Financial Analysts Journal | A Publication of CFA Institute
https://doi.org/10.1080/0015198X.2019.1600958

In Defense of Portfolio Optimization: What If We Can Forecast?

David Allen, Colin Lizieri, and Stephen Satchell

David Allen is head of long/short strategies at Plato Investment Management, Sydney, Australia, and an Honorary Associate of the University of Technology Sydney. Colin Lizieri is the Grosvenor Professor of Real Estate Finance at the University of Cambridge, Cambridge, United Kingdom. Stephen Satchell is a professor of finance at the University of Sydney Business School, Sydney, and a fellow of Trinity College, University of Cambridge, Cambridge, United Kingdom.

We challenge the academic consensus that estimation error makes mean–variance portfolio strategies inferior to passive equal-weighted approaches. We demonstrate analytically, via simulation, and empirically that investors endowed with modest forecasting ability benefit substantially from a mean–variance approach. An investor with some forecasting ability improves expected utility by increasing the number of assets considered. We frame our study realistically using budget constraints, transaction costs, and out-of-sample testing for a wide range of investments. We derive practical decision rules to choose between passive and mean–variance optimization and generate results consistent with much financial market practice and the original Markowitz formulation.

Disclosure: The authors report no conflicts of interest.

CE Credits: 1

Markowitz's influential *Portfolio Selection* (Markowitz 1959) underpins many advances in financial economics and is one of the most widely used quantitative approaches to portfolio construction in the industry.[1] Nonetheless, the approach of allocating capital to risky assets by maximizing expected return for a given level of risk has been subject to numerous criticisms (e.g., see Fisher and Statman 1997). One criticism, that the mean–variance optimization model requires either normally distributed returns or investor quadratic utility, sits in contradiction to both Markowitz's original formulation and his subsequent research.[2] Normal distributions are *not* required; that the distribution be characterized by location and scale suffices.[3] A second criticism, often attributed to Michaud (1989), involves the phenomenon of error maximization, in which return and covariance forecast errors are magnified in the estimated portfolio weights, which may lead to poor out-of-sample performance.

Given doubts (both normative and positive) about forecast ability and the presence of noise in return series, some authors have concluded that mean–variance optimization is inferior to more passive strategies, such as the equal-weighted 1/*N* approach. We examine this choice by exploring the relationship between forecasting ability and investor welfare. We demonstrate, analytically, by way of simulation, and empirically, that mean variance outperforms 1/*N* with very modest forecasting ability, contrary to views expressed in much of the literature, even accounting for the presence of error. We do not claim that we can eliminate estimation error in any practical way. Nonetheless, our results present a much more compelling case for the use of portfolio optimization techniques. Our findings have an intuitive logic, to the extent that the implications may seem obvious: If you have good

The authors would like to thank Professor Stephen Brown, Nick Baltas, Steven Thorley, the editorial team at the Financial Analysts Journal, *and one anonymous reviewer for their very helpful comments and suggestions. We would also like to acknowledge the contributions of Professor Mark Kritzman, Dr. Mark Thompson, and participants at seminars in Cambridge, Oxford, and Sydney in shaping the paper.*

forecasting skill, make use of it; if not, then equal-weighting or risk-weighting a portfolio may be preferable, particularly in the absence of good estimation windows for mean and covariance. We provide a formal demonstration of this intuition.

Our analytical approach allows us to contrast mean–variance with 1/*N* allocation and explicitly identify the drivers of expected utility. We show application of the research to a range of realistic investment problems and reexamine the influential model of DeMiguel, Garlappi, and Uppal (2009), which has supported advocacy of 1/*N*. We extend our analytical results in a high-dimension simulation framework (in the spirit of Kane, Kim, and White 2010), which provides strong confirmatory results.

We then step away from the theoretical (independent and identically distributed) normal framework to conduct out-of-sample empirical tests involving rolling portfolio rebalancing with short estimation windows and realistically large sets of potential assets. Mean–variance performs well in this context.

