Optimal Versus Naive Diversification: How Inefficient is the 1/*N* Portfolio Strategy?

Victor DeMiguel
London Business School

Lorenzo Garlappi
University of Texas at Austin

Raman Uppal
London Business School and CEPR

We evaluate the out-of-sample performance of the sample-based mean-variance model, and its extensions designed to reduce estimation error, relative to the naive $1/N$ portfolio. Of the 14 models we evaluate across seven empirical datasets, none is consistently better than the $1/N$ rule in terms of Sharpe ratio, certainty-equivalent return, or turnover, which indicates that, out of sample, the gain from optimal diversification is more than offset by estimation error. Based on parameters calibrated to the US equity market, our analytical results and simulations show that the estimation window needed for the sample-based mean-variance strategy and its extensions to outperform the $1/N$ benchmark is around 3000 months for a portfolio with 25 assets and about 6000 months for a portfolio with 50 assets. 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. (*JEL* G11)

In about the fourth century, Rabbi Issac bar Aha proposed the following rule for asset allocation: "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand."[1] After a "brief"

We wish to thank Matt Spiegel (the editor), two anonymous referees, and Ľuboš Pástor for extensive comments; 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&P 500 Index. We also gratefully acknowledge the comments from Pierluigi Balduzzi, John Birge, Michael Brennan, Ian Cooper, Bernard Dumas, Bruno Gerard, Francisco Gomes, Eric Jacquier, Chris Malloy, Francisco Nogales, 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, Manchester 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, the 2005 conference of the Institute for Quantitative Investment Research (Inquire UK), the 2006 NBER Summer Institute, the 2006 meeting of the European Finance Association, and the First Annual Meeting of the Swiss Finance Institute. Send correspondence to Lorenzo Garlappi, McCombs School of Business, University of Texas at Austin, Austin, TX 78712; telephone: (512) 471-5682; fax: (512) 471-5073. E-mail: lorenzo.garlappi@mccombs.utexas.edu.

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

© The Author 2007. Published by Oxford University Press on behalf of The Society for Financial Studies. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.
doi:10.1093/rfs/hhm075 Advance Access publication December 3, 2007

lull in the literature on asset allocation, there have been considerable advances starting with the pathbreaking work of Markowitz (1952),[2] 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. 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 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; Bawa, Brown, and Klein, 1979), to "shrinkage estimators" (Jobson, Korkie, and Ratti, 1979; Jobson and Korkie, 1980; Jorion, 1985, 1986), to the more recent approaches that rely on an asset-pricing model for establishing a prior (Pástor, 2000; Pástor and Stambaugh, 2000).[4] Equally rich is the set of *non-Bayesian* approaches to estimation error, which include "robust" portfolio allocation rules (Goldfarb and Iyengar, 2003; Garlappi, Uppal, and Wang, 2007); portfolio rules designed to optimally diversify across market *and* estimation risk (Kan and Zhou, 2007); portfolios that exploit the moment restrictions imposed by the factor structure of returns (MacKinlay and Pastor, 2000); methods that focus on reducing the error in estimating the covariance matrix (Best and Grauer, 1992; Chan, Karceski, and Lakonishok, 1999; Ledoit and Wolf, 2004a, 2004b); and, finally, portfolio rules that impose shortselling constraints (Frost and Savarino, 1988; Chopra, 1993; Jagannathan and Ma, 2003).[5]

Our objective in this paper is to understand the conditions under which mean-variance 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 its various extensions designed to reduce the effect of estimation error—relative to the performance of the *naive* portfolio diversification rule. We define the naive rule to be one in which a fraction $1/N$ of wealth is allocated to each of the N assets available for investment at each rebalancing date. There are two reasons for using the naive rule as a benchmark. First, it is easy to implement because it does not rely either on estimation of the moments of asset returns or on

[2] Some of the results on mean-variance portfolio choice in Markowitz (1952, 1956, 1959) and Roy (1952) had already been anticipated in 1940 by de Finetti, an English translation of which is now available in Barone (2006).

[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 (2007).

[4] Another approach, proposed by Black and Litterman (1990, 1992), combines two sets of priors—one based on an equilibrium asset-pricing model and the other on the subjective views of the investor—which is not strictly Bayesian, because a Bayesian approach combines a prior with the data.

[5] Michaud (1998) has advocated the use of resampling methods; Scherer (2002) and Harvey et al. (2003) discuss the various limitations of this approach.

Table 1
List of various asset-allocation models considered

#	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 (2007) "three-fund" model	mv-min
13.	Mixture of minimum-variance and $1/N$	ew-min
14.	Garlappi, Uppal, and Wang's (2007) multi-prior model	Not reported

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 (2007) is that they show that the optimal robust portfolio is a weighted average of the mean-variance and minimum-variance portfolios, the results for both of which are already being reported.

optimization. Second, despite the sophisticated theoretical models developed in the last 50 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 performance of various portfolio rules proposed in the literature.

We compare the out-of-sample performance of 14 different portfolio models relative to that of the $1/N$ policy across seven empirical datasets of monthly returns, using the following three performance criteria: (i) the out-of-sample Sharpe ratio; (ii) the certainty-equivalent (CEQ) return for the expected utility of a mean-variance investor; and (iii) the turnover (trading volume) for each portfolio strategy. The 14 models are listed in Table 1 and discussed in Section 1. The seven empirical datasets are listed in Table 2 and described in Appendix A.

[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.

Table 2
List of datasets considered

#	Dataset and source	N	Time period	Abbreviation
1	Ten sector portfolios of the S&P 500 and the US equity market portfolio Source: Roberto Wessels	$10+1$	01/1981–12/2002	S&P Sectors
2	Ten industry portfolios and the US equity market portfolio Source: Ken French's Web site	$10+1$	07/1963–11/2004	Industry
3	Eight country indexes and the World Index Source: MSCI	$8+1$	01/1970–07/2001	International
4	SMB and HML portfolios and the US equity market portfolio Source: Ken French's Web site	$2+1$	07/1963–11/2004	MKT/SMB/HML
5	Twenty size- and book-to-market portfolios and the US equity MKT Source: Ken French's Web site	$20+1$	07/1963–11/2004	FF-1-factor
6	Twenty size- and book-to-market portfolios and the MKT, SMB, and HML portfolios Source: Ken French's Web site	$20+3$	07/1963–11/2004	FF-3-factor
7	Twenty size- and book-to-market portfolios and the MKT, SMB, HML, and UMD portfolios Source: Ken French's Web site	$20+4$	07/1963–11/2004	FF-4-factor
8	Simulated data Source: Market model	$\{10, 25, 50\}$	2000 years	—

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 Web site). In the last column 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 five portfolios 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 US equity MKT; in Dataset #6, they are the MKT, SMB, and HML portfolios; and in Dataset #7, they are the MKT, SMB, HML, and UMD portfolios. Because the results for the "FF-3-factor" dataset are almost identical to those for "FF-1-factor," only the results for "FF-1-factor" are reported.

Our first contribution is to show that of the 14 models evaluated, 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 earlier models,[7] we demonstrate that this is true: (i) for a wide range of models that include several developed more recently; (ii) using three performance metrics; and (iii) across several datasets. In general, 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* on the sample-based mean-variance and Bayesian portfolio strategies leads to only a modest improvement in Sharpe ratios and CEQ returns, although it shows a substantial reduction in turnover. Of all the optimizing models studied here, the minimum-variance portfolio with constraints studied in Jagannathan and Ma

[7] 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.

(2003) performs best in terms of Sharpe ratio. But even this model delivers a Sharpe ratio that is statistically superior to that of the $1/N$ strategy in only one of the seven empirical datasets, a CEQ return that is not statistically superior to that of the $1/N$ strategy in any of these datasets, and a turnover that is always 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 *analytical* expression for the *critical length* of the estimation window that is needed for the sample-based mean-variance strategy to achieve a higher CEQ return than that of 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 US stock-market data, we find that the critical length of the estimation window is 3000 months for a portfolio with only 25 assets, and more than 6000 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.

