mean-variance strategy if*

$$kS_*^2 - S_{ew}^2 - h < 0, \tag{22}$$

in which:

$$h = \frac{NM(M-2)}{(M-N-1)(M-N-2)(M-N-4)} > 0. \tag{23}$$

From the inequality in (19), we see that if μ is unknown but Σ is known, then the $1/N$ strategy is more likely to outperform the sample-based mean-variance strategy if the number of periods over which the parameters are estimated, M, is low and if the number of available assets, N, is high. Because k in (21) is increasing in M and decreasing in N, the inequality in (20) shows that also for the case μ is known but Σ is unknown, $1/N$ is more likely to outperform the sample-based mean-variance policy as N increases and M decreases. Finally, for the case in which both parameters are unknown, we note that, because $h > 0$, the left hand side of (22) is always smaller than the left hand side of (20). Hence, all else being equal, it is more likely that the $1/N$ rule outperforms the mean-variance rule when both the means and covariances are unknown than when only the covariances are unknown.

To illustrate the implications of Proposition 1 above, we compute the critical value M_{mv}^*, as defined in (18), for the three cases considered in the proposition. In Figure 1, we plot the critical length of the estimation period for these three cases, as a function of the number of assets, for different values of the ex-ante Sharpe ratios of the tangency portfolio, S_*, and of the $1/N$ portfolio, S_{ew}. We calibrate our choice of S_* and S_{ew} to the Sharpe ratios reported in Table 3 for empirical data. From Table 3, we see that the in-sample Sharpe ratio for the mean-variance strategy is about 40% for the S&PSectors dataset, about 20% for the Industry and International datasets, and about 15% for the value-weighted market portfolio; so we consider these as the three representative values of the Sharpe ratio of the tangency portfolio: $S_* = 0.40$ (Panels A and B), $S_* = 0.20$ (Panels C and D), and $S_* = 0.15$ (Panels E and F). From Table 3, we also see that the Sharpe ratio for the $1/N$ strategy is about half of that for the in-sample mean-variance strategy. So, in Panel A, we set the Sharpe-ratio of the $1/N$ strategy to be $S_{\text{ew}} = 0.20$ and in Panel C we set this to be 0.10. We also wish to consider a more extreme setting in which the *ex ante* Sharpe ratio of the $1/N$ portfolio is much smaller than that for the mean-variance portfolio—only a quarter rather than a half of the Sharpe ratio of the in-sample mean-variance portfolio, S_*; so, in Panel B we set $S_{\text{ew}} = 0.10$ and in Panel D we set it to 0.05. Similarly, for Panels E and F, which are calibrated to data for the U.S. stock market, we set $S_{\text{ew}} = 0.12$ and $S_{\text{ew}} = 0.08$, respectively; these values are obtained from Tables 7 and 8 for simulated data; the details of this simulated data are provided in Section 6.1.

There are two interesting observations from Figure 1. First, as expected, a large part of the effect of estimation error is attributable to estimation of the mean. We can see this by noticing that the critical value for a given number of assets N increases going from the case in which the mean is known (dash-dotted line) to the case in which it is not known. Second, and most importantly, the magnitude of the critical number of estimation periods is striking. In Panel A, in which the ex-ante Sharpe ratio for the mean-variance policy is 0.40 and that for the $1/N$ policy is 0.20, we see that with 25 assets, the estimation window required for the mean-variance policy to outperform the $1/N$ strategy is more than 200 months; with 50 assets, this increases to about 600 months; and, with 100 assets it is more than 1,200 months. Even for the more extreme case considered in Panel B, in which the Sharpe ratio of the $1/N$ portfolio is only one-fourth that of the mean-variance portfolio, the critical length of the estimation period does not decrease substantially—it is 270 months for 25 assets, 530 months for 50 assets, and 1,060 months for 100 assets..

Reducing the ex-ante Sharpe ratio of the mean-variance portfolio increases the critical length of the estimation window required for it to outperform the $1/N$ benchmark; this explains, at least partly, the relatively good performance of the optimal strategies for the "FF-1-factor" and "FF-4-factor" datasets, for which the Sharpe ratio of the in-sample mean-variance policy is around 0.50, which is much higher than in the other datasets. From Panel C, in which the Sharpe ratio of the mean-variance portfolio is 0.20 and that for the $1/N$ portfolio is 0.10, we see that if there are 25 assets over which wealth is to be allocated, then for the mean-variance strategy that relies on estimation of both mean and covariances to outperform the $1/N$ rule on average, about 1,000 months of data are needed. If the number of assets is 50, the length of the estimation window increases to about 2,000 months. Even in the more extreme case considered in Panel D, in which the $1/N$ rule has a Sharpe ratio that is only one-quarter that of the mean-variance portfolio, with 50 assets the number of estimation periods required for the sample-based mean-variance model to outperform the $1/N$ policy is over 1,500 months.

This is even more striking in Panels E and F, which are calibrated to data for the U.S. stock market. In Panel E, we find that for a portfolio with 25 assets, the estimation window needed for the sample-based mean-variance policy to outperform the $1/N$ policy is more than 3,000 months, and for a portfolio with 50 assets it is more than 6,000 months. In Panel E, in which the Sharpe ratio for the $1/N$ portfolio is only 0.08, for a portfolio with 25 assets, the estimation window needed is more than 1,600 months, and for a portfolio with 50 assets it is more than 3,200 months.

6 Results from simulated data

The results in the section above, although limited to the case of the mean-variance strategy based on sample estimates of the parameters and Normally distributed returns, are nevertheless useful for assessing the loss in performance from having to estimate expected returns and the covariances of returns. In this section, we use simulated data to analyze how the performance of each of the strategies considered in our earlier empirical investigation, relative to the benchmark $1/N$ strategy, depends on the number of assets, N, and the length of the estimation window, M. The main advantage of using simulated data is that we understand exactly their economic and statistical properties. The data we simulate are based on a simple single-factor model, with returns that are IID Normal. Given that most of the models of optimal portfolio choice are derived under these assumptions, this setup should favor the the mean-variance model and its various extensions. It also means that the simulation results are not driven by the small-firm effect, calendar effects, momentum, mean-reversion, fat tails, or other anomalies that have been documented in the literature.

Below, we first describe the details of the simulated dataset and then discuss the results from our experiments.

6.1 Details about how the simulated data is generated

Our approach for simulating returns, and also our choice of parameter values, is similar to that in MacKinlay and Pastor (2000). We assume that the market is composed of a risk-free asset and N risky assets. The risk-free rate is assumed to follow a normal distribution with mean r_f and variance $\sigma^2_{r_f}$. The N risky assets include K factors. The excess returns of the remaining $N-K$ risky assets are generated by the following model,

$$R_{a,t} = \alpha + BR_{b,t} + \epsilon_t, \tag{24}$$

in which $R_{a,t}$ is the $(N-K)$ vector of excess asset returns, α is the $(N-K)$ vector of mispricing coefficients, B is the $(N-K)\times K$ matrix of factor loadings, $R_{b,t}$ is the K vector of excess returns on the factor ("benchmark") portfolios, $R_b \sim N(\mu_b, \Omega_b)$, and ϵ_t is the $(N-K)$ vector of noise, $\epsilon \sim N(0, \Sigma_\epsilon)$, which is serially and cross-sectionally uncorrelated with returns on the factor portfolios.

To initialize the simulation, we need to choose values for: (i) the average risk-free rate, r_f, and its variance, $\sigma^2_{r_f}$; (ii) the mispricing, α; (iii) the factor loading, B; (iv) the mean of the factors, μ_b, and their variance-covariance matrix, Ω_b; and, (v) the variance of the noise, Σ_ϵ.

We choose the risk-free rate to have an annual average of 2% and a standard deviation of 2%. We assume there is only one factor ($K = 1$), whose excess return has an annual average of 8% and standard deviation of 16%. This implies an annual Sharpe ratio of 0.5 for the factor, which is close to the annualized Sharpe ratio for the value-weighted market portfolio in the dataset with the S&P sectors reported in Table 3.[30] The mispricing α is set to zero and the factor loadings, B, for each of the risky assets are evenly spread between 0.5 and 1.5. Finally, the variance-covariance matrix of noise, Σ_ϵ, is assumed to be diagonal.

We consider two different cases for the magnitude of the noise. For the first case, we draw the elements of the diagonal of Σ_ϵ from a uniform distribution with support $[0.10, 0.30]$, so that the cross-sectional average annual idiosyncratic volatility is 20%, which corresponds to what is observed in U.S. stock-market data and is similar to the choice in MacKinlay and Pastor (2000).[31] But, in this case, the in-sample Sharpe ratio of the mean-variance portfolio is not very different from that for the $1/N$ policy, and so it is very difficult for the strategies from the optimizing models to outperform the naive benchmark. So, we consider a second case with much higher idiosyncratic volatility, which differs from what is observed empirically but allows us to study conditions in which the $1/N$ policy may perform poorly. For the second case, we draw the elements of the diagonal of Σ_ϵ from a uniform distribution with support $[0.65, 0.85]$; that is, the cross-sectional average annual idiosyncratic volatility is 75%.

For each of the above cases, we consider different number of assets $N = \{10, 25, 50\}$ and different estimation window lengths of $M = \{120, 360, 6000\}$ months, corresponding to 10, 30, and 500 years, respectively. We use Monte-Carlo sampling to generate monthly return data for $T = 24,000$ months (2,000 years).

6.2 Discussion of results from simulated data

We first discuss the results corresponding to the case with idiosyncratic volatility of 20%. The Sharpe ratios of the various portfolio policies for this case are reported in Tables 7.

Note that for the simulated dataset, we know the true values of the mean and covariance matrix of asset returns, and thus, we can compute the optimal (as opposed to estimated) mean-variance policy. This policy is labelled "mv-true." As expected, the Sharpe ratio of this policy is the highest of all policies, 0.1477, and is close to the Sharpe ratio of the "vw" policy for the "S&PSectors" dataset, 0.1444, which can be interpreted as the Sharpe ratio of the market portfolio.

[30] The monthly Sharpe ratio reported in Table 3 for the value-weighted portfolio (vw) is 0.1444. To obtain the annualized Sharpe ratio, one needs to multiply 0.1444 by $\sqrt{12}$, which gives 0.50.

[31] Note that MacKinlay and Pastor (2000) set the idiosyncratic volatility of *all* assets to be equal to 20%, while we assume that it is 20% on average.

The first thing to note is that the Sharpe ratio of the $1/N$ policy grows with the number of assets. This is because the greater the number of assets available, the larger the benefits from diversification, even if it is naive. The second observation is that the simulations confirm the findings based on the analytic results in the previous section: very long estimation windows are required before the sample-based mean-variance policy, "mv," can achieve an out-of-sample Sharpe ratio that is higher than that for the $1/N$ policy. Moreover, the length of the estimation window needed before the sample-based mean-variance policy outperforms $1/N$ increases substantially with the number of assets. In particular, for the case of 10 risky assets, the sample-based mean-variance policy does not achieve a Sharpe ratio that is higher than that of the $1/N$ policy when M is 120 or 360 months; only for the case of $M = 6,000$ months does it achieve a Sharpe ratio that is higher than that for the $1/N$ policy. And, for the cases with 25 and 50 assets, the sample-based mean-variance does not achieve the same Sharpe ratio as the $1/N$ policy even with an estimation window length of 6,000 months.[32]

Next, we examine the effectiveness of the Bayesian models in dealing with estimation error. Just as we observed in Table 3 for the empirical datasets, the performance of the two Bayesian policies considered is very similar to that of the sample-based mean-variance policy. In particular, the Bayes-Stein policy, "bs," outperforms the $1/N$ benchmark in only those cases as the sample-based mean-variance strategy. The "Data-and-Model" strategy, "dm," performs slightly better than the sample-based mean-variance and Bayes-Stein policies. But, these two Bayesian policies still need 50 years of data to achieve a higher Sharpe ratio than the $1/N$ policy for the case with 10 assets (this number is not reported in the table), and do not outperform the $1/N$ benchmark policy for the cases with 25 and 50 assets even if the estimation window is 6,000 months.