First, building on Grinold's (1989) work relating forecasting ability to the information ratio, we provide a utility-maximizing framework that unifies estimation error and forecasting ability effects. Our results show that the Kan and Zhou (2007) result that increasing the number of assets leads to performance deterioration does not necessarily hold when one allows for forecasting, even in the presence of estimation error. Second, we provide a closed-form solution for the amount of skill required for mean–variance optimization to outperform the 1/*N* approach *ex ante*. Third, our empirical results show the conditions in which mean–variance optimization can outperform.

Before proceeding with the analysis, we briefly review the existing empirical and analytical literature on mean–variance portfolio approaches.

Mean–Variance Optimization: Errors in Estimation and Forecasting Ability

The mean–variance approach seeks to overweight assets with low correlations, high expected returns, and low relative variance. We note Michaud's (1989) argument that in empirical settings, such assets may be subject to substantial estimation error. He coined the term "error maximization" for the resulting portfolios, in which overweighting leads to most assets being driven from the portfolio and results in "corner solutions." As Scherer (2002) noted, with two highly correlated assets, the optimization algorithm will tend to take long positions in the asset with the higher expected return and short positions in the other asset, even though the return difference may be within error margins.

Numerous studies have suggested that mean–variance optimization does not outperform passive benchmarks, including 1/*N* approaches, when tested in an out-of-sample context (e.g., Jobson and Korkie 1981; Jagannathan and Ma 2003). An important contribution comes from DeMiguel et al. (2009), who compared 14 portfolio allocation models (many of which were designed to reduce the impact of estimation error) for seven data sets and concluded that "none is consistently better than the 1/*N* rule" (p. 1). This result held when the Sharpe ratio, certainty equivalent,[4] and turnover measures were used. The authors suggested that their results cast significant doubt on the utility of active portfolio investment strategies.

The majority of those studies critical of the mean–variance approach took as their starting point *ex post* sample estimates as forecasts of expected returns and covariances, often over short time series, despite substantial evidence that sample moments have low predictive power. In Markowitz's original formulation, the proposed E-V rule starts from the *expected* return:

> To use the E-V rule in the selection of securities, we must have procedures for finding reasonable μ_i and σ_{ij}. These procedures, I believe, should combine statistical techniques and the judgment of practical men. (Markowitz 1952, p. 91)

Some 60 years after the original formulation, he reiterated this view:

> Judgment plays an essential role in the proper application of risk–return analysis for individual and institutional portfolios. For example, the estimates of mean, variance, and covariance of a mean–variance analysis should be forward-looking rather than purely historical. (Markowitz 2010, p. 7)

A branch of the literature uses long time series to generate sample estimates in place of short windows. For example, Kritzman, Page, and Turkington (2010), using long-term historical averages and covariances as inputs, suggested that mean–variance optimization generates higher out-of-sample Sharpe ratios than the 1/*N* approach.[5] Furthermore, the authors demonstrated that virtually any reasonable return forecast applied to mean–variance optimization would outperform 1/*N* and that DeMiguel et al. (2009), by using rolling five-year means as forecasts, were, in effect, assuming negative forecasting ability.

Commonly, a range of models and approaches has been used to reduce the impact of estimation error, including Bayesian shrinkage techniques, resampling, and weight constraints—see, for example, Black and Litterman (1991), Michaud (1998), and Frost and Savarino (1988). Although some of these techniques have been shown to reduce the impact of estimation error and accord with industry practice, we will not use them, in order to provide a pure evaluation of the performance of the mean–variance approach in the face of errors and noise.

A small body of research has used one-period-ahead return forecasts to examine the performance of mean–variance optimization—an approach that is closer in spirit to Markowitz's original proposal. Most of the research combining linear forecasting with mean–variance optimization does show outperformance relative to passive benchmarks.[6] Typically, however, such research used small asset sets (where the effect of estimation error is less extreme than in large sets). In practical applications, higher dimensions are needed—for example, for equity funds benchmarked to the Russell 3000 Index or the MSCI indexes—and in these situations, estimation error will become problematic. Therefore, we focused our simulations and empirical work on high-dimension problems where the impact of estimation error is most severe.