Because the above analytical results are available only for the sample-based mean-variance strategy, we use simulated data to examine its various extensions 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. 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. Moreover, other things being equal, a smaller number of assets makes naive diversification less effective relative to optimal diversification.

The intuition for our findings is that to implement the mean-variance model, both the vector of expected excess returns over the risk-free rate and the variance-covariance matrix of returns have to be estimated. It is well known (Merton, 1980) that a very long time series of data is required in order to estimate expected returns precisely; similarly, the estimate of the variance-covariance matrix is poorly behaved (Green and Hollifield, 1992; Jagannathan and Ma, 2003). The portfolio weights based on the sample estimates of these moments result in extreme positive and negative weights that are far from optimal.[8] As

[8] Consider the following extreme two-asset example. Suppose that the true per annum mean and volatility of returns 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 for the two assets would be 50%. If, on the other hand, the mean return on the first asset is 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,

a result, "allocation mistakes" caused by using the $1/N$ weights can turn out to be *smaller* than the error caused by using the weights from an optimizing model with inputs that have been estimated with error. Although the "error-maximizing" property of the mean-variance portfolio has been described in the literature (Michaud, 1989; 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 designed explicitly to reduce the effect of estimation error achieve only modest success.

A second reason why the $1/N$ rule performs well in the datasets we consider is that we are using it to allocate wealth across portfolios of stocks rather than individual stocks. Because diversified portfolios have lower idiosyncratic volatility than individual assets, the loss from naive as opposed to optimal diversification is much smaller when allocating wealth across portfolios. Our simulations show that optimal diversification policies will dominate the $1/N$ rule only for very high levels of idiosyncratic volatility. Another advantage of the $1/N$ rule is that it is straightforward to apply to a large number of assets, in contrast to optimizing models, which typically require additional parameters to be estimated as the number of assets increases.

In all our experiments, the choice of N has been dictated by the dataset. A natural question that arises then is: What is N? That is, for what number and kind of assets does the $1/N$ strategy outperform other optimizing portfolio models? The results show that the naive $1/N$ strategy is more likely to outperform the strategies from the optimizing models when: (i) N is large, because this improves the potential for diversification, even if it is naive, while at the same time increasing the number of parameters to be estimated by an optimizing model; (ii) the assets do not have a sufficiently long data history to allow for a precise estimation of the moments. In the empirical analysis, we consider datasets with $N = \{3, 9, 11, 21, 24\}$ and assets from equity portfolios that are based on industry classification, equity portfolios constructed on the basis of firm characteristics, and also international equity indices. In the simulations, $N = \{10, 25, 50\}$ and the asset returns are calibrated to match returns on portfolios of US stocks. The empirical and simulation-based results show that for an estimation window of $M = 120$ months, our main finding is not sensitive to the type of assets we considered or to the choice of the number of assets, N.

We draw two conclusions from the results. 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 expected returns. For this, methods that complement traditional classical and Bayesian statistical techniques by exploiting empirical regularities that are present for a particular set of assets (Brandt, Santa-Clara,

---

the *optimization* tries to exploit even the smallest difference in the two assets by taking extreme long and short positions *without* taking into account that these differences in returns may be the result of estimation error. As we describe in Section 3, the weights from mean-variance optimization when using actual data and more than just two assets are even more extreme than the weights in the given example.

and Valkanov, 2007) 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 asset-allocation models and for "active" portfolio-management strategies offered by the investment industry.

The rest of the paper is organized as follows. In Section 1, we describe the various models of optimal asset allocation and evaluate their performance. In Section 2, 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 3. Section 4 contains the analytical 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 5 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 the findings are described briefly in Section 6, with the details reported in a separate appendix titled "Implementation Details and Robustness Checks," which is available from the authors. Our conclusions are presented in Section 7. The empirical datasets we use are described in Appendix A and the proof for the main analytical result is given in Appendix B.

1. Description of the Asset-Allocation Models Considered

In this section, we discuss the various models from the portfolio-choice literature that we consider. 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 the separate appendix to this paper.

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 $N \times N$ identity 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. The vector of *relative* weights in the portfolio with only-risky assets is

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

where the normalization by the absolute value of the sum of the portfolio weights, $|\mathbf{1}_N^\top \mathrm{x}_t|$, guarantees that the direction of the portfolio position is preserved in the few cases where the sum of the weights on the risky assets is negative.

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

1.1 Naive portfolio

The naive ("ew" or "1/N") 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 Equation (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.

1.2 Sample-based mean-variance portfolio

In the mean-variance ("mv") model of Markowitz (1952), the investor optimizes the tradeoff between the mean and variance of portfolio returns. To implement this model, we follow the classic "plug-in" approach; that is, we solve the problem in Equation (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 possibility of estimation error.

1.3 Bayesian approach to estimation error

Under the Bayesian approach, the estimates of μ and Σ are computed using the *predictive distribution* of asset returns. This distribution is obtained by

[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.

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. In the following sections, we describe three common implementations we consider.

1.3.1 Bayesian diffuse-prior portfolio. Barry (1974), Klein and Bawa (1976), and Brown (1979) 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.

1.3.2 Bayes-Stein shrinkage portfolio. The Bayes-Stein ("bs") 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

$$\hat{\mu}_t^{\text{bs}} = (1 - \hat{\phi}_t)\,\hat{\mu}_t + \hat{\phi}_t\,\hat{\mu}_t^{\min}, \tag{4}$$

$$\hat{\phi}_t = \frac{N + 2}{(N + 2) + M(\hat{\mu}_t - \mu_t^{\min})^{\top}\hat{\Sigma}_t^{-1}(\hat{\mu}_t - \mu_t^{\min})}, \tag{5}$$

in which $0 < \hat{\phi}_t < 1$, $\hat{\Sigma}_t = \frac{1}{M-N-2}\sum_{s=t-M+1}^{t}(R_s - \hat{\mu}_t)(R_s - \hat{\mu}_t)^{\top}$, and $\hat{\mu}_t^{\min} \equiv \hat{\mu}_t^{\top}\hat{w}_t^{\min}$ is the average excess return on the sample global minimum-variance portfolio, $\hat{w}_t^{\min}$. 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]

1.3.3 Bayesian portfolio based on belief in an asset-pricing model. 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 belief relative to the information contained in the data. These portfolios are a further refinement of shrinkage

[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.

portfolios because they address the arbitrariness of the choice of a shrinkage target, $\bar{\mu}$, and of the shrinkage factor, ϕ, by using the investor's belief about the validity of an asset-pricing model. We implement the Data-and-Model approach using three different asset-pricing models: the Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model, and the Carhart (1997) four-factor model. In our empirical analysis, we consider a Bayesian investor whose belief in the asset-pricing model is captured by a prior about the extent of mispricing. Let the variable α reflect this mispricing. We assume the prior to be normally distributed around $\alpha = 0$, and with the benchmark value of its tightness being $\sigma_\alpha = 1\%$ per annum. Intuitively, this implies that the investor believes that with 95% probability the mispricing is approximately between -2% and $+2\%$ on an annual basis.

1.4 Portfolios with moment restrictions

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

1.4.1 Minimum-variance portfolio. Under the minimum-variance ("min") strategy, we choose the portfolio of risky assets that minimizes the variance of returns; that is,

$$\min_{\mathrm{w}_t} \ \mathrm{w}_t^\top \Sigma_t \mathrm{w}_t, \quad \text{s.t.} \quad \mathbf{1}_N^\top \mathrm{w}_t = 1. \tag{6}$$

To implement this policy, we use only the estimate of the covariance matrix of asset returns (the sample covariance matrix) and completely ignore the estimates of the expected returns.[11] 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 Equation (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, $\mu_t \propto \mathbf{1}_N$.

1.4.2 Value-weighted portfolio implied by the market model. The optimal strategy in a CAPM world is the value-weighted ("vw") 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.