Studying the policies with moment restrictions, we find that the minimum-variance policy, "min," does not beat $1/N$ for any of the cases considered if average idiosyncratic volatility is 20%. The performance of the minimum-variance policy, however, is remarkably stable with respect to the estimation window length; that is, it does not seem to benefit from longer estimation windows as much as the sample-based mean-variance and Bayesian policies. This confirms our previous insight that most of the estimation error is associated with estimating the mean return rather than the covariances. Because the minimum-variance policy uses only estimates of the covariance matrix, it does not benefit much a longer estimation window of 6,000 months. On the other hand, the "mp" policy proposed by MacKinlay and Pastor (2000) does quite well and its performance is similar to that of the $1/N$ policy. The reason for this is that, although this approach does estimate mean asset returns, the estimation of these mean returns is done in conjunction with the estimation of the covariance matrix and, as a result, the overall impact of estimation error is reduced.

[32]One can also use the analytic results in Proposition 1 to compute the critical length of the estimation period that is needed for the sample-based mean-variance policy to outperform the $1/N$ benchmark in terms of CEQ returns. Calibrating to the Sharpe ratios reported in Table 7, from equation (22) we can compute that: if $S_* = 0.1477$, $S_{ew} = 0.1356$, and $N = 10$, then the critical length of the estimation window is 3,022 months; if $N = 25$ and $S_{ew} = 0.1447$, then the critical length of the estimation window is 29,227 months; and, if $N = 50$ and $S_{ew} = 0.1466$, then the critical length of the estimation window is 158,043 months.

The imposition of constraints improves the performance of the sample-based mean-variance policy, "mv-c," only for small estimation window lengths, but worsens its performance for large estimation windows. The intuition for this is that when the estimation window is long, the estimation error (especially for estimating expected returns) is smaller, and therefore, constraints reduce performance. Consequently, the constrained sample-based mean-variance policy, does not outperform the $1/N$ benchmark policy for any of the cases considered. Similarly, imposing shortsale constraints on the Bayes-Stein policy improves the performance relative to its unconstrained counterpart only for short estimation windows, and thus, even with constraints this policy does not outperform the $1/N$ benchmark. Just as in the empirical data, imposing constraints improves the performance of the minimum-variance policy relative to its unconstrained counterpart, but even then the constrained minimum-variance policy, "min-c," does not outperform the $1/N$ benchmark for any of the cases considered. The minimum-variance policy with generalized constraints, "g-min-c," also does not outperform the $1/N$ policy for any of the cases considered, though it does outperform "min-c" for N equal to 25 and 50.

Finally, we examine the mixture portfolio strategies. The performance of the Kan and Zhou (2005) policy, "mv-min," which is a mixture of the mean-variance and minimum-variance policies produces Sharpe ratios that are very similar to those for the sample-based mean-variance policy, "mv," and so it outperforms the $1/N$ only for the case of $N = 10$ when the estimation window is $M = 6000$ months. The second mixture policy, which is a combination of the equally-weighted and minimum-variance portfolios, "ew-min," performs better than the mixture policy for short estimation windows, but is dominated by the minimum-variance policy with constraints, "min-c."

Overall, of all the *optimizing* models considered in Table 7, the one with the highest Sharpe ratio is the MacKinlay and Pastor (2000) strategy, "mp," followed by the minimum-variance-constrained strategy, "min-c."

The results for the case with the average idiosyncratic volatility of 75% are given in Table 8. Our motivation for considering such a high level of idiosyncratic volatility is to identify and analyze situations in which $1/N$ is likely to perform poorly. In particular, note that the Sharpe ratio of the true mean-variance policy, "mv-true," is independent of the idiosyncratic volatility and is still equal to 0.1477, because it is simply the Sharpe ratio of the factor. The Sharpe ratio of the $1/N$ policy with $N = 10$, on the other hand, decreases from 0.1356 to 0.0842 when the idiosyncratic volatility changes from 20% (in Table 7) to 75%.

The main insight from Table 8 is that for the case of $N = 10$ the sample-based mean-variance policy now needs only 360 months of data to deliver a Sharpe ratio that is higher than that for the $1/N$ benchmark. But note that the sample-based mean-variance policy still needs an estimation window length of 6,000 months for the case with 25 assets to outperform the $1/N$ policy, and when there are 50 risky assets even 6,000 months of data is not enough.

Just as before, the performance of both the Bayesian policies, "bs" and "dm," is only slightly superior to that of the sample-based mean-variance policy, but these policies now outperform $1/N$ for $N = \{10, 25, 50\}$ if the estimation window is $M = 6,000$ months. The minimum-variance policy, on the other hand, now outperforms the $1/N$ benchmark in all cases except when the number of assets is large and the estimation window is short ($N = 50$ and $M = 120$ months). The performance of the approach proposed by MacKinlay and Pastor (2000) is quite insensitive to the estimation window length and comparable to that of $1/N$ for the cases with 25 and 50 assets, whereas it is outperformed by the $1/N$ policy in the case of 10 assets.

Imposing constraints on the sample-based mean-variance policy, "mv-c," improves performance relative to the unconstrained policy but still leaves it short of the $1/N$ benchmark except in the case in which the number of assets is small and the estimation window is very long ($N = 10$ and $M = 6000$). The constrained minimum-variance, "min-c," generalized constrained minimum-variance, "g-min-c," and the mixture portfolio of $1/N$ and minimum-variance, "ew-min," outperform the $1/N$-benchmark policy for almost all cases with the exception of the case with 50 assets and an estimation window of only 120 months, while the Kan and Zhou (2005) mixture policy, "mv-min," outperforms $1/N$ only for very long estimation windows. Overall, of all the *optimizing* models considered in Table 8 in which idiosyncratic volatility is quite high, the one with the highest Sharpe ratio is the minimum-variance-constrained strategy, "min-c."

In summary, the simulation results in this section show that for reasonable parameter values, the models of optimal portfolio choice that have been developed specifically to deal with the problem of estimation error reduce only moderately the critical length of the estimation window needed to outperform the $1/N$ policy. Thus, the models of optimal portfolio choice are likely to outperform the $1/N$ benchmark only when the estimation window is long, when the number of assets is small, and when the ex-ante Sharpe ratio of the mean-variance strategy is substantially higher than that of the $1/N$ policy.

7 Results for other specifications: Robustness checks

In the benchmark case reported in Tables 3–6, we have assumed that: (1) the length of the estimation window is $M = 120$ months; (2) the estimation window is rolling, rather than increasing with time; (3) the holding period is one month; (4) the portfolios evaluated are those consisting of only-risky assets; (5) one can invest in also the factor portfolios; (6) the performance is measured relative to the $1/N$-with-rebalancing strategy, rather than the $1/N$-buy-and-hold strategy; (7) the investor has a risk aversion of $\gamma = 1$; (8) the investor's level of confidence in the asset-pricing model is $\omega = 0.50$; and, (9) the investor uses only moments of asset returns to form portfolios, and not also asset-specific characteristics.

Below, we describe the experiments we have undertaken to verify the robustness of our findings to the assumptions listed above. Just as we do for the benchmark case, for each of these experiments, we compute the four tables for the Sharpe ratio, CEQ returns, turnover, and rankings. Because of the large number of tables for these robustness experiments, we have collected them in a separate appendix titled, "Tables with Results for Robustness Checks," that can be downloaded from our website. The section numbers of that document correspond to the subsection numbers below.

7.1 Different lengths of the estimation window

We have reported results for the empirical datasets only the case in which the length of the estimation window is $M = 120$ months. We considered also the case of $M = 60$ months. Having a smaller estimation window typically reduces the performance of the optimal models of asset allocation. As the analytical results of Section 5 and the simulation-based results of Section 6 show, large changes in the length of the estimation window are needed before there is a significant effect on the performance of the optimal models of portfolio choice. This is especially true for the best-performing strategies, which are ones that rely on estimates of the variance-covariance matrix but ignore estimates of expected returns, such as minimum-variance, minimum-variance-constrained, and the generalized-minimum-variance-constrained. Thus, the rankings of the different strategies for the case of $M = 60$ are virtually the same as for the case of $M = 120$ reported in this paper.

7.2 Increasing estimation window, rather than rolling window

All the results reported in the paper are for the case in which estimation is done using a rolling window, that is, where we drop the earliest observation when adding a new observation. We also considered the case in which the estimation window is *increasing* over time, that is, where we do *not* drop the last observation when we add a new observation. We find that the performance of the optimizing strategies improves only slightly in this case. The intuition for this is that, as mentioned above, small changes in the estimation window have only a negligible effect on the results.

7.3 Different holding periods

The results we present are for a holding period of one month, given that the frequency of returns in our datasets is monthly. Because Chan, Karceski, and Lakonishok (1999) and Jagannathan and Ma (2003) consider a holding period of one year, we also considered a holding period of one year and found that it does not affect our qualitative results. The results are almost the same as the ones reported for the benchmark case in the paper.

Also, the Sharpe ratio we report for the myopic portfolio choice models is a one-period Sharpe ratio; that is, it is computed for the case in which agents care about wealth next period. In principle, one might be interested also in the long-horizon Sharpe ratio, that is, the Sharpe ratio computed over the entire holding period of the portfolio. We address this concern in our simulations by considering also the experiment in which we simulate many *paths* of returns, and compute the Sharpe ratio for terminal wealth across these paths (instead of across realizations for a single path); we find that the qualitative insights from these long-horizon Sharpe ratios are similar to the ones that are reported in the paper. Note that these results, because they rely on simulations, are not included in the separate appendix, "Tables with Results for Robustness Checks."

7.4 Portfolios that include the risk-free asset

In the paper, we report results for the performance of the fund of *only-risky assets*, rather than the performance of the overall portfolio, which consists of both the riskless asset and the risky assets. The reason for making this choice was that we wanted to focus on the effect of asset allocation alone, and if one considered the performance of the overall portfolio, then that would depend also on market-timing ability.

If one can hold also the risk-free asset, then for some portfolios there is no change at all in the Sharpe ratio, CEQ returns, and turnover; these are the minimum-variance portfolio, the minimum-variance-constrained portfolio, the generalized-minimum-variance-constrained portfolio, the value-weighted portfolio, and the mixture of the equally-weighted and the minimum-variance portfolio. For the other policies, there are small changes. In terms of Sharpe ratio, the $1/N$ strategy maintains its fifth rank. In terms of CEQ returns, the rank of the $1/N$ strategy drops from second to ninth, though there is still no optimal strategy that consistently outperforms $1/N$.