The expected utility of using mean–variance optimization with forecasting will depend on the quality of those forecasts. Grinold (1989) and Grinold and Kahn (1999) developed expressions for expected utility. The best known is Grinold and Kahn's "fundamental law of active management," which relates the information ratio (IR)—the active return divided by the active risk—to the information coefficient (IC)—the correlation between forecast and realized returns—and to the number of stock positions:

$$IR = IC\sqrt{N}.$$

They concluded, "it takes only a modest amount of skill to win so long as that skill is deployed frequently and across a large number of stocks" (Grinold and Kahn 1999, p. 162).

At issue, however, is how long the sample period and large the number of assets must be for active management to outperform passive strategies. Zhou (2008) argued that more than 10,000 months of data are required for mean–variance optimization to achieve 90% of true maximum utility. DeMiguel et al. (2009) argued that for a portfolio containing 25 assets, 3,000 months of estimation data are required for mean–variance optimization to outperform the 1/*N* approach; this period increases to 12,000 months for a portfolio of 100 assets. They concluded that this infeasible requirement means that the 1/*N* approach should be preferred to mean–variance optimization.

The diametrically opposed conclusions of Grinold and Kahn (1999) and DeMiguel et al. (2009) result from their different underlying assumptions. DeMiguel et al. assumed estimation error but no forecasting ability. Thus, in their view, increasing the number of assets increases estimation error effects, damaging expected utility. In contrast, Grinold and Kahn assumed forecasting ability but not estimation error. Hence, in their view, increasing the breadth of investment opportunities increases utility. We seek to reconcile these differences by considering *both* forecasting ability *and* estimation error in the same modeling approach.

From the 1980s, academics have increasingly reported evidence of partial predictability of returns from using publicly available information and of anomalies that cast doubt on pure market efficiency. This burgeoning literature is familiar to the reader, and we will not discuss it in depth here. Nonetheless, we note that documented relationships between dividend yield or equity returns with inflation, interest rates, and credit spreads and the presence of autocorrelation in financial series (Ball 1978; Rozeff 1984; Shiller 1984) led Fama and French (1988) to conclude "there is much evidence that stock returns are predictable," despite their commitment to the efficient market paradigm. Thus, to assume *some* forecasting ability seems reasonable. Those forecasts will be noisy, however, so estimation error must be considered.

We compare the performance of mean–variance optimization with that of the 1/*N* approach for a number of reasons. First, because the 1/*N* rule does not require any estimations, it carries no estimation risk. Second, it is easy to apply in the market; hence, it represents a practical strategy for investors. Moreover, others have used 1/*N* as a basis for benchmarking evaluations of mean–variance optimization, so we follow in that tradition. Finally, behavioral evidence suggests that investors, in the absence of specific skills, tend to equal-weight investment choices; for example, Benartzi and Thaler (2001) documented such behavior on the part of US investors choosing defined-contribution investment plans. These authors were not alone. Markowitz, questioned about his own allocations for retirement in his TIAA-CREF account, confessed, "I should have computed the historic covariances of the asset classes and drawn an efficient frontier. Instead . . . I split my contributions fifty-fifty between bonds and equities" (Zweig 1998, p. 115).

The 1/*N* allocation plan will be optimal in certain contexts. If the distribution of future returns is independent of current information and each asset has the same mean, variance, and pairwise correlation, then the mean–variance portfolio will coincide with the 1/*N* portfolio. In practice, the 1/*N* approach will be close to optimal for cases not too different from these assumptions.

We turn now to a consideration of mean–variance utility with both forecasting ability and error estimation.

Mean–Variance Utility: An Analytical Approach

In this section, we describe two models that include the factors that drive investors' expected utility. Our focus is the interplay between estimation error, forecasting ability, and the budget constraint. The models are based on one-period-ahead mean–variance optimization, which is consistent with commercially available portfolio optimization software and follows a long line of academic research.[7]

For simplicity and consistency with prior research, we assumed a constant IC across all assets. Previous literature suggested values of IC in the range of zero to 0.1 (which equates to R^2 values of 0% to 1%). This assumption implies little forecasting ability, which is intuitively problematic, because the mean–variance optimizer treats expected returns as certain. A reviewer pointed out to us, however, that an individual's ICs are likely to vary by market and sector because of that individual's varying expertise in different markets and sectors. This point favors 1/*N* investing, in which only *N* need vary with the investment set being considered.