1.4.3 Portfolio implied by asset-pricing models with unobservable factors. MacKinlay and Pastor (2000) show that if returns have an exact factor structure but some factors are not observed, then the resulting mispricing is contained

[11] Note that expected returns, μ_t, *do* appear in the likelihood function needed to estimate Σ_t. However, under the assumption of normally distributed asset returns, it is possible to show (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 for constructing the MLE estimator of Σ_t.

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 Pastor show that, in this case, the covariance matrix of returns takes the following form[12]:

$$\Sigma = \nu\mu\mu^{\top} + \sigma^2 I_N, \tag{7}$$

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). We denote this portfolio strategy by "mp."

1.5 Shortsale-constrained portfolios

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 nonnegativity constraint on the portfolio weights in the corresponding optimization problems.

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

$$\mathcal{L} = x_t^{\top}\mu_t - \frac{\gamma}{2}x_t^{\top}\Sigma_t x_t + x_t^{\top}\lambda_t, \tag{8}$$

in which λ_t is the $N \times 1$ vector of Lagrange multipliers for the constraints on shortselling. Rearranging Equation (8), 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 its expected return is low. 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 the average.

Similarly, Jagannathan and Ma (2003) show that imposing a shortsale constraint on the minimum-variance portfolio is equivalent 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); Chan, Karceski,

[12] 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.

and Lakonishok (1999); and Ledoit and Wolf (2004a, 2004b)—that have been developed to deal with the problems associated with estimating the covariance matrix.[13]

Motivated by the desire to examine whether the out-of-sample portfolio performance can be improved by ignoring expected returns (which are difficult to estimate) but still taking into account the correlations between returns, we also consider a new strategy that has not been considered in the existing literature. This strategy, denoted by "g-min-c," is a combination of the $1/N$ policy and the constrained-minimum-variance strategy, and it can be interpreted as a simple generalization of the shortsale-constrained minimum-variance portfolio. It is obtained by imposing an additional constraint on the minimum-variance problem (6): $\mathrm{w} \geq a\mathbf{1}_N$, 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.

1.6 Optimal combination of portfolios

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 (nonnormalized) portfolios of the form

$$\mathrm{x}^{\mathrm{S}} = c\,\mathrm{x}^{c} + d\,\mathrm{x}^{d}, \quad \text{s.t.} \quad \mathbf{1}_N^{\top}\mathrm{x}^{\mathrm{S}} = 1, \tag{9}$$

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. The two mixture portfolios that we consider are described as follows.

1.6.1 The Kan and Zhou (2007) three-fund portfolio. In order to improve on the models that use Bayes-Stein shrinkage estimators, Kan and Zhou (2007) propose a "three-fund" ("mv-min") 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 also benefit from 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 nonnormalized

[13] 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.

weights of this mixture portfolio are

$$\hat{\mathrm{x}}_t^{\text{mv-min}} = \frac{1}{\gamma}\big(c\,\hat{\Sigma}_t^{-1}\hat{\mu}_t + d\,\hat{\Sigma}_t^{-1}\mathbf{1}_N\big), \tag{10}$$

in which c and d are chosen optimally to maximize the expected utility of a mean-variance investor. The weights in the risky assets used in our implementation are given by normalizing the expression in Equation (10); that is, $\hat{\mathrm{w}}_t^{\text{mv-min}} = \hat{\mathrm{x}}_t^{\text{mv-min}}/|\mathbf{1}_N^\top \hat{\mathrm{x}}_t^{\text{mv-min}}|$.

1.6.2 Mixture of equally weighted and minimum-variance portfolios. Finally, we consider a new portfolio strategy denoted "ew-min" that has not been studied in the existing literature. This strategy is a combination of the naive $1/N$ portfolio and the minimum-variance portfolio, rather than the mean-variance portfolio and the minimum-variance portfolio considered in Kan and Zhou (2007) and Garlappi, Uppal, and Wang (2007).[14] Again, our motivation for considering this portfolio is that because expected returns are more difficult to estimate than covariances, one may want to ignore the estimates of mean returns but 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

$$\hat{\mathrm{w}}^{\text{ew-min}} = c\,\frac{1}{N}\mathbf{1}_N + d\,\hat{\Sigma}^{-1}\mathbf{1}_N, \quad \text{s.t. } \mathbf{1}_N^\top \hat{\mathrm{w}}^{\text{ew-min}} = 1, \tag{11}$$

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

2. Methodology for Evaluating Performance

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

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 $M = 120$ months.[15] In each month t, starting from $t = M + 1$, we use the data in the previous M months to estimate the parameters needed to

[14] Garlappi, Uppal, and Wang (2007) consider an investor who is averse not just to risk but also to uncertainty, in the sense of Knight (1921). They 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 aversion to uncertainty. 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 extreme portfolios, we do not report separately the performance of robust portfolio strategies.

[15] The insights from the results for the case of $M = 60$ are not very different from those for the case of $M = 120$, and hence, in the interest of conserving space, are reported only in the separate appendix.

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 $t + 1$. 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. 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 three 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, $\hat{\sigma}_k$:

$$\widehat{\mathrm{SR}}_k = \frac{\hat{\mu}_k}{\hat{\sigma}_k}. \tag{12}$$

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).[16]

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

$$\widehat{\mathrm{SR}}_k^{\mathrm{IS}} = \frac{\mathrm{Mean}_k}{\mathrm{Std}_k} = \frac{\hat{\mu}_k^{\mathrm{IS}\top} \hat{\mathrm{w}}_k}{\sqrt{\hat{\mathrm{w}}_k^{\top} \hat{\Sigma}_k^{\mathrm{IS}} \hat{\mathrm{w}}_k}}, \tag{13}$$

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

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

$$\widehat{\mathrm{CEQ}}_k = \hat{\mu}_k - \frac{\gamma}{2}\hat{\sigma}_k^2, \tag{14}$$

[16] 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:

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

Note that this statistic holds asymptotically under the assumption that returns are distributed independently and identically (IID) over time with a normal distribution. This assumption is typically violated in the data. We address this in Section 5, where we simulate a dataset with $T = 24{,}000$ monthly returns that are IID normal.

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 the risk aversion.[17] The results we report are for the case of $\gamma = 1$; results for other values of γ are discussed in the separate appendix with robustness checks. 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 for means and variance.[18]

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:

$$\text{Turnover} = \frac{1}{T-M} \sum_{t=1}^{T-M} \sum_{j=1}^{N} \left(\left| \hat{w}_{k,j,t+1} - \hat{w}_{k,j,t^+} \right| \right), \tag{15}$$

in which $\hat{w}_{k,j,t}$ is the portfolio weight in asset j at time t under strategy k; $\hat{w}_{k_j,t^+}$ is the portfolio weight *before* rebalancing at $t+1$; and $\hat{w}_{k,j,t+1}$ is the desired portfolio weight at time $t+1$, after rebalancing. For example, in the case of the $1/N$ strategy, $w_{k,j,t} = w_{k,j,t+1} = 1/N$, 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. For the $1/N$ benchmark strategy we report its absolute turnover, and for all the other strategies their turnover relative to that of the benchmark strategy.

In addition to reporting the raw turnover for each strategy, we also report an economic measure of this by reporting how proportional transactions costs generated by this turnover affect the returns from a particular strategy.[19] We set the proportional transactions cost equal to 50 basis points per transaction as assumed in Balduzzi and Lynch (1999), based on the studies of the cost per transaction for individual stocks on the NYSE by Stoll and Whaley (1983), Bhardwaj and Brooks (1992), and Lesmond, Ogden, and Trzcinka (1999).

Let $R_{k,p}$ be the return from strategy k on the portfolio of N assets before rebalancing; that is, $R_{k,p} = \sum_{j=1}^{N} R_{j,t+1} \hat{w}_{k,j,t}$. When the portfolio is

[17] To be precise, the definition in Equation (14) 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.