7.5 Portfolios that exclude the factors as investable assets

In the results that we have reported, we have assumed that the factor portfolios driving returns are available also as investable assets. The factor portfolios in the various datasets have been the U.S. equity market portfolio, the World market portfolio, and the Fama-French factor portfolios SMB, HML and MOM. When we consider the case in which these factor portfolios are *not* available as investable assets, the variance-covariance matrix is better behaved, and so the performances of strategies that rely on just minimizing the portfolio variance improve. These strategies are the minimum-variance portfolio, the Bayes-Stein shrinkage portfolio, the minimum-variance-constrained strategy, and the mixture of the equally-weighted and minimum-variance portfolios. Despite this improvement, the Sharpe ratio of the $1/N$ slips in rank only from fifth to sixth, and in terms of CEQ from second to fourth, while continuing to be the best in terms of turnover.

Note also that most of the investable assets in our datasets are financial assets – stocks and the 90-day treasury bill. One might wonder whether the results would be different if other asset classes, such as commodities and real estate, which have a lower correlation with equities, were included. Note that if wealth was being allocated across assets that were relatively uncorrelated, then keeping all else fixed, the loss from naive rather than optimal diversification when using the $1/N$ rule would be *smaller* rather than larger.

7.6 Benchmark is $1/N$-buy-and-hold rather than $1/N$-with rebalancing

The results reported in the paper compare the strategies from the various models of optimal portfolio choice to the $1/N$ policy in which each period the $1/N$ portfolio is rebalanced to ensure that, given the change in prices of assets, the relative weights are still $1/N$. One could also consider a different benchmark in which investors have inertia and *do not rebalance* even the naive $1/N$ portfolio after the initial date on which the position is established. We find that the results reported are similar to the case in which one uses as the benchmark portfolio the $1/N$-buy-and-hold portfolio, and there is virtually no change in the ranking of the $1/N$ policy relative to the strategies from the optimizing models.

7.7 Different levels of risk aversion

For all the asset allocation policies considered, we have reported results for only the case in which risk aversion is equal to 1. We also considered the following levels of risk aversion: $\gamma = \{2, 3, 4, 5, 10\}$. We find that the results are not very different across risk aversion levels. In particular, the Sharpe ratios for the unconstrained policies are not affected at all by the level of risk aversion, which is a consequence of two-fund separation (and because we are looking at the performance of the fund of just risky assets). The level of risk aversion affects the Sharpe ratio of only the mean-variance-constrained policy ("min-c") and the Bayes-Stein-constrained policy ("bs-c") for which two-fund separation does not hold because the level of risk aversion determines whether the constraints will be binding or not. So, a change in risk aversion has almost no effect on the rank of the $1/N$ strategy evaluated in terms of Sharpe ratios. The only quantity influenced significantly by the choice of risk aversion is the CEQ return. Because there is no optimization in the $1/N$ strategy, as risk aversion increases the rank of the $1/N$ strategy drops: for γ equal to 1 and 2, the $1/N$ strategy ranks second; for γ equal to 3, the $1/N$ strategy ranks fourth; for γ equal to 4, the $1/N$ strategy ranks sixth; and, for γ equal to 5 and 10, the $1/N$ strategy ranks seventh.

7.8 Different levels of confidence in the asset-pricing model

In the Bayesian Data-and-Model ("dm") approach developed in Pástor (2000) and Pástor and Stambaugh (2000), the investor forms a prior based on her subjective belief in a particular asset-pricing model. In the results that we report in the paper, the subjective prior belief in the asset-pricing model is assumed to be $\omega = 0.5$. We also consider the values of $\omega = \{0.25, 0.75\}$ and find that although a change in ω has an effect on the Sharpe ratio, CEQ return, and turnover, the relative performance of the Data-and-Model strategy is not very sensitive to the choice of ω. In particular, in the base case in which $\omega = 0.50$, the rank of the Data-and-Model approach is twelfth in terms of Sharpe ratio, eleventh in terms of CEQ returns, and thirteenth in terms of turnover; for ω equal to 0.25 and 0.75, these ranks are almost unchanged.

7.9 Portfolios that use the cross-sectional characteristics of stocks

In this paper, we have limited ourselves to comparing the performance of models of optimal asset allocation that consider moments of asset returns but not other characteristics of the assets. Brandt, Santa-Clara, and Valkanov (2005) propose a new approach for constructing the optimal portfolio that exploits the cross-sectional characteristics of equity returns. Their idea is to model the portfolio weights in firm i as a benchmark weight plus a linear function of firm i's characteristics:

$$\hat{\mathrm{w}}_{i,t} = \overline{\mathrm{w}}_{i,t} + \frac{1}{N_t}\,\theta^\top\,\hat{y}_{i,t}, \tag{25}$$

in which $\overline{\mathrm{w}}_{i,t}$ is the weight of firm i in the value-weighted portfolio, N_t is the number of firms at time t, and $\hat{y}_{i,t}$ is a vector of firm-specific characteristics. The optimal portfolio at time t is obtained by finding the parameters θ that maximize expected utility over the next period:

$$\theta = \operatorname{argmax}\, E_t\left[u\left(\sum_{i=1}^{N_t}\left(\overline{\mathrm{w}}_{i,t} + \frac{1}{N_t}\,\theta^\top\,\hat{y}_{i,t}\right)R_{i,t+1}\right)\right], \tag{26}$$

in which $R_{i,t+1}$ is the return on firm i. The construction of such a portfolio, hence, reduces to a statistical estimation problem, and the low dimensionality of the problem allows one to avoid problems of over-fitting. In their application to the universe of CRSP stocks (1964–2002), Brandt, Santa-Clara, and Valkanov (2005) find that portfolios tend to load on small, value, and past-winners stocks.

In order to compare the out-of-sample performance of the Brandt, Santa-Clara, and Valkanov (2005) methodology relative to the benchmark $1/N$ portfolio, we apply their approach to the two datasets in our paper for which the investable assets have asset-specific characteristics similar to the ones that they use in their analysis. These two datasets, taken from Kenneth French's website, are: (1) 10 industry portfolios, which are described in Appendix B.1, and (2) 25 size- and book-to-market-sorted portfolios, which are described in Appendix B.5. As we did in our empirical analysis, we use the rolling-window approach with the length of the estimation period being 120 months.

For the 10 industry portfolios, the Brandt, Santa-Clara, and Valkanov model has an out-of-sample Sharpe ratio of 0.181 while that of the $1/N$ strategy is only 0.139, with the P-value for the difference being 0.212. And, the turnover for the Brandt, Santa-Clara, and Valkanov model is about 66 times greater than that for the $1/N$ strategy. For the 25 size- and book-to-market-sorted portfolios, we find that using the Brandt, Santa-Clara, and Valkanov (2005) model gives an out-of-sample Sharpe ratio of 0.234, which is higher than that for the $1/N$ strategy, 0.165, but the P-value for the difference is 0.131. Again, the turnover for the Brandt, Santa-Clara, and Valkanov portfolio is about 100 times greater than that for the $1/N$ strategy. So, for both datasets the Sharpe ratio of the Brandt, Santa-Clara, and Valkanov strategy is higher but statistically indistinguishable from that of the $1/N$ benchmark, while the turnover is substantially lower for the $1/N$ strategy.

However, for both datasets, the performance of the Brandt, Santa-Clara, and Valkanov (2005) approach improves if one can partition more finely the assets so that the number of assets available for investment is much larger. For instance, if one has 48 instead of 10 industry portfolios, the Sharpe ratio increases to 0.225 relative to only 0.139 for the $1/N$ strategy, and the P-value for the difference is 0.094. However, the turnover for the Brandt, Santa-Clara, and Valkanov is now 135 times that for the $1/N$ strategy. Similarly, if one has a 100 instead of just 25 size- and book-to-market-sorted portfolios, then the Sharpe ratio improves to 0.364, while that of the $1/N$ strategy is 0.122, with the difference being statistically significant; however, the turnover for the Brandt, Santa-Clara, and Valkanov strategy is now 143 times greater than that for the $1/N$ benchmark strategy.

We conclude from this experiment that using information about the cross-sectional characteristics of assets, rather than just statistical information about the moments of asset returns, *does* lead to an improvement in Sharpe ratios. If the number of investable assets is relatively small, then the improvement in performance relative to the benchmark $1/N$ strategy may not be statistically significant. However, when the number of investable assets is large, and hence, the potential for over-fitting more severe, then the difference in Sharpe ratios is statistically significant. However, the turnover of the Brandt, Santa-Clara, and Valkanov (2005) approach is substantially higher than that for the $1/N$ strategy, and this difference increases with the number of investable assets. Moreover, it may not be possible to use the methodology of Brandt, Santa-Clara, and Valkanov for all asset classes; for example, if one wished to allocate wealth across international stock indexes, then it is not clear what cross-sectional characteristics to associate with these country indexes.

8 Conclusions

We have compared the performance of fourteen models of optimal asset allocation, relative to that of the benchmark $1/N$ policy. This comparison is undertaken using seven different empirical datasets as well as using simulated data. We find that the *out-of-sample* Sharpe ratio of the mean-variance strategy is much lower than that of the $1/N$ strategy, indicating that the errors in estimating means and covariances destroys all the gains from optimal, relative to naive, diversification. We also find that various models proposed in the literature to deal with the problem of estimation error typically do not outperform the $1/N$ benchmark for the seven empirical datasets. For instance, the the out-of-sample performance of the Bayesian strategies of Jorion (1985), Pástor (2000), and Pástor and Stambaugh (2000) is similar to that of the sample-based mean-variance model. Similarly, the minimum-variance strategy, which ignores expected returns but exploits the correlation structure, is unable to achieve a Sharpe ratio that is higher than that of the $1/N$ strategy. Even the strategies proposed recently in Kan and Zhou (2005) and MacKinlay and Pastor (2000) have Sharpe ratios that are lower than those attained by the $1/N$ strategy. Imposing shortselling constraints on the sample-based mean-variance and Bayes-Stein strategies improves performance only moderately. However, imposing constraints on the minimum-variance strategy does lead to a significant improvement in performance. In summary, our results indicate that, of the various optimizing models in the literature, there is no single model that consistently delivers a Sharpe ratio or a certainty-equivalent return that is higher than that of the $1/N$ portfolio, which also has a very low turnover.

To understand the poor performance of the optimizing models, we derive analytically the length of the estimation period needed before the sample-based mean-variance strategy can be expected to achieve a higher certainty-equivalent return than the $1/N$ benchmark. For parameters calibrated to U.S. stock-market data, we find that for a portfolio with only 25 assets, the estimation window needed is more than 3,000 months, and for a portfolio with 50 assets it is more than 6,000 months, while typically these parameters are estimated using 120 months of data. This result illustrates the severity of the problem of estimating the moments of asset returns for making asset-allocation decisions. Using simulated data we show that the various extensions to the sample-based mean-variance model that have been designed to incorporate estimation error reduce only moderately the estimation window needed for these models to outperform the naive $1/N$ benchmark.