The starting point of our model is a set of return forecasts, the covariances between those return forecasts, and a covariance vector relating forecast to actual returns. In setting up the model, we made several assumptions that are common in the portfolio strategy literature, including that forecast errors are independent and normal and that the covariance matrix is known (we relax this assumption in the simulation and the out-of-sample evaluation, presented later in the article). We directly modeled the budget constraint faced by investors,[8] and we used the constrained weights to derive optimum allocations. We provide the mathematical derivations in Appendix A of the online supplemental material.[9] The analysis allowed us to present two important models that explore the benefits of a mean–variance strategy for an investor who has some forecasting ability but is faced with estimation error and a budget constraint.

Model 1. The unconditional expected utility of the mean–variance investor under the assumptions of multiple forecasts, one for each asset, with a constant forecasting ability level (*IC*), estimation error, and a budget constraint, in the unconditional covariance matrix, is given by

$$E[U] = \frac{\alpha + (N-1)\left(IC^2 - \frac{1}{T}\right) - \frac{(\beta-\lambda)^2}{\gamma}}{2\lambda}, \tag{1}$$

where

$\alpha = \boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$ and refers to the squared Sharpe ratio of the unconstrained mean–variance portfolio

$\beta = \boldsymbol{\mu}'\Sigma^{-1}\mathbf{i}$

$\gamma = \mathbf{i}'\Sigma^{-1}\mathbf{i}$

λ = level of risk aversion

$\boldsymbol{\mu}$ = vector of returns

Σ = covariance matrix

$\mathbf{i}$ = vector of ones

T = length of the estimation window

This model brings together two strands of the portfolio literature—the impact of estimation error and the role of forecasting—while, by incorporating the budget constraint, allowing a fair comparison with the 1/*N* allocation rule. Equation 1 shows that expected utility is positively related to the Sharpe ratio of the unconstrained Markowitz portfolio and to forecasting level, IC. Expected utility is also positively related to the length of the estimation window, T, because expected utility increases as estimation error decreases. The relationship between the number of assets, N, and expected utility, however, is more complex. When other inputs are held constant, the effect on utility of increasing the number of assets depends on the level of forecasting ability relative to the number of estimation periods. For example, if forecasting ability is sufficiently large relative to T, utility will increase with an increase in available assets. This aspect helps reconcile the contradictory results of Grinold (1989), who concluded that the information ratio increases with the number of assets, and DeMiguel et al. (2009), who suggested that expected utility declines with additional assets.[10]

How much forecasting ability is needed for the mean–variance approach to generate returns superior to those of 1/N? We define the critical level of forecasting ability, IC^*, as the level of forecasting ability where the expected utility of mean–variance is equal to the expected utility of 1/N. If $IC < IC^*$, then the investor would be better off using 1/N.[11]

Setting Equation 1 equal to the expected utility of the 1/N investor, $V_{1/N}$, and solving for IC, we arrive at the forecasting ability level required for mean–variance to outperform 1/N. This resolution provides a useful decision rule for practitioners. If their forecasting ability exceeds IC^*, they should use mean–variance optimization; otherwise, they are better off using 1/N:

$$IC^* = \sqrt{\frac{2\lambda V_{1/N} - \alpha + \dfrac{N-1}{T} + \dfrac{(\beta-\lambda)^2}{\gamma}}{N-1}}. \tag{2}$$

Note that as the length of our estimation window, T, increases, IC^* falls and hence less forecasting skill is needed to beat 1/N. IC^* is increasing in risk aversion,[12] however, which may push more conservative investors toward 1/N.