[18] If v denotes the vector of moments $v = (\mu_i, \mu_n, \sigma_i^2, \sigma_n^2)$, $\hat{v}$ 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 $f(v)$ (Greene, 2002) is $\sqrt{T}(f(\hat{v}) - f(v)) \to \mathcal{N}(0, \frac{\partial f}{\partial v}^\top \Theta \frac{\partial f}{\partial v})$, in which

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

[19] Note that while the turnover of each strategy is related to the transactions costs incurred in implementing that strategy, it is important to realize that in the presence of transactions costs, it would not be optimal to implement the same portfolio strategy.

rebalanced at time $t+1$, it gives rise to a trade in each asset of magnitude $|\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t^+}|$. Denoting by c the proportional transaction cost, the cost of such a trade over all assets is $c \times \sum_{j=1}^{N} |\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t^+}|$. Therefore, we can write the evolution of wealth for strategy k as follows:

$$W_{k,t+1} = W_{k,t}(1 + R_{k,p})\left(1 - c \times \sum_{j=1}^{N} |\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t^+}|\right), \qquad (16)$$

with the return *net* of transactions costs given by $\frac{W_{k,t+1}}{W_{k,t}} - 1$.

For each strategy, we compute the *return-loss* with respect to the $1/N$ strategy. The return-loss is defined as the additional return needed for strategy k to perform as well as the $1/N$ strategy in terms of the Sharpe ratio. To compute the return-loss per month, suppose μ_{ew} and σ_{ew} are the monthly out-of-sample mean and volatility of the net returns from the $1/N$ strategy, and μ_k and σ_k are the corresponding quantities for strategy k. Then, the return-loss from strategy k is

$$\text{return-loss}_k = \frac{\mu_{ew}}{\sigma_{ew}} \times \sigma_k - \mu_k. \qquad (17)$$

3. 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 $1/N$ strategy. For each strategy, we compute across all the datasets listed in Table 2, the in-sample and out-of-sample Sharpe ratios (Table 3), the CEQ return (Table 4), and the turnover (Table 5). In each of these tables, the various strategies being examined are listed in rows, while the columns refer to the different datasets.

3.1 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.[20] 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&P Sectors"), 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

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

Table 3
Sharpe ratios for empirical data

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	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	0.0679	−0.0332	0.2186	0.0128	0.1841
	(0.12)	(0.17)	(0.03)	(0.46)	(0.02)	(0.45)
bs	0.0811	0.0719	−0.0297	0.2536	0.0138	0.1791
	(0.09)	(0.19)	(0.03)	(0.25)	(0.02)	(0.48)
dm ($\sigma_\alpha = 1.0\%$)	0.1410	0.0581	0.0707	0.0016	0.0004	0.2355
	(0.08)	(0.14)	(0.08)	(0.00)	(0.01)	(0.17)
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	0.0533	0.0984	−0.0002	0.1238	0.1230
	(0.44)	(0.04)	(0.15)	(0.00)	(0.08)	(0.03)
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	0.0772	−0.0353	0.2546	−0.0079	0.1757
	(0.05)	(0.21)	(0.01)	(0.22)	(0.01)	(0.50)
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)

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 parentheses 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 2. The results for the "FF-3-factor" dataset are not reported because they are very similar to those for the "FF-1-factor" dataset.

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.

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. The difference between the mean-variance strategy's in-sample and out-of-sample Sharpe ratios allows us to gauge the severity of the 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 than that for the $1/N$ strategy for all but one of the datasets, with the exception being the "FF-4-factor" dataset (though the difference is statistically insignificant). 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&P Sectors," 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. Similarly, for the "International" dataset, the in-sample Sharpe ratio for the mean-variance strategy is 0.2090, which drops to -0.0332 out of sample, while the Sharpe ratio of the $1/N$ strategy is 0.1277.

The comparisons of Sharpe ratios confirm the well-known perils of using classical sample-based estimates of the moments of asset returns to implement Markowitz's mean-variance portfolios. Thus, our first observation is that *out of sample*, the $1/N$ strategy typically 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. In Table 3, 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 "FF-4-factor," and even in these cases the difference is not statistically significant at conventional levels (the p-values are 0.25 and 0.48, respectively). In fact, the Sharpe ratios for the Bayes-Stein portfolios are only slightly better than that for the sample-based mean-variance portfolio. The reason why in our datasets the Bayes-Stein strategy yields 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.[21] The Data-and-Model strategy, "dm," in which the investor's prior on the mispricing α of the model (CAPM; Fama and French, 1993; or Carhart, 1997) has a tightness of 1% per annum ($\sigma_\alpha = 1.0\%$), improves over the Bayes-Stein approach for three datasets—"S&P Sectors," "International," and "FF-4-factor." However, the "dm" strategy outperforms the $1/N$ strategy only for the "FF-4-factor" dataset, in which the "dm" strategy with $\sigma_\alpha = 1\%$ achieves a Sharpe ratio of 0.2355, which is larger than the Sharpe ratio of 0.1753 for the $1/N$ strategy, but the difference is statistically insignificant (the p-value is 0.17). As we document in the appendix "Implementation Details and Robustness Checks" that is available from the authors, the improved performance of the "dm" strategy for the "FF-4-factor" dataset is because the Carhart (1997) model provides a good description of the cross-sectional returns for the size- and book-to-market portfolios.

Our third observation is about the portfolios that are based on restrictions on the moments of returns. From the row in Table 3 for the minimum-variance strategy titled "min," we see that ignoring the estimates of expected returns

[21] The factor that determines the shrinkage of expected returns toward the mean return on the minimum-variance portfolio is $\hat{\phi}$ [see Equation (4)]. For the datasets we are considering, $\hat{\phi}$ ranges from a low of 0.32 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, $\hat{\mu}$.

altogether but exploiting the information about correlations does lead to better performance, relative to the out-of-sample mean-variance strategy "mv" in all datasets but "FF-4-factor." 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 $+124\%$ rather than ranging from -148195% to $+116828\%$ as it did for the mean-variance strategy. Although the $1/N$ strategy has a higher Sharpe ratio than the minimum-variance strategy for the datasets "S&P Sectors," and "FF-4-factor," for the "Industry," "International," and "MKT/SMB/HML" datasets, the minimum-variance strategy has a higher Sharpe ratio, but 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, which is partly because of the small-firm effect. The out-of-sample Sharpe ratio for the "mp" approach proposed by MacKinlay and Pastor (2000) is also 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&P Sectors," "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 for the "FF-1-factor" and "FF-4-factor" datasets, but again with the p-value significant only for the "FF-1-factor" dataset.

Our fifth observation is that strategies that *combine* portfolio constraints with some form of shrinkage of expected returns are usually much more effective in reducing the effect of estimation error. This can be seen, for example, by examining 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. The results indicate that while the $1/N$ strategy has a higher Sharpe ratio than the "min-c" strategy for the "S&P Sectors" 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

Table 4
Certainty-equivalent returns for empirical data

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	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	−0.7816	−0.1365	0.0045	−2.7142	−0.0829
	(0.28)	(0.00)	(0.00)	(0.31)	(0.00)	(0.01)
bs	0.0030	−0.3157	−0.0312	0.0043	−0.6504	−0.0362
	(0.16)	(0.00)	(0.00)	(0.32)	(0.00)	(0.06)
dm ($\sigma_\alpha = 1.0\%$)	0.0052	−0.0319	0.0021	−0.0084	−0.0296	0.0110
	(0.11)	(0.01)	(0.08)	(0.04)	(0.00)	(0.11)
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	0.0014	0.0034	−0.0026	0.0054	0.0053
	(0.19)	(0.05)	(0.17)	(0.04)	(0.09)	(0.10)
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	−0.2337	−0.0066	0.0044	−0.0875	−0.0318
	(0.07)	(0.00)	(0.01)	(0.28)	(0.00)	(0.07)
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)

For each of the empirical datasets listed in Table 2, this table reports the monthly CEQ return for the $1/N$ strategy, the in-sample CEQ return of the mean-variance strategy, and the out-of-sample CEQ returns for the strategies from the models of optimal asset allocation listed in Table 1. In parentheses 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 2. The results for the "FF-3-factor" dataset are not reported because these are very similar to those for the "FF-1-factor" dataset.