These findings have two implications. One, while there has been considerable progress in the design of optimal portfolios, more energy needs to be devoted to improving the estimation of the moments of asset returns and using not just statistical but also other available information about stock returns. As our evaluation of the approach proposed in Brandt, Santa-Clara, and Valkanov (2005) shows, exploiting information about the cross-sectional characteristics of assets, may be a promising direction to pursue. Two, when evaluating the performance of a particular strategy for optimal asset allocation, the $1/N$ naive-diversification rule should serve at least as a first obvious benchmark. We emphasize, however, that we are not proposing the $1/N$ rule as the best asset-allocation policy for all situations but merely as a reasonable benchmark to assess the performance of several alternative portfolio strategies. In fact, we have identified the circumstances in which the $1/N$ strategy will underperform strategies from optimizing models: when the number of investable assets is small, when the estimation window is long, and when the in-sample Sharpe ratio of the mean-variance portfolio is substantially higher than that for the $1/N$ portfolio.

A Details for some of the models of optimal portfolio choice

In this section, we provide additional details for some of the models of optimal portfolio selection that were described in the main text.

A.1 Bayes-Stein shrinkage portfolio

In our analysis of the Bayes-Stein shrinkage portfolio, we use the estimator proposed by Jorion (1985, 1986) who takes the grand mean, $\overline{\mu}$, to be the mean of the minimum-variance portfolio, $\mu^{\min}$. More specifically, following Jorion (1986), the estimator we use for expected returns is:

$$\hat{\mu}_t^{\text{bs}} = (1-\hat{\phi}_t)\,\hat{\mu}_t + \hat{\phi}_t\,\hat{\mu}_t^{\min}, \tag{A1}$$

$$\hat{\phi}_t = \frac{N+2}{(N+2) + M(\hat{\mu}_t - \mu_t^{\min})^\top \hat{\Sigma}_t^{-1}(\hat{\mu}_t - \mu_t^{\min})}, \tag{A2}$$

in which $\hat{\Sigma}_t = \frac{1}{M-N-2}\sum_{s=t-M+1}^{t}(R_s - \hat{\mu}_t)(R_s - \hat{\mu}_t)^\top$, and $\hat{\mu}_t^{\min} \equiv \hat{\mu}_t^\top \hat{\text{w}}_t^{\min}$ is the average excess return on the sample global minimum-variance portfolio, $\hat{\text{w}}_t^{\min}$.

The implementation in Jorion (1986) accounts *also* for estimation error in the covariance matrix via a traditional Bayesian estimation. To accomplish this, Jorion (1986), derives first the predictive variance of asset returns starting from an informative prior on μ with precision τ, and then uses the sample estimates of Σ and τ to obtain the following expression for the covariance matrix that is utilized in portfolio construction:

$$\hat{\Sigma}_t^{\text{bs}} = \hat{\Sigma}_t\left(1 + \frac{1}{M+\hat{\tau}_t}\right) + \frac{\hat{\tau}_t}{M(M+1+\hat{\tau}_t)}\frac{\mathbf{1}_N \mathbf{1}_N^\top}{\mathbf{1}_N^\top \hat{\Sigma}_t^{-1}\mathbf{1}_N}, \tag{A3}$$

$$\hat{\tau}_t = M\frac{\hat{\phi}_t}{1-\hat{\phi}_t}. \tag{A4}$$

The portfolio obtained by using $\hat{\mu}_t^{\text{bs}}$ and $\hat{\Sigma}_t^{\text{bs}}$ is:

$$\hat{\text{w}}_t^{\text{bs}} = \frac{\left(\hat{\Sigma}_t^{\text{bs}}\right)^{-1}\hat{\mu}_t^{\text{bs}}}{\mathbf{1}_N^\top \left(\hat{\Sigma}_t^{\text{bs}}\right)^{-1}\hat{\mu}_t^{\text{bs}}}, \tag{A5}$$

which combines both a shrinkage approach and a traditional Bayesian estimation, and hence, is known as the "Bayes-Stein" portfolio.

A.2 Bayesian "Data-and-Model" portfolio

For evaluating the Bayesian "Data-and-Model" approach in Pástor (2000), we assume that the market is composed of N risky assets and one risk-free asset. The N risky assets include K factors. The excess returns of the remaining $N-K$ risky assets follow the factor model defined below:

$$R_{a,t} = \alpha + BR_{b,t} + \epsilon_t. \tag{A6}$$

Then, the "true" mean vector and variance-covariance matrix of returns are:

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}, \quad \Sigma = \begin{pmatrix} B\Sigma_{bb}B^\top + \Sigma_\epsilon & B\Sigma_{bb} \\ \Sigma_{bb}B^\top & \Sigma_{bb} \end{pmatrix}, \tag{A7}$$

in which Σ_{bb} is the variance-covariance matrix of the returns on the factors. If the asset pricing model is correct, then $\mu_a = B\mu_b$ and the mispricing term α is zero.

Let $\hat{B}$ and $\hat{\Sigma}_\epsilon$ be the maximum likelihood estimators of the factor loadings B and of the variance-covariance matrix of residuals, Σ_ϵ, obtained by estimating the unrestricted regression in (A6). These will be the estimators chosen by an investor who completely ignores the prediction of the asset pricing model. Similarly, let $\overline{B}$ and $\overline{\Sigma}_\epsilon$ be the quantities obtained by estimating the model (24) with the restriction that $\alpha = 0$. These would be the estimators chosen by an investor who dogmatically believes in the asset pricing model.

A Bayesian investor expresses his confidence in the asset-pricing model by postulating a *prior* belief on the mispricing term, α. The prior on α, conditional on Σ_ϵ, is assumed to have a Normal distribution with a mean of zero and variance $\tau\Sigma_\epsilon$, that is, $p(\alpha|\Sigma_\epsilon) = \mathcal{N}(0, \tau\Sigma_\epsilon)$, with τ determining the precision of the prior belief over the validity of the asset-pricing model.[33] Finally, let $\hat{\mu}_a$ denote the sample mean of the return on the assets, R_{at}, and $\hat{\mu}_b$ and $\hat{\Sigma}_{bb}$ be, respectively, the sample mean and variance of returns on the factors, R_{bt}.

Under the assumptions described above, Wang (2005) shows how to obtain estimators for the mean vector and covariance matrix of asset returns that account for the belief of a Bayesian investor about the validity of a particular asset pricing model. More specifically, (see Wang (2005, Theorem 1)) if one denotes by $\hat{S}^2 = \hat{\mu}_b^\top \hat{\Sigma}_{bb}^{-1} \hat{\mu}_b$ the square of the highest Sharpe ratio of the efficient frontier spanned by the mean and variance of the factor portfolios, and by $\omega = \frac{1}{1+M\tau/(1+\hat{S}^2)}$ the degree of confidence a Bayesian investor places in the asset pricing model (that is, $\omega = 1$ implies dogmatic belief in the model), then a Bayesian Data-and-Model ("dm") investor with a degree of confidence ω in the model will use the following *shrinkage* estimators of the expected return and variance-covariance matrix of the investable assets:

$$\mu^{\text{dm}} = \omega \begin{pmatrix} \overline{B}\hat{\mu}_a \\ \hat{\mu}_b \end{pmatrix} + (1-\omega) \begin{pmatrix} \hat{\mu}_a \\ \hat{\mu}_b \end{pmatrix}, \tag{A8}$$

$$\Sigma^{\text{dm}} = \begin{pmatrix} \Sigma_{aa}^{\text{dm}}(\omega) & \Sigma_{ab}^{\text{dm}}(\omega) \\ \Sigma_{ab}^{\text{dm}}(\omega)^\top & \Sigma_{bb}^{\text{dm}}(\omega) \end{pmatrix}, \tag{A9}$$

in which

$$\begin{aligned} \Sigma_{aa}^{\text{dm}}(\omega) &= \kappa\,(\omega\overline{B} + (1-\omega)\hat{B})\,\hat{\Sigma}_{bb}\,(\omega\overline{B} + (1-\omega)\hat{B})^\top \\ &\quad + h\,(\omega\overline{\delta} + (1-\omega)\hat{\delta})\,(\omega\overline{\Sigma}_\epsilon + (1-\omega)\hat{\Sigma}_\epsilon), \end{aligned} \tag{A10}$$

$$\Sigma_{ab}^{\text{dm}}(\omega) = \kappa\,(\omega\overline{B} + (1-\omega)\hat{B})\,\hat{\Sigma}_{bb}, \tag{A11}$$

$$\Sigma_{bb}^{\text{dm}}(\omega) = \kappa\hat{\Sigma}_{bb}, \tag{A12}$$

[33] The priors on the factor loadings, B, on the variance-covariance of the residuals, Σ_ϵ, as well as on the expected returns and variance-covariance matrix of the factors, are assumed to be non-informative because the asset-pricing model does not impose any restrictions on these parameters.

and

$$\bar{\delta} = \frac{M(M-2)+K}{M(M-K-2)} - \frac{k+3}{M(M-K-2)} \cdot \frac{\hat{S}^2}{1+\hat{S}^2}, \quad \text{(A13)}$$

$$\hat{\delta} = \frac{(M-2)(M+1)}{M(M-K-2)}, \quad \text{(A14)}$$

$$\kappa = \frac{M+1}{M-K-2}, \quad \text{(A15)}$$

$$h = \frac{M}{M-N-1}. \quad \text{(A16)}$$

Observe, from Equation (A8) that the parameter ω acts as a *linear* shrinkage factor for the mean returns, and from Equation (A9) as a *quadratic* shrinkage factor for the variance of the returns on the assets. The resulting portfolio weights at each time t are, given by

$$\hat{\mathrm{w}}_t^{\mathrm{dm}} = \frac{\left(\hat{\Sigma}_t^{\mathrm{dm}}\right)^{-1} \hat{\mu}_t^{\mathrm{dm}}}{\mathbf{1}_N^\top \left(\hat{\Sigma}_t^{\mathrm{dm}}\right)^{-1} \hat{\mu}_t^{\mathrm{dm}}}. \quad \text{(A17)}$$

A.3 Portfolio implied by asset pricing models with unobservable factors

The optimal portfolio strategy from MacKinlay and Pastor (2000), denoted, x^{mp}, is

$$\mathrm{x}^{\mathrm{mp}} = \frac{1}{\gamma}\Sigma^{-1}\mu = \frac{1}{\gamma}\frac{1}{\sigma^2}\left(I - \frac{\nu\mu\mu^\top}{\sigma^2+\nu\mu^\top\mu}\right)\mu = \frac{1}{\gamma(\sigma^2+\nu\mu^\top\mu)}\mu, \quad \text{(A18)}$$

in which ν, μ and σ^2 are obtained by maximizing the likelihood function obtained by imposing the structure described in (5) on the residual covariance matrix. Instead of implementing this strategy by numerically optimizing the log-likelihood function, we follow the approach suggested by Kan and Zhou (2005), who show that the MacKinlay and Pastor portfolio can be approximated as

$$\hat{\mathrm{x}}_t^{\mathrm{mp}} \approx \frac{\tilde{\mu}_t}{\gamma(\hat{\eta}_{1t} - \tilde{\mu}_t^\top \tilde{\mu}_t)}, \quad \text{(A19)}$$

in which $\hat{\eta}_{1t}$ is the largest eigenvalue of the matrix $\widehat{\Sigma}_t + \hat{\mu}_t\hat{\mu}_t^\top$, $\hat{\mu}_t$ and $\widehat{\Sigma}_t$ are the sample mean and covariance, and $\tilde{\mu}_t = (\hat{q}_{1t}\hat{\mu}_t\hat{q}_{1t})$, with $\hat{q}_{1t}$ the eigenvector associated with $\hat{\eta}_{1t}$. The corresponding normalized portfolio weights simplify to $\hat{\mathrm{w}}_t^{\mathrm{mp}} \approx \tilde{\mu}_t/\mathbf{1}_N^\top\tilde{\mu}_t$. Note that for $\tilde{\mu}_t \propto \mathbf{1}_N$, this strategy corresponds to the $1/N$ portfolio. Clearly the performance of this rule depends on how reasonable is the restriction (5).