From Model 1, we are also able to derive an expression for expected utility for a simplified case where volatility and pairwise correlation are constant across assets. This derivation allows us to shed further light on what drives expected utility.

Model 2. Our Model 2 states as follows:[13] The unconditional expected utility of the mean–variance investor under the assumptions of multiple forecasting variables a, constant forecasting ability level IC, a constant pairwise correlation ρ, a constant volatility σ, and a cross-sectional dispersion of mean returns σ_μ^2; and in the presence of estimation error in the unconditional covariance matrix and with a budget constraint, is given by:

$$E[U] = \frac{(N-1)\left[\dfrac{\sigma_\mu^2}{\sigma^2}(1-\rho) - \dfrac{1}{T} + IC^2\right] + O(1)}{2\lambda}. \tag{3}$$

This result shows that expected utility falls as both individual asset volatility and the correlation between assets increase. An important aspect, however, is that as the cross-sectional dispersion of expected returns increases, so too does utility. This pattern is consistent with the empirical results of Petajisto (2013). Investors using a strategy based on a portfolio optimizer benefit from heterogeneity in returns as measured by cross-sectional mean dispersion, σ_μ^2. This variable, then, is a key variable for active managers because it provides a measure of investment opportunity.

Applying the Models. We now use Model 1 to investigate the expected utility of an investor using a mean–variance approach, with some defined forecasting ability and in the presence of estimation error and a budget constraint.

We are able to compare our results directly with those of DeMiguel et al. (2009) by using the same data sets; the sole exception is that we did not have access to their 10 Standard & Poor's sectors. We anticipated that the results would be similar for the 10 Fama–French industries that we do include. When we refer to DeMiguel et al. in this section, we are referring to their analytical model of utility, not to their simulation and rolling portfolio rebalancing work, which we discuss later.

The *international* data set includes eight developed-market MSCI indexes and the MSCI World Index. The *industries* data set includes 10 US value-weighted industry portfolios. The MKT/SMB/HML data set includes the Fama–French market, size (small-minus-big), and book-to-market (high-minus-low) long–short factor portfolios. The *FF-1* data set includes the 20 portfolios formed by the intersection of the Fama–French size portfolios with the book-to-market portfolios and the market portfolio. The *FF-3* data set augments the *FF-1* data set with the market, SMB, and HML factor portfolios. The *FF-4* data set augments *FF-3* with the momentum factor, UMD (up minus down). We used the same inception points as in DeMiguel et al. (2009).

Table 1 shows that mean–variance optimization generates higher utility than the 1/*N* approach for every single data set even if forecasting ability is zero. We followed DeMiguel et al. (2009) in using truncated data sets for the majority of the tests (omitting almost 40 years of data prior to 1963). Adding these omitted data expands the estimation period and makes the conclusion still more decisive. Our results are derived from a risk aversion level of 5, but our conclusions are robust to other levels of risk aversion.

We attempted to reconcile these results with those of DeMiguel et al. (2009). Their analytical approach requires estimates of the squared Sharpe ratios of the 1/*N* and mean–variance optimization portfolios. These estimates were taken from their existing data sets, which contain between 3 and 24 assets. While DeMiguel et al. allowed the portfolios to grow to up to 100 assets, however, they left these estimates constant; the available set of assets greatly expanded while the distribution of returns remained unchanged. In practice, were we to add more countries or use more granular definitions of sectors, the Sharpe ratios of the mean–variance and 1/*N* portfolios would surely change.

A more natural approach would be to use the actual number of assets, the available estimation window, and the observed Sharpe ratios for different investment problems. As **Table 2** shows, with such a revised approach, use of the DeMiguel et al. (2009) model now shows mean–variance outperforming 1/*N* whatever the assumption about investor knowledge. The DeMiguel et al. result, therefore, seems to relate to the specific way the authors applied their model.