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 has a higher Sharpe ratio than $1/N$ in all but two datasets, "S&P Sectors" and "FF-1-factor," although the superior performance is statistically significant for only the "FF-4-factor" dataset.[22]

Finally, the two mixture portfolios, "mv-min" and "ew-min," described in Sections 1.6.1 and 1.6.2, do not outperform $1/N$ in a statistically significant way.

3.2 Certainty equivalent returns

The comparison of CEQ returns in Table 4 confirms the conclusions from the analysis of Sharpe ratios: the in-sample mean-variance 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 $1/N$

[22] 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.

Table 5
Portfolio turnovers for empirical data

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	Mkt/ SMB/HML $N = 3$	FF-1-factor $N = 21$	FF-4-factor $N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
			Panel A: Relative turnover of each strategy			
mv (in sample)	–	–	–	–	–	–
mv	38.99	606594.36	4475.81	2.83	10466.10	3553.03
bs	22.41	10621.23	1777.22	1.85	11796.47	3417.81
dm ($\sigma_\alpha = 1.0\%$)	1.72	21744.35	60.97	76.30	918.40	32.46
min	6.54	21.65	7.30	1.11	45.47	6.83
vw	0	0	0	0	0	0
mp	1.10	11.98	6.29	59.41	2.39	2.07
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	9927.09	760.57	2.61	4292.16	4857.19
ew-min	4.82	15.66	4.24	1.11	34.10	6.80
			Panel B: Return-loss relative to $1/N$ (per month)			
mv (in sample)	–	–	–	–	–	–
mv	0.0145	231.8504	1.1689	0.0003	7.4030	1.5740
bs	0.0092	9.4602	0.3798	−0.0004	2.0858	1.1876
dm ($\sigma_\alpha = 1.0\%$)	0.0021	8.9987	0.0130	0.0393	0.1302	−0.0007
min	0.0048	0.0015	0.0000	−0.0004	−0.0008	0.0024
vw	−0.0001	0.0037	0.0012	0.0157	0.0021	0.0028
mp	0.0001	0.0050	0.0021	0.0227	0.0023	0.0030
mv-c	0.0085	0.0048	0.0034	0.0041	−0.0005	0.0002
bs-c	0.0061	0.0038	0.0030	0.0023	−0.0004	−0.0000
min-c	0.0042	−0.0001	−0.0007	−0.0004	0.0006	−0.0025
g-min-c	0.0019	−0.0003	−0.0006	−0.0003	0.0001	−0.0029
mv-min	0.0085	6.8115	0.1706	−0.0003	0.9306	1.8979
ew-min	0.0030	0.0008	−0.0001	−0.0004	−0.0011	0.0024

For each of the empirical datasets listed in Table 2, the first line of this table reports the monthly turnover for the $1/N$ strategy, panel A reports the turnover for the strategies from each optimizing model *relative* to the turnover of the $1/N$ model, and panel B reports the return-loss, which is the extra return a strategy needs to provide in order that its Sharpe ratio equal that of the $1/N$ strategy in the presence of proportional transactions costs of 50 basis points. 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. 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 constrained-mean-variance 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, with the p-values of the differences being 0.03 and 0.05, respectively.

3.3 Portfolio turnover

Table 5 contains the results for portfolio turnover, our third metric of performance. The first line reports the actual turnover of the $1/N$ strategy. Panel A reports the turnover of all the strategies relative to that of the $1/N$ strategy, and in panel B we report the return-loss, as defined in Equation (17).

From panel A of Table 5, we see that in all cases the turnover for the portfolios from the optimizing models is much higher than for the benchmark $1/N$ strategy. Comparing the turnover across the various datasets, 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 ($N = 3$), the estimation problem is less severe. This is also confirmed by panel B of the table, where for the "MKT/SMB/HML" dataset several strategies have a return-loss that is slightly negative, implying that even in the presence of proportional transactions costs, these strategies attain a higher Sharpe ratio than that of the $1/N$ strategy.

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 usually successful in reducing turnover, relative to 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. From panel B of Table 5, we see that for some of the datasets the "min-c" and "g-min-c" strategies have a slightly negative return-loss, implying that in these cases these strategies achieve a higher Sharpe ratio than that of the $1/N$ strategy even in the presence of proportional transactions costs.

3.4 Summary of findings from the empirical datasets

From the above discussion, we conclude that of the strategies from the optimizing models, there is no single strategy that always dominates the $1/N$ strategy 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.

4. Results from Studying Analytically the Estimation Error

This section examines analytically some of the determinants of the empirical results identified previously. 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.

As in Kan and Zhou (2007), 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)]:

$$U(\mathrm{x}) = \mathrm{x}^\top \mu - \frac{\gamma}{2}\mathrm{x}^\top \Sigma \mathrm{x}. \tag{18}$$

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

$$U(\mathrm{x}^*) = \frac{1}{2\gamma}\mu\Sigma^{-1}\mu \equiv \frac{1}{2\gamma}S_*^2, \tag{19}$$

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:

$$\hat{\mathrm{x}} = f(R_1, R_2, \ldots, R_M). \tag{20}$$

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

$$L(\mathrm{x}^*, \hat{\mathrm{x}}) = U(\mathrm{x}^*) - E[U(\hat{\mathrm{x}})], \tag{21}$$

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 (2007), 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,

$$M^*_{\text{mv}} \equiv \inf\{M : L_{\text{mv}}(\mathrm{x}^*, \hat{\mathrm{x}}) < L_{\text{ew}}(\mathrm{x}^*, \mathrm{w}^{\text{ew}})\}. \tag{22}$$

Just as in Kan and Zhou, we consider three 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 sample-based mean-variance rule to outperform the $1/N$ 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 sample-based mean-variance strategy has a lower expected loss than the 1/N strategy if:*

$$S_*^2 - S_{ew}^2 - \frac{N}{M} > 0. \tag{23}$$

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

$$kS_*^2 - S_{ew}^2 > 0, \tag{24}$$

$$\textit{where} \qquad k = \left(\frac{M}{M-N-2}\right)\left(2 - \frac{M(M-2)}{(M-N-1)(M-N-4)}\right) < 1. \tag{25}$$

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

$$kS_*^2 - S_{ew}^2 - h > 0, \tag{26}$$

$$\textit{where} \qquad h = \frac{NM(M-2)}{(M-N-1)(M-N-2)(M-N-4)} > 0. \tag{27}$$

From the inequality (23), we see that if μ is unknown but Σ is known, then the sample-based mean-variance strategy is more likely to outperform the $1/N$ strategy if the number of periods over which the parameters are estimated, M, is high and if the number of available assets, N, is low. Because k in Equation (25) is increasing in M and decreasing in N, the inequality (24) shows that also for the case where μ is known but Σ is unknown, the sample-based mean-variance policy is more likely to outperform the $1/N$ strategy as M increases and N decreases. Finally, for the case in which both parameters are unknown, we note that because $h > 0$, the left-hand side of Equation (26) is always smaller than the left-hand side of Equation (24).

To illustrate the implications of Proposition 1 above, we compute the critical value M^*_{mv}, as defined in Equation (22), 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&P Sectors 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 US stock market, we set $S_{\text{ew}} = 0.12$ and $S_{\text{ew}} = 0.08$, respectively; these values are obtained from Table 6 for simulated data; the details of the simulated data are provided in Section 5.1.