A.4 Optimal "Three-Fund" portfolio

After choosing the parameters c and d in (8) to maximize the expected utility of a mean-variance investor, Kan and Zhou (2005) propose the following "three-fund" (mv-min) weights:

$$\hat{\mathrm{x}}^{\mathrm{mv\text{-}min}} = \frac{(M-N-1)(M-N-4)}{\gamma\, M\,(M-2)}\left[\left(\frac{\hat{\psi}_a^2}{\hat{\psi}_a^2+\frac{N}{M}}\right)\hat{\Sigma}^{-1}\hat{\mu} + \left(\frac{\frac{N}{M}}{\hat{\psi}_a^2+\frac{N}{M}}\right)\hat{\mu}^{\min}\hat{\Sigma}^{-1}\mathbf{1}_N\right], \quad \text{(A20)}$$

in which, defining the Incomplete Beta function by $B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}dy$,

$$\hat{\psi}_a^2 = \frac{(M-N-1)\hat{\psi}^2 - (N-1)}{M} + \frac{2(\hat{\psi}^2)^{\frac{N-1}{2}}(1+\hat{\psi}^2)^{-\frac{M-2}{2}}}{M\, B_{\hat{\psi}^2/(1+\hat{\psi}^2)}((N-1)/2, (M-N+1)/2)}, \tag{A21}$$

$$\hat{\psi}^2 = (\hat{\mu} - \hat{\mu}^{\min})^\top \hat{\Sigma}^{-1}(\hat{\mu} - \hat{\mu}^{\min}). \tag{A22}$$

A.5 Combination of minimum-variance strategy and the naive $1/N$ strategy

Because the investor is assumed to be ignoring expected returns, the expected-utility maximization is equivalent to minimizing the portfolio variance in which Σ is unknown. Therefore, the objective is to minimize the expected portfolio variance:

$$E[U(\hat{\mathrm{w}}(c,d))] = \frac{1}{2}E\left[\left(c\frac{1}{N}\mathbf{1}_N + d\,\hat{\Sigma}^{-1}\mathbf{1}_N\right)^\top \Sigma \left(c\frac{1}{N}\mathbf{1}_N + d\,\hat{\Sigma}^{-1}\mathbf{1}_N\right)\right] \tag{A23}$$

$$= \frac{1}{2}E\left[\frac{c^2}{N^2}\mathbf{1}_N^\top\Sigma\mathbf{1}_N + d^2\mathbf{1}_N^\top\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{1}_N + 2\frac{c\,d}{N}\mathbf{1}_N^\top\Sigma\hat{\Sigma}^{-1}\mathbf{1}_N\right]. \tag{A24}$$

Let $W \equiv \Sigma^{-\frac{1}{2}}\widehat{\Sigma}\Sigma^{-\frac{1}{2}}$. Then, $W^{-1} = \Sigma^{\frac{1}{2}}\widehat{\Sigma}^{-1}\Sigma^{\frac{1}{2}}$ and $W^{-2} = \Sigma^{\frac{1}{2}}\widehat{\Sigma}^{-1}\Sigma\widehat{\Sigma}^{-1}\Sigma^{\frac{1}{2}}$. The expectation in (A24) may hence be written as

$$\frac{c^2}{2N^2}\mathbf{1}_N^\top\Sigma\mathbf{1}_N + E\left[\frac{1}{2}d^2\mathbf{1}_N^\top\Sigma^{-1/2}W^{-2}\Sigma^{-1/2}\mathbf{1}_N + \frac{c\,d}{N}\mathbf{1}_N^\top\Sigma\Sigma^{-1/2}W^{-1}\Sigma^{-1/2}\mathbf{1}_N\right]. \tag{A25}$$

Because $M\,W \sim \mathcal{W}_N(M-1, I_N)$, it follows that (see Haff (1979)):

$$E[W^{-1}] = \left(\frac{M}{M-N-2}\right)I_N \tag{A26}$$

$$E[W^{-2}] = kI_N, \tag{A27}$$

in which

$$k = \frac{M^2(M-2)}{(M-N-1)(M-N-2)(M-N-4)}. \tag{A28}$$

Thus,

$$E[U(\hat{\mathrm{w}}(c,d))] = \frac{c^2}{2N^2}\mathbf{1}_N^\top\Sigma\mathbf{1}_N + \frac{1}{2}k\,d^2\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N + \frac{M\,c\,d}{M-N-2}. \tag{A29}$$

We want to find the values of c and d that minimize the expected value in (A29) subject to the constraint that the portfolio weights sum up to one. By writing the corresponding first-order optimality conditions, and simplifying, one gets: $c = 1 - d\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N$, and

$$d = \frac{(M-N-2)(\mathbf{1}_N^\top\Sigma\mathbf{1}_N)(\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N) - N^2M}{N^2(M-N-2)k(\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N) - 2M\,N^2(\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N) + (M-N-2)(\mathbf{1}_N^\top\Sigma^{-1}\mathbf{1}_N)^2(\mathbf{1}_N^\top\Sigma\mathbf{1}_N)}.$$

B Description of the seven empirical datasets

This appendix describes the seven empirical datasets considered in our study. Each dataset contains excess monthly returns over the 90-day T-bill (from Ken French's website). A list of the datasets considered is given in Table 2.

B.1 Sector portfolios ("S&PSectors")

This dataset consists of monthly excess returns on ten portfolios tracking the sectors composing the S&P500 index.[34] The data span from January, 1981 to December, 2002. We augment the dataset by adding as a factor the excess return on the U.S. equity market portfolio, MKT, defined as the value-weighted return on all NYSE, AMEX and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate. The choice of this dataset was motivated by the fact that if one wished to consider an allocation across just domestic equities, then one logical class of assets to consider would be sector portfolios.

B.2 Industry portfolios ("Industry")

This dataset consists of monthly excess returns on ten industry portfolios in the United States. The ten industries considered are: Consumer-Discretionary, Consumer-Staples, Manufacturing, Energy, High-Tech, Telecommunication, Wholesale and Retail, Health, Utilities, and Other. The monthly returns range from July 1963 to November 2004 and were obtained from Kenneth French's website. We augment the dataset by adding as a factor the excess return on the U.S. equity market portfolio, MKT. The motivation for considering this dataset is similar to that for the dataset considered above: if one wished to consider an allocation across only domestic equities, then one way to classify assets would be by industry.

B.3 International equity indexes ("International")

The eight international equity indices considered are for: Canada, France, Germany, Italy, Japan, Switzerland, United Kingdom, are United States. In addition to these country indexes, the World index is used as the factor portfolio. Returns are computed based on the month-end US-dollar value of the country equity index for the period January 1970 to July 2001. Data are from MSCI (Morgan Stanley Capital International). There are two motivations for considering this data: one, it does not limit the investor to just domestic assets; two, this data is similar to that used in Jorion (1985, 1986), where the Bayes-Stein shrinkage model is proposed.

B.4 MKT, SMB and HML portfolios ("MKT/SMB/HML")

This dataset is an updated version of the one used by Pástor (2000) for evaluating the Bayesian "Data-and-Model" approach to asset allocation. The assets are represented by three broad portfolios: (i) MKT, that is, the excess return on the U.S. equity market, (ii) HML, a zero-cost

[34] We are grateful to Roberto Wessels for creating this dataset and for making it available to us.

portfolio that is long in high book-to-market stocks and short in low book-to-market stocks, and (iii) SMB, a zero-cost portfolio that is long in small-cap stocks and short in large-cap stocks. The data consists of monthly returns from July 1963 to November 2004. The data are taken from Kenneth French's website. The Data-and-Model approach is implemented by assuming that the investor takes into account his beliefs in an asset pricing model (CAPM or three- or four-factor APT) when constructing expected asset returns.

B.5 Size- and Book-to-Market-sorted portfolio ("FF-1-factor," "FF-3-factor," "FF-4-factor")

The data consist of monthly returns on the twenty portfolios sorted by size and book-to-market.[35] The data are obtained from Kenneth French's website and span from July 1963 to December 2004. This dataset is the one used by Wang (2005) to analyze the shrinkage properties of the Data-and-Model approach. We use this dataset for three different experiments. In the first, we augment the dataset by adding the market portfolio (MKT). We then impose that a Bayesian investor takes into account his beliefs in the CAPM to construct estimates of expected returns. In the second, we augment the dataset by adding the market portfolio (MKT), and the zero cost portfolios HML and SMB. We now assume that a Bayesian investor uses the Fama-French three-factor model to construct estimates of expected returns. In the third experiment, we augment the size- and book-to-market-sorted portfolios with four factor portfolios: MKT, HML, SMB and the momentum portfolio, MOM, which is also obtained from Kenneth French's website. For this dataset, the investor is assumed to estimate expected returns using a four-factor model.

C Proof for Proposition 1

Assuming that the distribution of returns is jointly Normal, Kan and Zhou (2005) derive the following expression for the expected loss from using the sample-based mean-variance policy with estimated rather than the true parameters: When μ is unknown but Σ is known, the expected loss is:

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}|\Sigma) = \frac{1}{2\gamma}\frac{N}{M}. \tag{C1}$$

When μ is known but Σ is unknown, the expected loss is:

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}|\mu) = \frac{1}{2\gamma}(1-k)S_*^2, \tag{C2}$$

with k given in Equation (21). When both μ and Σ are unknown, the expected loss is:

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}) = \frac{1}{2\gamma}\left((1-k)S_*^2 + h\right), \tag{C3}$$

with k given in (21) and h given in (23).

[35] As in Wang (2005), we exclude the five portfolios containing the largest firms because the market, SMB and HML are almost a linear combination of the twenty-five Fama-French portfolios.

Following a similar approach, we derive the expected loss from using the $1/N$ rule. Formally, let x^{ew} denote the equally-weighted policy:

$$\mathrm{x}^{\mathrm{ew}} = c\,\mathbf{1}_N, \quad c \in \mathbb{R}. \tag{C4}$$

Note that the weights in the above expressions do not necessarily correspond to $1/N$, but the *normalized* weights do, independent of the choice of the scalar $c \in \mathbb{R}$. To clarify, if initial wealth is one dollar, then Nc represents the fraction invested globally in the risky assets, and $1 - Nc$ the fraction invested in the risk-free asset.