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 more 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 1200 months. Even for the more extreme

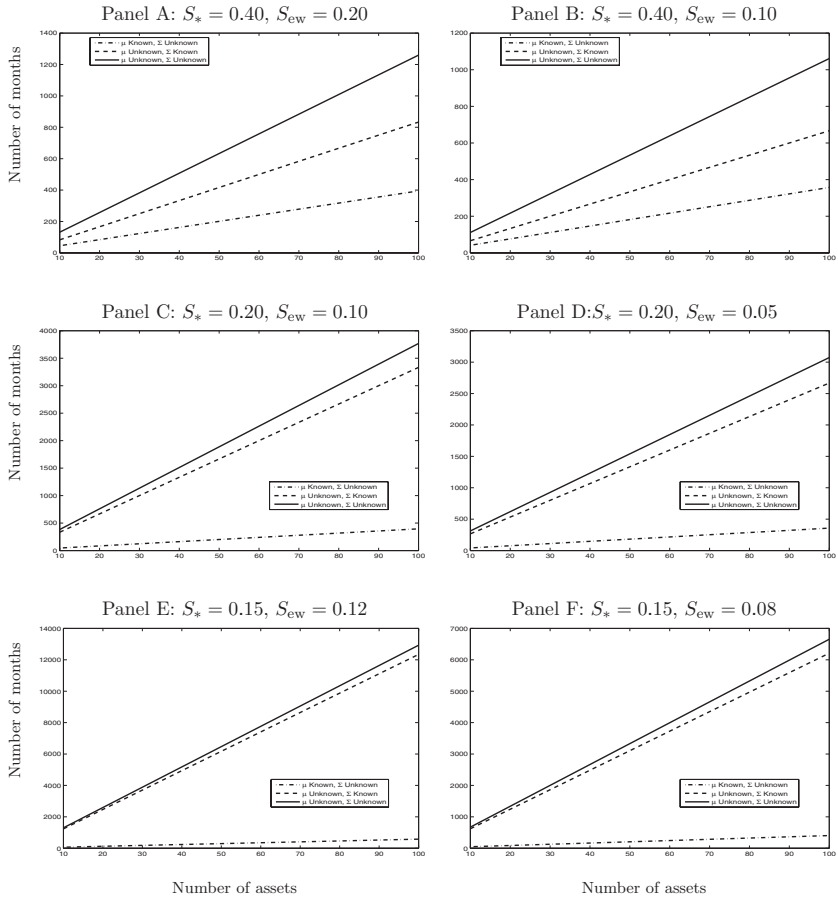

Figure 1
Number of estimation months for mean-variance portfolio to outperform the $1/N$ benchmark
The six panels in this figure show the critical number of estimation months required for the sample-based mean-variance strategy to outperform the $1/N$ rule on average, as a function of the number of assets, N. Each panel is drawn for different levels of the Sharpe ratios for the *ex ante* mean-variance portfolio, S_*, and the equally weighted portfolio, S_{ew}. Critical values are computed using the definition in Equation (22). 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 neither the means nor the covariances are known.

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 1060 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 of Figure 1, 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 1000 months of data are needed. If the number of assets is 50, the length of the estimation window increases to about 2000 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 1500 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 3000 months, and for a portfolio with 50 assets, it is more than 6000 months. Even in panel F, 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 1600 months, and for a portfolio with 50 assets, it is more than 3200 months.

5. Results from Simulated Data

The results in Section 4 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 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 that we simulate are based on a simple single-factor model, with returns that are distributed independently and identically over time with a normal distribution. Given that most of the models of optimal portfolio choice are derived under these assumptions, this setup should favor the mean-variance model and its various extensions. It also means that the results for the simulated data 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.

5.1 Details about how the simulated data are 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, which include K factors. The excess returns of the remaining $N - K$ risky assets are generated by the

factor model $R_{a,t} = \alpha + BR_{b,t} + \epsilon_t$, where $R_{a,t}$ is the excess asset returns vector, α is the mispricing coefficients vector, B is the factor loadings matrix, $R_{b,t}$ is the vector of excess returns on the factor portfolios, $R_b \sim N(\mu_b, \Omega_b)$, and ϵ_t is the vector of noise, $\epsilon \sim N(0, \Sigma_\epsilon)$, which is independent with respect to the factor portfolios.

For our simulations, we assume that the risk-free rate follows a normal distribution, with an annual average of 2% and a standard deviation of 2%. We assume that there is only one factor ($K = 1$), whose annual excess return has an annual average of 8% and standard deviation of 16%. The mispricing α is set to zero, and the factor loadings, B, are evenly spread between 0.5 and 1.5. Finally, the variance-covariance matrix of noise, Σ_ϵ, is assumed to be diagonal, with elements drawn from a uniform distribution with support [0.10, 0.30], so that the cross-sectional average annual idiosyncratic volatility is 20%. We consider cases with number of assets $N = \{10, 25, 50\}$, and estimation window lengths $M = \{120, 360, 6000\}$ months, which correspond to 10, 30, and 500 years. We use Monte Carlo sampling to generate monthly return data for T =24,000 months.

5.2 Discussion of results from simulated data

The Sharpe ratios of the various portfolio policies are reported in Table 6. 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 labeled "mv-true," and, as expected, has the highest Sharpe ratio of all policies.

The simulation results confirm the insights obtained from the analytical results in Section 4: very long estimation windows are required before the sample-based mean-variance policy, "mv," achieves a higher out-of-sample Sharpe ratio than the $1/N$ policy. Moreover, the critical estimation window length 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 Sharpe ratio of the sample-based mean-variance policy is higher than that of the $1/N$ policy only for the case of $M = 6000$ months. For the cases with 25 and 50 assets, on the other hand, it does not achieve the same Sharpe ratio as the $1/N$ policy even for an estimation window length of 6000 months.

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 Bayes-Stein policy is very similar to that of the sample-based mean-variance policy. In particular, the Bayes-Stein policy, "bs," outperforms the $1/N$ benchmark only in the same cases as the sample-based mean-variance strategy. The Data-and-Model strategy, "dm," performs much better than the sample-based mean-variance and Bayes-Stein policies if $\sigma_\alpha = 1\%$ per annum. But the "dm" policy still needs more than 120 months of data to achieve a higher Sharpe ratio than the $1/N$ policy for the case with 10 assets, and does

Table 6
Sharpe ratios for simulated data

	$N = 10$			$N = 25$			$N = 50$		
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	−0.0019	0.0077	0.1416	0.0027	0.0059	0.1353	0.0078	−0.0030	0.1212
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
bs	−0.0021	0.0087	0.1416	0.0031	0.0074	0.1363	0.0076	−0.0035	0.1229
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
dm	0.0725	0.1475	0.1464	0.0133	0.1473	0.1457	0.0201	0.0380	0.1430
($\sigma_\alpha = 1.0\%$)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)	(0.29)	(0.00)	(0.00)	(0.02)
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	0.1171	0.1349	0.1354	0.1265	0.1442	0.1446	0.1312	0.1460	0.1465
	(0.00)	(0.24)	(0.40)	(0.00)	(0.21)	(0.43)	(0.00)	(0.10)	(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	−0.0029	0.0106	0.1414	0.0087	0.0172	0.1361	0.0016	−0.0068	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)

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 parentheses 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 2. These quantities are computed for simulated data that are described in Section 5.1 for different numbers of investable assets, N, and different lengths of the estimation window, M, measured in months.

not outperform the $1/N$ benchmark policy for the case of 50 assets even with an estimation window of 6000 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. On the other hand, the "mp" policy does quite well, and its performance is similar to that of the $1/N$ policy. One reason why this policy performs well is that the data are simulated assuming the market model with no unobservable factors, which are ideal conditions for this policy.

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 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 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, 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.

Finally, we examine the mixture portfolio strategies. The performance of the Kan and Zhou (2007) policy, "mv-min," produces Sharpe ratios that are very similar to those for the sample-based mean-variance policy, "mv." The second mixture policy, "ew-min," performs better than the "mv-min" mixture policy for short estimation windows, but is dominated by the minimum-variance policy with constraints, "min-c."