Suppose the investor uses the rule (C4) for a generic c. The expected loss from using such a rule instead of the one that relies on perfect knowledge of the parameters is then

$$L(\mathrm{x}^*, \mathrm{x}^{\mathrm{ew}}) = U(\mathrm{x}^*) - c\mathbf{1}_N^\top \mu + \frac{\gamma}{2} c^2 \mathbf{1}_N \Sigma \mathbf{1}_N. \tag{C5}$$

In order to fully isolate the cost of using the $1/N$ rule and avoid the effects of poor market timing, we assume that the investor chooses c optimally, i.e., in such a way that the loss in (C5) is minimized. Since the loss function is convex in c, the lowest possible loss is obtained by choosing $c^* = \frac{\mathbf{1}_N^\top \mu}{\gamma \mathbf{1}_N \Sigma \mathbf{1}_N}$, which delivers the following *lowest bound* on the loss from using the $1/N$ portfolio rule:

$$L_{\mathrm{ew}}(\mathrm{x}^*, \mathrm{x}^{\mathrm{ew}}) = \frac{1}{2\gamma}\left(\mu^\top \Sigma^{-1} \mu - \frac{(\mathbf{1}_N^\top \mu)^2}{\mathbf{1}_N^\top \Sigma \mathbf{1}_N}\right) \equiv \frac{1}{2\gamma}\left(S_*^2 - S_{\mathrm{ew}}^2\right), \tag{C6}$$

in which $S_{\mathrm{ew}}^2 = \frac{(\mathbf{1}_N^\top \mu)^2}{\mathbf{1}_N^\top \Sigma \mathbf{1}_N}$ is the squared Sharpe ratio of the $1/N$ portfolio. Comparing the expressions (C1), (C2), and (C3) to the expression (C6) gives the result in the proposition. ∎

Table 1: List of various asset-allocation models considered

This table lists the various asset-allocation models we consider. The last column of the table gives the abbreviation used to refer to the strategy in the tables where we compare the performance of the optimal portfolio strategies to that of the $1/N$ strategy. The results for two strategies are not reported. The reason for not reporting the results for the Bayesian diffuse-prior strategy is that for an estimation period that is of the length that we are considering (60 or 120 months), the Bayesian diffuse-prior portfolio is very similar to the sample-based mean-variance portfolio. The reason for not reporting the results for the multi-prior robust portfolio described in Garlappi, Uppal, and Wang (2006) is that they show that the optimal robust portfolio is a weighted average of the mean-variance and minimum-variance portfolios, both of which are already being reported.

#	Model	Abbreviation
Naive		
0.	$1/N$ with rebalancing *(benchmark strategy)*	ew or $1/N$
Classical approach that ignores estimation error		
1.	Sample-based mean-variance	mv
Bayesian approach to estimation error		
2.	Bayesian diffuse-prior	Not reported
3.	Bayes-Stein	bs
4.	Bayesian Data-and-Model	dm
Moment restrictions		
5.	Minimum-variance	min
6.	Value-weighted market portfolio	vw
7.	MacKinlay and Pastor's (2000) missing-factor model	mp
Portfolio constraints		
8.	Sample-based mean-variance with shortsale constraints	mv-c
9.	Bayes-Stein with shortsale constraints	bs-c
10.	Minimum-variance with shortsale constraints	min-c
11.	Minimum-variance with generalized constraints	g-min-c
Optimal combinations of portfolios		
12.	Kan and Zhou's (2005) "three-fund" model	mv-min
13.	Mixture of minimum-variance and $1/N$	ew-min
14.	Garlappi, Uppal, and Wang's (2006) multi-prior model	Not reported

Table 2: List of datasets considered

This table lists the various datasets analyzed, the number of risky assets N in each dataset (where the number after the "+" indicates the number of factor portfolios available), and the time period spanned. Each dataset contains monthly excess returns over the 90-day nominal US T-bill (from Ken French's website). In parenthesis is the abbreviation used to refer to the dataset in the tables evaluating the performance of the various portfolio strategies. Note that as in Wang (2005), of the 25 size- and book-to-market-sorted portfolios we exclude the 5 portfolio containing the largest firms, because the market, SMB and HML are almost a linear combination of the 25 Fama-French portfolios. Note also that in Datasets #5, 6, and 7, the only difference is in the factor portfolios that are available: in Dataset#5, it is the U.S. equity market portfolio; in Dataset#6, they are the U.S. equity market, SMB, and HML portfolios; in Dataset#7, they are the U.S. equity market, SMB, HML, and MOM portfolios. Because the results for the "FF-3-factor" dataset are virtually identical to those for "FF-1-factor," only the results for "FF-1-factor" are reported.

#	Dataset (Abbreviation)	N	Time Period	Source
1	Ten sector portfolios of the S&P500 and the U.S. equity market portfolio (S&PSectors)	10+1	01/1981–12/2002	Roberto Wessels
2	Ten industry portfolios and the U.S. equity market portfolio (Industry)	10+1	07/1963–11/2004	Ken French's website
3	Eight country indexes and the World Index (International)	8+1	01/1970–07/2001	MSCI
4	SMB and HML portfolios and the U.S. equity market portfolio (MKT/SMB/HML)	2+1	07/1963–11/2004	Ken French's website
5	Twenty Size and Book-to-Market portfolios and the U.S. equity market portfolio (FF-1-factor)	20+1	07/1963–11/2004	Ken French's website
6	Twenty Size and Book-to-Market portfolios and the MKT, SMB, and HML portfolios (FF-3-factor)	20+3	07/1963–11/2004	Ken French's website
7	Twenty Size and Book-to-Market portfolios and the MKT, SMB, HML & MOM portfolios (FF-4-factor)	20+4	07/1963–11/2004	Ken French's website
8	Simulated data with $N = \{10, 25, 50\}$		2000 years	Market Model

Table 3: Sharpe Ratios for Empirical Data

For each of the empirical datasets listed in Table 2, this table reports the monthly Sharpe ratio for the $1/N$ strategy, the in-sample Sharpe ratio of the mean-variance strategy, and the out-of-sample Sharpe ratios for the strategies from the models of optimal asset allocation listed in Table 1. In parenthesis is the P-value of the difference between the Sharpe ratio of each strategy from that of the $1/N$ benchmark, which is computed using the Jobson and Korkie (1981) methodology described in Section 3. The results for the FF-3-factor dataset are not reported because these are very similar to those for the FF-1-factor dataset.

Strategy	S&P Sectors $N = 11$	Industry Portf. $N = 11$	Inter'l Portf. $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$
$1/N$	0.1876	0.1353	0.1277	0.2240	0.1623	0.1753
mv (in sample)	0.3848	0.2124	0.2090	0.2851	0.5098	0.5364
mv	0.0794	-0.0363	-0.0719	0.2186	-0.0684	-0.0031
	(0.12)	(0.01)	(0.01)	(0.46)	(0.00)	(0.01)
bs	0.0811	-0.0319	-0.0528	0.2536	-0.0636	-0.0042
	(0.09)	(0.01)	(0.01)	(0.25)	(0.00)	(0.01)
dm ($\omega = 0.50$)	0.1036	0.0188	0.0811	0.1998	-0.0615	-0.0456
	(0.12)	(0.06)	(0.26)	(0.32)	(0.00)	(0.00)
min	0.0820	0.1554	0.1490	0.2493	0.2778	-0.0183
	(0.05)	(0.30)	(0.21)	(0.23)	(0.01)	(0.01)
vw	0.1444	0.1138	0.1239	0.1138	0.1138	0.1138
	(0.09)	(0.01)	(0.43)	(0.00)	(0.01)	(0.00)
mp	0.1863	0.1249	0.1209	0.0558	0.1525	0.1516
	(0.44)	(0.02)	(0.10)	(0.00)	(0.00)	(0.00)
mv-c	0.0892	0.0678	0.0848	0.1084	0.1977	0.2024
	(0.09)	(0.03)	(0.17)	(0.02)	(0.02)	(0.27)
bs-c	0.1075	0.0819	0.0848	0.1514	0.1955	0.2062
	(0.14)	(0.06)	(0.15)	(0.09)	(0.03)	(0.25)
min-c	0.0834	0.1425	0.1501	0.2493	0.1546	0.3580
	(0.01)	(0.41)	(0.16)	(0.23)	(0.35)	(0.00)
g-min-c	0.1371	0.1451	0.1429	0.2467	0.1615	0.3028
	(0.08)	(0.31)	(0.19)	(0.25)	(0.47)	(0.00)
mv-min	0.0683	-0.0267	0.0572	0.2546	-0.0141	-0.0051
	(0.05)	(0.02)	(0.13)	(0.22)	(0.01)	(0.01)
ew-min	0.1208	0.1576	0.1407	0.2503	0.2608	-0.0161
	(0.07)	(0.21)	(0.18)	(0.17)	(0.00)	(0.01)

Table 4: Certainty-Equivalent Returns for Empirical Data

For each of the empirical datasets listed in Table 2, this table reports the monthly certainty-equivalent return for the $1/N$ strategy, the in-sample certainty-equivalent return of the mean-variance strategy, and the out-of-sample certainty-equivalent returns for the strategies from the models of optimal asset allocation listed in Table 1. In parenthesis is the P-value of the difference between the Sharpe ratio of each strategy from that of the $1/N$ benchmark, which is computed using the Jobson and Korkie (1981) methodology described in Section 3. The results for the FF-3-factor dataset are not reported because these are very similar to those for the FF-1-factor dataset.

Strategy	S&P Sectors $N = 11$	Industry Portf. $N = 11$	Inter'l Portf. $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$
$1/N$	0.0069	0.0050	0.0046	0.0039	0.0073	0.0072
mv (in sample)	0.0478	0.0106	0.0096	0.0047	0.0300	0.0304
mv	0.0031	-0.9222	-0.1549	0.0045	-2.8916	-0.2078
	(0.28)	(0.00)	(0.00)	(0.31)	(0.00)	(0.00)
bs	0.0030	-0.4076	-0.0362	0.0043	-0.7370	-0.1325
	(0.16)	(0.00)	(0.00)	(0.32)	(0.00)	(0.00)
dm ($\omega = 0.50$)	0.0043	-5.6991	0.0021	0.0044	-29.3230	-16.9880
	(0.24)	(0.00)	(0.37)	(0.35)	(0.00)	(0.00)
min	0.0024	0.0052	0.0054	0.0039	0.0100	-0.0002
	(0.03)	(0.45)	(0.23)	(0.45)	(0.12)	(0.00)
vw	0.0053	0.0042	0.0044	0.0042	0.0042	0.0042
	(0.12)	(0.04)	(0.39)	(0.44)	(0.00)	(0.00)
mp	0.0073	0.0047	0.0044	0.0014	0.0069	0.0069
	(0.19)	(0.14)	(0.20)	(0.22)	(0.05)	(0.29)
mv-c	0.0040	0.0023	0.0032	0.0030	0.0090	0.0075
	(0.29)	(0.10)	(0.29)	(0.28)	(0.03)	(0.42)
bs-c	0.0052	0.0031	0.0031	0.0038	0.0088	0.0074
	(0.36)	(0.15)	(0.23)	(0.46)	(0.05)	(0.44)
min-c	0.0024	0.0047	0.0054	0.0039	0.0060	0.0051
	(0.01)	(0.40)	(0.21)	(0.45)	(0.12)	(0.17)
g-min-c	0.0044	0.0048	0.0051	0.0038	0.0067	0.0070
	(0.04)	(0.41)	(0.28)	(0.40)	(0.17)	(0.45)
mv-min	0.0021	-0.3149	0.0013	0.0044	-0.0900	-0.1227
	(0.07)	(0.00)	(0.23)	(0.28)	(0.00)	(0.00)
ew-min	0.0037	0.0052	0.0050	0.0039	0.0093	-0.0002
	(0.04)	(0.42)	(0.24)	(0.43)	(0.12)	(0.00)

Table 5: Portfolio Turnovers for Empirical Data

For each of the empirical datasets listed in Table 2, Panel A of this table reports the monthly turnover for the $1/N$ strategy and the turnovers for the strategies from the models of optimal asset allocation listed in Table 1. Panel B reports the turnover for the strategies from each optimizing model *relative* to the turnover of the $1/N$ model. The results for the FF-3-factor dataset are not reported because these are very similar to those for the FF-1-factor dataset.