All else being equal, the performance of the sample-based mean-variance (and that of the optimizing policies in general) would improve relative to that of the $1/N$ policy if the idiosyncratic asset volatility was much higher than 20%. To see this, note first that while the higher idiosyncratic volatility would not affect the Sharpe ratio of the true mean-variance policy because this is simply the Sharpe ratio of the market factor, the Sharpe ratio of the $1/N$ policy would decrease. A second reason for the optimizing policies to perform relatively better is that with higher idiosyncratic volatility the covariance matrix of returns is less likely to be singular, and hence, easier to invert. For instance, we find that for a cross-sectional average idiosyncratic volatility of 75% per annum, the portfolio strategies from the optimizing models outperform the $1/N$ strategy if the number of assets is about 10 and the estimation window is longer than 120 months.[23]

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.

6. Results for Other Specifications: Robustness Checks

In the benchmark case reported in Tables 3–5, we have assumed that: (i) the length of the estimation window is $M = 120$ months rather than $M = 60$; (ii) the estimation window is rolling, rather than increasing with time; (iii) the holding period is one month rather than one year; (iv) the portfolio evaluated is that consisting of only-risky assets rather than one that also includes the risk-free asset; (v) one can invest also in the factor portfolios; (vi) the performance is measured relative to the $1/N$-with-rebalancing strategy, rather than the $1/N$-buy-and-hold strategy; (vii) the investor has a risk aversion of $\gamma = 1$, rather

[23] In the interest of space, the table with the results for the case with idiosyncratic volatility of 75% is not reported in the paper.

than some other value of risk aversion, say $\gamma = \{2, 3, 4, 5, 10\}$; and (viii) the investor's level of confidence in the asset-pricing model is $\sigma_\alpha = 1\%$ per annum, rather than 2% or 0.5%. To check whether our results are sensitive to these assumptions, we generate tables for the Sharpe ratio, CEQ returns, and turnover for all policies and empirical datasets considered after relaxing each of the assumptions aforementioned. In addition, based on each of these three measures, we also report the rankings of the various strategies. Because of the large number of tables for these robustness experiments, we have collected the results for these experiments in a separate appendix titled "Implementation Details and Robustness Checks," which is available from the authors. The main insight from these robustness checks is that the relative performance reported in the paper for the various strategies is not very sensitive to any of these assumptions.

Two other assumptions that we have made are that the investor uses only moments of asset returns to form portfolios, but not asset-specific characteristics, and that the moments of asset returns are constant over time. We discuss these as follows.

6.1 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 (2007) 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. 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 find that portfolios tend to load on small stocks, value stocks, and past winners.

In order to compare the out-of-sample performance of the Brandt, Santa-Clara, and Valkanov (2007) 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 Web site, are (i) 10 industry portfolios; and (ii) 25 size- and book-to-market-sorted portfolios. As we did in our empirical analysis, we use the rolling-window approach, with the length of the estimation period being $M = 120$ months.

For the 10 industry portfolios, the Brandt, Santa-Clara, and Valkanov (2007) model has an out-of-sample Sharpe ratio of 0.1882, while that of the $1/N$ strategy is only 0.1390, but the p-value for the difference is 0.11. And, the turnover for the Brandt, Santa-Clara, and Valkanov model is about 52 times greater than that for the $1/N$ strategy. This corresponds to a return-loss of

0.0025, which implies that in the presence of proportional transactions costs of 50 basis points, the Brandt, Santa-Clara, and Valkanov strategy would need to earn an additional return of 0.0025 per month in order to attain the same Sharpe ratio as the $1/N$ strategy. For the 25 size- and book-to-market-sorted portfolios, we find that using the Brandt, Santa-Clara, and Valkanov model gives an out-of-sample Sharpe ratio of 0.1824, which is higher than that for the $1/N$ strategy, 0.1649, but the p-value for the difference is 0.38. Again, the turnover for the Brandt, Santa-Clara, and Valkanov portfolio is about 127 times greater than that for the $1/N$ strategy and the return-loss is 0.0092. 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 (2007) 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 0.2120 relative to only 0.1387 for the $1/N$ strategy. However, the p-value for the difference is 0.1302, and the turnover for the Brandt, Santa-Clara, and Valkanov strategy is now 167 times rather than 52 times that for the $1/N$ strategy, with the return-loss being 0.0047 instead of 0.0025. So, in this case, the improvement is not striking. On the other hand, if one has 100 instead of just 25 size- and book-to-market-sorted portfolios, then the Sharpe ratio improves to 0.3622, while that of the $1/N$ strategy is 0.1217, with p-value for the difference now being 0.0; moreover, even though the turnover for the Brandt, Santa-Clara, and Valkanov strategy is now 158 times greater than that for the $1/N$ benchmark strategy, the return-loss is -0.0795, indicating that this strategy would earn a higher Sharpe ratio than the $1/N$ strategy even after a proportional transactions cost of 50 basis points.

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. But, the turnover of the Brandt, Santa-Clara, and Valkanov (2007) 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 explain returns on country indexes.

6.2 Time-varying moments of asset returns

Another limitation of the analysis described in this paper is that all the optimizing models considered assume that the moments of returns are *constant* over time. If the first and second moments of returns vary over time, then these models may perform poorly. While a formal model with strategies that account for time-varying moments could potentially outperform the naive $1/N$ rule, because of the larger number of parameters that need to be estimated for such a model, it is not clear that these gains can be achieved out of sample. In an earlier version of this paper, we partially address this issue by considering two models of *dynamic* asset allocation that allow for (i) stochastic interest rates, as in Campbell and Viceira (2001); and (ii) time-varying expected returns, as in Campbell and Viceira (1999) and Campbell, Chan, and Viceira (2003). We find that the estimation error is exacerbated by the large number of parameters that need to be estimated and that, in general, the dynamic strategies do not outperform the $1/N$ rule.

7. Conclusions

We have compared the performance of 14 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 simulated data. We find that the *out-of-sample* Sharpe ratio of the sample-based mean-variance strategy is much lower than that of the $1/N$ strategy, indicating that the errors in estimating means and covariances erode all the gains from optimal, relative to naive, diversification. We also find that the various extensions to the sample-based mean-variance model that have been 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. In summary, we find that of the various optimizing models in the literature, there is no single model that consistently delivers a Sharpe ratio or a CEQ 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 US stock-market data, we find that for a portfolio with only 25 assets, the estimation window needed is more than 3000 months, and for a portfolio with 50 assets, it is more than 6000 months, while typically these parameters are estimated using 60–120 months of data. Using simulated data, we show that the various extensions to the sample-based mean-variance model that have been designed to deal with estimation error reduce only moderately the estimation window needed for these models to outperform the naive $1/N$ benchmark.

These findings have two important implications. First, 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 to 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 (2007) shows, exploiting information about the cross-sectional characteristics of assets may be a promising direction to pursue. Second, in order to evaluate the performance of a particular strategy for optimal asset allocation, proposed either by academic research or by the investment-management industry, the $1/N$ naive-diversification rule should serve at least as a first obvious benchmark.

Appendix A: 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 Web site). A list of the datasets considered is given in Table 2.

A.1 Sector portfolios

The "S&P Sectors" dataset consists of monthly excess returns on 10 value-weighted industry portfolios formed by using the Global Industry Classification Standard (GICS) developed by Standard & Poor's (S&P) and Morgan Stanley Capital International (MSCI). The dataset has been created by Roberto Wessels, and we are grateful to him for making it available to us. The 10 industries considered are Energy, Material, Industrials, Consumer-Discretionary, Consumer-Staples, Healthcare, Financials, Information-Technology, Telecommunications, and Utilities. The data span from January 1981 to December 2002. We augment the dataset by adding as a factor the excess return on the US 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.

A.2 Industry portfolios

The "Industry" dataset consists of monthly excess returns on 10 industry portfolios in the United States. The 10 industries considered are Consumer-Discretionary, Consumer-Staples, Manufacturing, Energy, High-Tech, Telecommunication, Wholesale and Retail, Health, Utilities, and Others. The monthly returns range from July 1963 to November 2004 and were obtained from Kenneth French's Web site. We augment the dataset by adding as a factor the excess return on the US equity market portfolio, MKT.