Strategy	S&P Sectors $N = 11$	Industry Portf. $N = 11$	Inter'l Portf. $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$
Panel A: Turnover of each strategy						
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
mv (in sample)	–	–	–	–	–	–
mv	1.1894	13132.0286	124.1466	0.0672	152.4460	52.9029
bs	0.6838	218.1655	51.5994	0.0438	170.3013	55.0974
dm ($\omega = 0.50$)	0.6219	756.4779	40.8928	0.0715	249.2670	147.1048
min	0.1996	0.4680	0.2140	0.0263	0.7368	0.1352
vw	0	0	0	0	0	0
mp	0.0336	0.0227	0.0323	0.1046	0.0177	0.0194
mv-c	0.1381	0.1550	0.2120	0.0976	0.2841	0.2735
bs-c	0.1109	0.1562	0.1789	0.0864	0.2806	0.2586
min-c	0.0753	0.0557	0.0665	0.0263	0.0636	0.0348
g-min-c	0.0396	0.0328	0.0432	0.0258	0.0288	0.0336
mv-min	0.6046	214.9602	24.0112	0.0618	60.9671	84.8745
ew-min	0.1469	0.3386	0.1242	0.0264	0.5525	0.1347
Panel B: Turnover relative to $1/N$ strategy						
$1/N$	1	1	1	1	1	1
mv (in sample)	–	–	–	–	–	–
mv	38.99	607479.61	4236.10	2.83	9408.16	2672.71
bs	22.41	10092.20	1760.67	1.85	10510.09	2783.58
dm ($\omega = 0.50$)	20.39	34994.20	1395.34	3.02	15383.43	7431.89
min	6.54	21.65	7.30	1.11	45.47	6.83
vw	0	0	0	0	0	0
mp	1.10	1.05	1.10	4.41	1.09	0.98
mv-c	4.53	7.17	7.23	4.12	17.53	13.82
bs-c	3.64	7.22	6.10	3.65	17.32	13.07
min-c	2.47	2.58	2.27	1.11	3.93	1.76
g-min-c	1.30	1.52	1.47	1.09	1.78	1.70
mv-min	19.82	9943.93	819.30	2.61	3762.56	4287.95
ew-min	4.82	15.66	4.24	1.11	34.10	6.80

Table 6: Ranking of Strategies by Sharpe Ratio, CEQ and Turnover

This table reports the rank of the fourteen strategies for each of the six datasets, based on the strategy's Sharpe ratio, its certainty-equivalent return (CEQ), and its turnover. The ranks are then summed across all the datasets to determine the aggregate ranking for each performance metric. Note that no adjustment is made for the magnitude and statistical significance of the difference in performance.

Strategy	S&P Sectors $N = 11$	Industry Portf. $N = 11$	Inter'l Portf. $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$	Total of ranks	Final Rank
Panel A: Rank based on Sharpe ratio								
$1/N$	2	6	6	8	6	6	34	5
mv (in sample)	1	1	1	1	1	1	6	1
mv	13	14	14	9	14	9	73	14
bs	12	13	13	3	13	10	64	12
dm ($\omega = 0.50$)	8	11	11	10	12	14	66	13
min	11	3	3	5	2	13	37	6
vw	4	8	7	12	10	8	49	9
mp	3	7	8	14	9	7	48	8
mv-c	9	10	9	13	4	5	50	10
bs-c	7	9	10	11	5	4	46	7
min-c	10	5	2	6	8	2	33	4
g-min-c	5	4	4	7	7	3	30	2
mv-min	14	12	12	2	11	11	62	11
ew-min	6	2	5	4	3	12	32	3
Panel B: Rank based on CEQ								
$1/N$	3	4	6	7	6	4	30	2
mv (in sample)	1	1	1	1	1	1	6	1
mv	10	13	14	2	13	13	65	14
bs	11	12	13	5	12	12	65	13
dm ($\omega = 0.50$)	7	14	11	3	14	14	63	12
min	13	3	2	9	2	10	39	4
vw	4	8	8	6	10	8	44	8
mp	2	6	7	14	7	6	42	6
mv-c	8	10	9	13	4	2	46	10
bs-c	5	9	10	11	5	3	43	7
min-c	12	7	3	8	9	7	46	9
g-min-c	6	5	4	12	8	5	40	5
mv-min	14	11	12	4	11	11	63	11
ew-min	9	2	5	10	3	9	38	3
Panel C: Rank based on Turnover								
$1/N$	2	2	2	2	2	3	13	2
mv	13	13	13	9	11	10	69	12
bs	12	11	12	7	12	11	65	11
dm ($\omega = 0.50$)	11	12	11	10	13	13	70	13
min	9	9	9	5	9	7	48	8
vw	1	1	1	1	1	1	6	1
mp	3	3	3	13	3	2	27	4
mv-c	7	6	8	12	7	9	49	9
bs-c	6	7	7	11	6	8	45	7
min-c	5	5	5	4	5	5	29	5
g-min-c	4	4	4	3	4	4	23	3
mv-min	10	10	10	8	10	12	60	10
ew-min	8	8	6	6	8	6	42	6

Table 7: Sharpe Ratios for Simulated Data with 20% Idiosyncratic Volatility

This table reports the monthly Sharpe ratio for the $1/N$ strategy, the in-sample Sharpe ratio of the mean-variance strategy, and the out-of-sample Sharpe ratios for the strategies from the models of optimal asset allocation listed in Table 1. In parenthesis is the P-value of the difference between the Sharpe ratio of each strategy from that of the $1/N$ benchmark, which is computed using the Jobson and Korkie (1981) methodology described in Section 3. These quantities are computed for simulated data that is described in Section 6.1 assuming that idiosyncratic volatility is 20%, for different number of investable assets, N, and different lengths of the estimation window, M, measured in months.

	N=10			N=25			N=50		
Strategy	M=120	M=360	M=6000	M=120	M=360	M=6000	M=120	M=360	M=6000
$1/N$	0.1356	0.1356	0.1356	0.1447	0.1447	0.1447	0.1466	0.1466	0.1466
mv (true)	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477
	(0.00)	(0.00)	(0.00)	(0.03)	(0.03)	(0.03)	(0.15)	(0.15)	(0.15)
mv	-0.0032	-0.0039	0.1416	0.0007	-0.0084	0.1353	0.0040	0.0123	0.1212
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
bs	-0.0024	-0.0038	0.1416	0.0010	-0.0082	0.1363	0.0045	0.0125	0.1229
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dm ($\omega = 0.5$)	-0.0045	0.1205	0.1456	-0.0085	-0.0038	0.1441	0.0039	0.0013	0.1394
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.38)	(0.00)	(0.00)	(0.00)
min	0.1113	0.1181	0.1208	0.0804	0.0911	0.0956	0.0491	0.0676	0.0696
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
mp	0.1356	0.1353	0.1354	0.1449	0.1446	0.1446	0.1467	0.1466	0.1465
	(0.49)	(0.38)	(0.40)	(0.33)	(0.45)	(0.43)	(0.39)	(0.47)	(0.42)
mv-c	0.0970	0.1121	0.1276	0.1011	0.1150	0.1315	0.1111	0.1194	0.1355
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
bs-c	0.1039	0.1221	0.1317	0.1095	0.1222	0.1350	0.1162	0.1251	0.1381
	(0.00)	(0.00)	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
min-c	0.1284	0.1324	0.1335	0.1181	0.1227	0.1248	0.1224	0.1277	0.1292
	(0.00)	(0.08)	(0.17)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
g-min-c	0.1289	0.1312	0.1320	0.1311	0.1336	0.1348	0.1364	0.1402	0.1415
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
mv-min	0.0037	-0.0053	0.1414	0.0028	-0.0050	0.1361	0.0103	0.0098	0.1229
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ew-min	0.1116	0.1184	0.1211	0.0810	0.0918	0.0964	0.0496	0.0684	0.0706
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

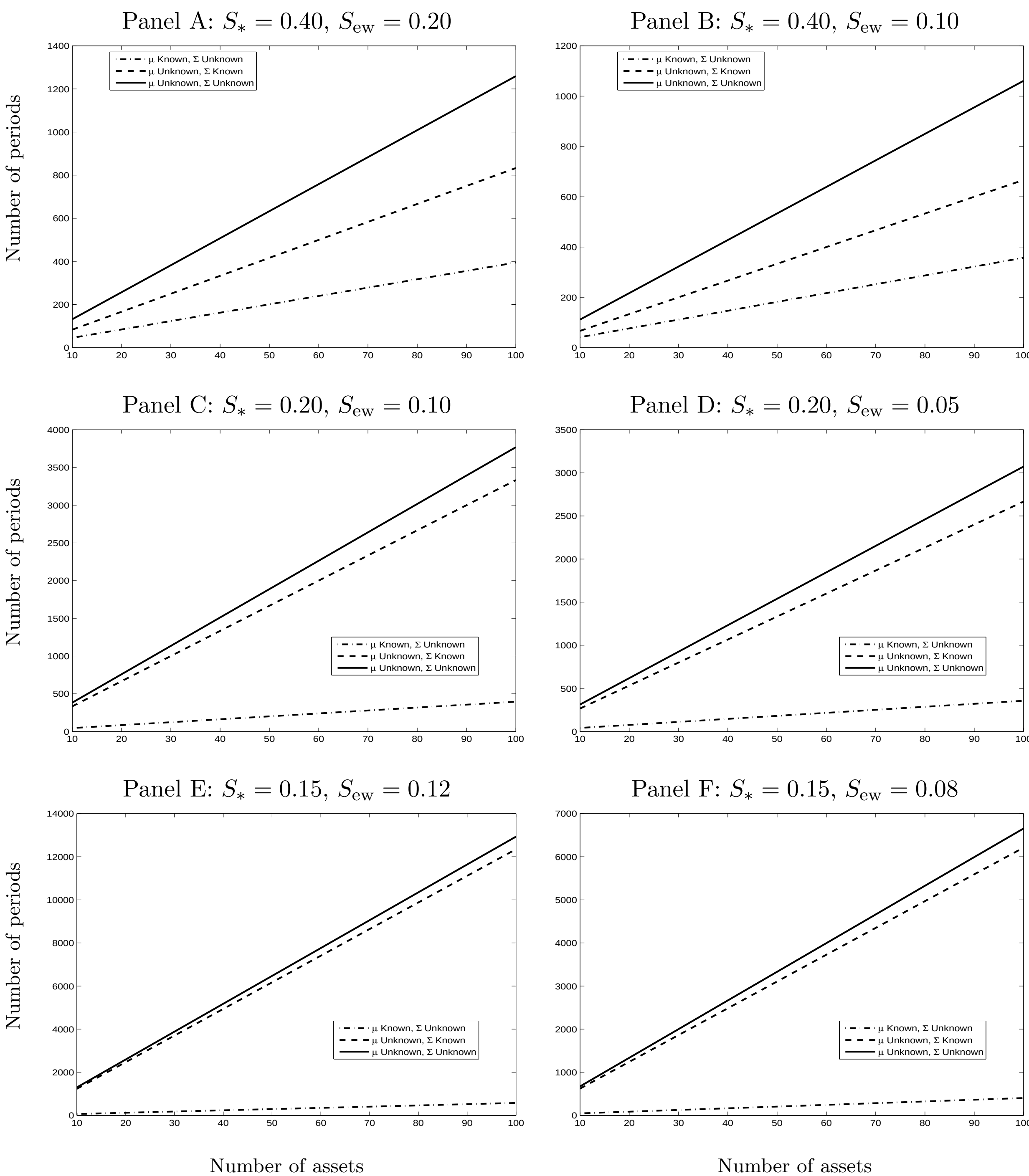
Table 8: Sharpe Ratios for Simulated Data with 75% Idiosyncratic Volatility

This table reports the monthly Sharpe ratio for the $1/N$ strategy, the in-sample Sharpe ratio of the mean-variance strategy, and the out-of-sample Sharpe ratios for the strategies from the models of optimal asset allocation listed in Table 1. In parenthesis is the P-value of the difference between the Sharpe ratio of each strategy from that of the $1/N$ benchmark, which is computed using the Jobson and Korkie (1981) methodology described in Section 3. These quantities are computed for simulated data that is described in Section 6.1 assuming that idiosyncratic volatility is 75%, for different number of investable assets, N, and different lengths of the estimation window, M, measured in months.