A.3 International equity indexes

The "International" dataset includes eight international equity indices: Canada, France, Germany, Italy, Japan, Switzerland, the UK, and the US. 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).

A.4 MKT, SMB, and HML portfolios

The "MKT/SMB/HML" 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 US equity market; (ii) HML, a zero-cost 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 consist of monthly returns from July 1963 to November 2004. The data are taken from Kenneth French's Web site. The Data-and-Model approach is implemented by assuming that the investor takes into account his beliefs in an asset-pricing model (CAPM; Fama and French, 1993; or Carhart, 1997) when constructing the expected asset returns.

A.5 Size- and book-to-market-sorted portfolios

The data consist of monthly returns on the 20 portfolios sorted by size and book-to-market.[24] The data are obtained from Kenneth French's Web site 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, denoted by "FF-1-factor," we augment the dataset by adding the 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, denoted by "FF-3-factor," we augment the dataset by adding the 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, denoted by "FF-4-factor," we augment the size- and book-to-market-sorted portfolios with four-factor portfolios: MKT, HML, SMB, and the momentum portfolio, UMD, which is also obtained from Kenneth French's Web site. For this dataset, the investor is assumed to estimate expected returns using a four-factor model.

Appendix B: Proof for Proposition 1

Assuming that the distribution of returns is jointly normal, Kan and Zhou (2007) derive the following expression for the expected loss from using the sample-based mean-variance policy with estimated rather than true parameters: when

[24] 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 25 Fama-French portfolios.

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

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}|\Sigma) = \frac{1}{2\gamma}\frac{N}{M}. \tag{B1}$$

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

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}|\mu) = \frac{1}{2\gamma}(1-k)S_*^2, \tag{B2}$$

with k given in Equation (25). When both μ and Σ are unknown, the expected loss is

$$L(\mathrm{x}^*, \hat{\mathrm{x}}^{\mathrm{mv}}) = \frac{1}{2\gamma}\left((1-k)S_*^2 + h\right), \tag{B3}$$

with k given in Equation (25) and h given in Equation (27). Following a similar approach, we derive the expected loss from using the $1/N$ rule. Formally, let x^{ew} denote the equally weighted policy:

$$\mathrm{x}^{\mathrm{ew}} = c\,\mathbf{1}_N, \quad c \in \mathbb{R}. \tag{B4}$$

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 $1 - Nc$ is the fraction invested in the risk-free asset.

Suppose the investor uses the rule (B4) 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

$$L(\mathrm{x}^*, \mathrm{x}^{\mathrm{ew}}) = U(\mathrm{x}^*) - c\mathbf{1}_N^\top\mu + \frac{\gamma}{2}c^2\mathbf{1}_N\Sigma\mathbf{1}_N. \tag{B5}$$

In order to fully isolate the cost of using the $1/N$ rule and avoid the effects of market timing, we assume that the investor chooses c optimally; that is, in such a way that the loss in Equation (B5) 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:

$$L_{\mathrm{ew}}(\mathrm{x}^*, \mathrm{x}^{\mathrm{ew}}) = \frac{1}{2\gamma}\left(\mu^\top\Sigma^{-1}\mu - \frac{\left(\mathbf{1}_N^\top\mu\right)^2}{\mathbf{1}_N^\top\Sigma\mathbf{1}_N}\right) \equiv \frac{1}{2\gamma}\left(S_*^2 - S_{\mathrm{ew}}^2\right), \tag{B6}$$

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 Equations (B1), (B2), and (B3) to Equation (B6) gives the result in the proposition. ■

References

Balduzzi, P., and A. W. Lynch. 1999. Transaction Costs and Predictability: Some Utility Cost Calculations. *Journal of Financial Economics* 52:47–78.

Barone, L. 2006. Bruno de Finetti, The Problem of "Full-Risk Insurances." *Journal of Investment Management* 4:19–43.

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*. Amsterdam, The Netherlands: North Holland.

Benartzi, S., and R. Thaler. 2001. Naive Diversification Strategies in Defined Contribution Saving Plans. *American Economic Review* 91:79–98.

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

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.

Bhardwaj, R., and L. Brooks. 1992. The January Anomaly: Effects of Low Share Price, Transaction Costs, and Bid-Ask Bias. *Journal of Finance* 47:553–74.

Black, F., and R. Litterman. 1990. Asset Allocation: Combining Investor Views with Market Equilibrium. Discussion Paper, Goldman, Sachs & Co.

Black, F., and R. Litterman. 1992. Global Portfolio Optimization. *Financial Analysts Journal* 48:28–43.

Bloomfield, T., R. Leftwich, and J. Long. 1977. Portfolio Strategies and Performance. *Journal of Financial Economics* 5:201–18.

Brandt, M. W. 2007. Portfolio Choice Problems, in Y. Ait-Sahalia, and L. P. Hansen (eds.), *Handbook of Financial Econometrics*. St. Louis, MO: Elsevier, forthcoming.

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2007. Parametric Portfolio Policies: Exploiting Characteristics in the Cross Section of Equity Return. Working Paper, UCLA.

Brown, S. 1979. The Effect of Estimation Risk on Capital Market Equilibrium. *Journal of Financial and Quantitative Analysis* 14:215–20.

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* 114:433–95.

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*. New York: Oxford University Press.

Carhart, M. 1997. On the Persistence in Mutual Fund Performance. *Journal of Finance* 52:57–82.

Chan, L. K. C., J. Karceski, and J. Lakonishok. 1999. On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *The 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.

Fama, E. F., and K. R. French. 1993. Common Risk Factors in the Returns on Stock and Bonds. *Journal of Financial Economics* 33:3–56.

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.

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

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *The Review of Financial Studies* 20:41–81.

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

Green, R., and B. Hollifield. 1992. When Will Mean-Variance Efficient Portfolios Be Well Diversified? *Journal of Finance* 47:1785–809.

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

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 (eds.), *Modern Developments in Investment Management*, 2nd ed. Hinsdale, IL: Dryden Press.

Huberman, G., and W. Jiang. 2006. Offering vs. Choice in 401(k) Plans: Equity Exposure and Number of Funds. *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* 58:1651–84.

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

Jobson, J. D., and R. Korkie. 1980. Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association* 75:544–54.

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

Jobson, J. D., R. Korkie, and V. Ratti. 1979. Improved Estimation for Markowitz Portfolios Using James-Stein Type Estimators. *Proceedings of the American Statistical Association* 41:279–92.

Jorion, P. 1985. International Portfolio Diversification with Estimation Risk. *Journal of Business* 58:259–78.

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. 2007. Optimal Portfolio Choice with Parameter Uncertainty. *Journal of Financial and Quantitative Analysis* 42:621–56.

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*. Boston, MA: Houghton Mifflin.

Ledoit, O., and M. Wolf. 2004a. Honey, I Shrunk the Sample Covariance Matrix: Problems in Mean-Variance Optimization. *Journal of Portfolio Management* 30:110–19.

Ledoit, O., and M. Wolf. 2004b. A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis* 88:365–411.

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka. 1999. A New Estimate of Transaction Costs. *The Review of Financial Studies* 12:1113–41.

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

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

Markowitz, H. M. 1952. Portfolio Selection. *Journal of Finance* 7:77–91.

Markowitz, H. M. 1956. The Optimization of a Quadratic Function Subject to Linear Constraints. *Naval Research Logistics Quarterly* 3:111–33.

Markowitz, H. M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. Cowles Foundation Monograph #16. New York: Wiley.

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

Merton, R. C. 1980. On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics* 8:323–61.

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

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

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

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.

Roy, A. D. 1952. Safety First and the Holding of Assets. *Econometrica* 20:431–49.

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 Berkeley Symposium on Probability and Statistics* 1, 197–206. Berkeley, CA: University of California Press.

Stoll, H., and R. Whaley. 1983. Transaction Costs and the Small Firm Effect. *Journal of Financial Economics* 12:57–79.

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

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.