	N=10			N=25			N=50		
	M=120	M=360	M=6000	M=120	M=360	M=6000	M=120	M=360	M=6000
Strategy									
$1/N$	0.0842	0.0842	0.0842	0.1139	0.1139	0.1139	0.1269	0.1269	0.1269
mv (true)	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477	0.1477
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
mv	0.0063	0.0857	0.1420	-0.0049	-0.0034	0.1360	0.0039	0.0048	0.1244
	(0.00)	(0.43)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.30)
bs	0.0091	0.1202	0.1444	-0.0037	-0.0021	0.1416	0.0039	0.0057	0.1359
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
dm ($\omega = 0.5$)	-0.0003	0.1199	0.1457	0.0009	0.0221	0.1443	0.0083	0.0434	0.1402
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
min	0.1322	0.1406	0.1444	0.1186	0.1350	0.1411	0.0992	0.1279	0.1376
	(0.00)	(0.00)	(0.00)	(0.20)	(0.00)	(0.00)	(0.00)	(0.42)	(0.01)
mp	-0.0047	0.0090	0.0765	0.1008	0.1111	0.1129	0.1252	0.1270	0.1270
	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)	(0.26)	(0.19)	(0.48)	(0.49)
mv-c	0.0487	0.0700	0.1268	0.0549	0.0702	0.1071	0.0651	0.0737	0.1012
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.00)	(0.00)	(0.00)
bs-c	0.0676	0.1003	0.1401	0.0683	0.0864	0.1287	0.0797	0.0873	0.1184
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
min-c	0.1404	0.1446	0.1459	0.1327	0.1408	0.1446	0.1211	0.1365	0.1429
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)
g-min-c	0.1194	0.1200	0.1200	0.1285	0.1351	0.1371	0.1198	0.1325	0.1380
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
mv-min	0.0050	0.1177	0.1433	0.0023	0.0022	0.1414	-0.0042	0.0133	0.1386
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ew-min	0.1322	0.1406	0.1444	0.1189	0.1353	0.1413	0.0999	0.1287	0.1383
	(0.00)	(0.00)	(0.00)	(0.19)	(0.00)	(0.00)	(0.00)	(0.36)	(0.00)

Figure 1: Number of estimation periods for mean-variance rule to outperform $1/N$

The figure shows the critical number of estimation periods required for the sample-based mean-variance strategy to outperform the $1/N$ rule on average, as a function of the number of assets N. Critical values are computed using the definition in (18). The dashed-dotted line reports the critical value of the estimation window for the case in which the means are known, but the covariances are not. The dashed line refers to the case in which the covariances are known, but the means are not. And, the solid line refers to the case in which both the means and covariances are not known.



References

Barry, C. B., 1974, "Portfolio Analysis under Uncertain Means, Variances, and Covariances," *Journal of Finance*, 29, 515–22.

Bawa, V. S., S. Brown, and R. Klein, 1979, *Estimation Risk and Optimal Portfolio Choice*, North Holland, Amsterdam.

Benartzi, S., and R. Thaler, 2001, "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 7998.

Best, M. J., and R. R. Grauer, 1991, "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies*, 4, 315–42.

Best, M. J., and R. R. Grauer, 1992, "Positively Weighted Minimum-Variance Portfolios and the Structure of Asset Expected Returns," *Journal of Financial and Quantitative Analysis*, 27, 513–37.

Bloomfield, T., R. Leftwich, and J. Long, 1977, "Portfolio Strategies and Performance," *Journal of Financial Economics*, 5, 201–218.

Brandt, M. W., 2004, "Portfolio Choice Problems," in Y. Ait-Sahalia, and L. P. Hansen (ed.), *Handbook of Financial Econometrics*, Elsevier, forthcoming.

Brandt, M. W., P. Santa-Clara, and R. Valkanov, 2005, "Parametric Portfolio Policies: Exploiting Characteristics in the Cross Section of Equity Returns," Working paper, UCLA.

Brown, S., 1979, "The Effect of Estimation Risk on Capital Market Equilibrium," *Journal of Financial and Quantitative Analysis*, 14, 215–220.

Campbell, J. Y., Y. L. Chan, and L. M. Viceira, 2003, "A Multivariate Model of Strategic Asset Allocation," *Journal of Financial Economics*, 67, 41–80.

Campbell, J. Y., and L. M. Viceira, 1999, "Consumption and Portfolio Decisions when Expected Returns are Time Varying," *Quarterly Journal of Economics*, 114, 433–495.

Campbell, J. Y., and L. M. Viceira, 2001, "Who Should Buy Long-Term Bonds?," *American Economic Review*, 91, 99–127.

Campbell, J. Y., and L. M. Viceira, 2002, *Strategic Asset Allocation*, Oxford University Press, New York.

Chan, L. K. C., J. Karceski, and J. Lakonishok, 1999, "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model," *Review of Financial Studies*, 12, 937–74.

Chopra, V. K., 1993, "Improving Optimization," *Journal of Investing*, 8, 51–59.

Dumas, B., and B. Jacquillat, 1990, "Performance of Currency Portfolios Chosen by a Bayesian Technique: 1967-1985," *Journal of Banking and Finance*, 14, 539–58.

Frost, P., and J. Savarino, 1988, "For Better Performance Constrain Portfolio Weights," *Journal of Portfolio Management*, 15, 29–34.

Frost, P. A., and J. E. Savarino, 1986, "An Empirical Bayes Approach to Efficient Portfolio Selection," *Journal of Financial and Quantitative Analysis*, 21, 293–305.

Garlappi, L., R. Uppal, and T. Wang, 2006, "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," forthcoming in *Review of Financial Studies*.

Goldfarb, D., and G. Iyengar, 2003, "Robust Portfolio Selection Problems," *Mathematica of Operations Research*, 28, 1–38.

Greene, W. H., 2002, *Econometric Analysis*, Prentice Hall, New York.

Haff, L., 1979, "An Identity for the Wishart Distribution with Applications," *Journal of Multivariate Analysis*, 9, 531–544.

Harvey, C. R., J. Liechty, M. Liechty, and P. Müller, 2003, "Portfolio Selection with Higher Moments," Working Paper, Duke University.

Hodges, S. D., and R. A. Brealey, 1978, "Portfolio Selection in a Dynamic and Uncertin World," in James H. Lorie, and R. A. Brealey (ed.), *Modern Developments in Investment Management*, Dryden Press, Hinsdale, Illionois, 2nd edn.

Huberman, G., and W. Jiang, 2006, "Offering vs. Choice in 401(k) Plans: Equity Exposure and Number of Funds Choice in 401(k) Plans," *Journal of Finance*, 61, 763–801.

Jagannathan, R., and T. Ma, 2003, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1684.

James, W., and C. Stein, 1961, "Estimation with Quadratic Loss," in *Proceedings of the 4th Berkeley Symposium on Probability and Statistics 1*. Berkeley: University of California Press.

Jobson, J. D., and R. Korkie, 1980, "Estimation for Markowitz Efficient Portfolios," *Journal of the American Statistical Association*, 75, 544–554.

Jobson, J. D., and R. Korkie, 1981, "Performance Hypothesis Testing with the Sharpe and Treynor Measures,," *Journal of Finance*, 36, 889–908.

Jorion, P., 1985, "International Portfolio Diversification with Estimation Risk," *Journal of Business*, 58, 259–278.

Jorion, P., 1986, "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, 21, 279–92.

Jorion, P., 1991, "Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection," *Journal of Banking and Finance*, 15, 717–27.

Kan, R., and G. Zhou, 2005, "Optimal Portfolio Choice with Parameter Uncertainty," forthcoming in *Journal of Financial and Quantitative Analysis*.

Klein, R. W., and V. S. Bawa, 1976, "The Effect of Estimation Risk on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–31.

Knight, F., 1921, *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston.

Ledoit, O., 1996, "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices," Working Paper, Anderson School, UCLA.

Ledoit, O., and M. Wolf, 2003, "Honey, I Shrunk the Sample Covariance Matrix," Working Paper, Department of Economics and Business, Universitat Pompeu Fabra.

Litterman, R., 2003, *Modern Investment Management: An Equilibrium Approach*, Wiley.

MacKinlay, A. C., and L. Pastor, 2000, "Asset Pricing Models: Implications for Expected Returns and Portfolio Selection," *Review of Financial Studies*, 13, 883–916.

Markowitz, H. M., 1952, "Mean-Variance Analysis in Portfolio Choice and Capital Markets," *Journal of Finance*, 7, 77–91.

Memmel, C., 2003, "Performance Hypothesis Testing with the Sharpe Ratio," *Finance Letters*, 1, 21–23.

Merton, R. C., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51, 247–257.

Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

Michaud, R. O., 1989, "The Markowitz Optimization Enigma: Is Optimized Optimal," *Financial Analysts Journal*, 45, 31–42.

Michaud, R. O., 1998, *Efficient Asset Management*, Harvard Business School Press, Boston.

Morrison, D. F., 1990, *Multivariate Statistical Methods*, McGraw-Hill, New York.

Pástor, Ľ., 2000, "Portfolio Selection and Asset Pricing Models," *Journal of Finance*, 55, 179–223.

Pástor, Ľ., and R. F. Stambaugh, 2000, "Comparing Asset Pricing Models: An Investment Perspective," *Journal of Financial Economics*, 56, 335–81.

Samuelson, P., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239–246.

Scherer, B., 2002, "Portfolio Resampling: Review and Critique," *Financial Analysts Journal*, 58, 98–109.

Stein, C., 1955, "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in *3rd Berkely Symposium on Probability and Statistics*, , vol. 1, pp. 197–206, Berkeley. University of California Press.

Wang, Z., 2005, "A Shrinkage Approach to Model Uncertainty and Asset Allocation," *The Review of Financial Studies*, 18, 673–705.