Table 3 sets out means, correlations, cross-sectional dispersions, and unconstrained Sharpe ratios for value-weighted industry portfolios of increasing granularity. The data are from Kenneth French's data library (with observations running from 1926 to 2013).[14] As the number of industries included increases from 5 to 48, the mean pairwise correlation falls by 28%, from 0.80 to 0.58, and the cross-sectional dispersion doubles. This result is intuitive: As granularity increases, idiosyncratic differences become more important.[15] The result also casts serious doubt on the validity of assuming a constant multivariate distribution of returns as asset numbers increase. The lower correlations and higher cross-sectional dispersions drive an increase in the mean–variance expected Sharpe ratio from 0.16 to 0.37. Hence, they increase the likelihood of mean–variance optimization outperforming the 1/*N* approach. As shown in Table 3, the DeMiguel et al. (2009) model indicates that mean–variance optimization will outperform a 1/*N* strategy in all cases, even though the model assumes no forecasting ability.

We acknowledge a limitation of our analytical approach in that, unlike DeMiguel et al. (2009),

Table 1. Utility of Mean–Variance Optimization vs. the 1/*N* Approach Using Model 1

	International	Industries	MKT/SMB/HML	*FF-1* Factor	*FF-4* Factor
Assets, *N*	9	11	3	21	24
Estimation window, *T*	379	497	497	497	497
Information coefficient					
IC = 0	✓	✓	✓	✓	✓
IC = 0.025	✓	✓	✓	✓	✓
IC = 0.05	✓	✓	✓	✓	✓
IC = 0.075	✓	✓	✓	✓	✓
IC = 0.10	✓	✓	✓	✓	✓

Notes: The estimation window, *T*, is the number of months used to estimate the expected return vector. A check mark signifies that the expected utility of the mean–variance approach exceeded the expected utility of the 1/*N* strategy.

Table 2. Utility of Mean–Variance Optimization vs. the 1/*N* Approach Using Proposition 1 of DeMiguel et al. (2009)

	S&P Sectors	International	Industries	MKT/SMB/HML	*FF-1* Factor	*FF-4* Factor
Assets, *N*	11	9	11	3	21	24
Estimation window, *T*	276	379	497	497	497	497
$SR_{1/N}$	0.19	0.13	0.14	0.22	0.16	0.18
SR_{mv}	0.39	0.21	0.21	0.29	0.51	0.54
Conditions						
μ is unknown, Σ is known	✓	✓	✓	✓	✓	✓
μ is known, Σ is unknown	✓	✓	✓	✓	✓	✓
μ is unknown, Σ is unknown	✓	✓	✓	✓	✓	✓

Notes: $SR_{1/N}$ and SR_{mv} refer to the in-sample monthly Sharpe ratios of, respectively, the 1/*N* and mean–variance portfolios. The estimation window, *T*, is the number of months used to estimate the expected return vector. A check mark signifies that the expected utility of the mean–variance approach exceeded the expected utility of the 1/*N* strategy.

Table 3. Expected Utility of Mean–Variance Optimization vs. the 1/*N* Approach Using Proposition 1 of DeMiguel et al. (2009): Industries Data, 1926–2013

	5 Industries	10 Industries	30 Industries	48 Industries
Assets, *N*	5	10	30	48
Estimation window, *T*	1,050	1,050	1,050	1,050
$\sigma_{\bar{X}}$	0.07	0.09	0.12	0.14
$\bar{\rho}_{i\neq j}$	0.80	0.72	0.64	0.58
$SR_{1/N}$	0.13	0.14	0.13	0.15
SR_{mv}	0.16	0.19	0.24	0.37
Conditions				
μ is unknown, Σ is known	✓	✓	✓	✓
μ is known, Σ is unknown	✓	✓	✓	✓
μ is unknown, Σ is unknown	✓	✓	✓	✓

Notes: $\sigma_{\bar{X}}$ refers to the cross-sectional dispersion of mean returns; $\bar{\rho}_{i\neq j}$ refers to the average pairwise correlation. See also the definitions in the Table 1 and Table 2 notes. A check mark signifies that the expected utility of the mean–variance approach exceeded the expected utility of the 1/*N* strategy.

we ignored the effect of estimation error in the covariance matrix. In theory, however, we can increase the sampling frequency of the covariance estimator to reduce estimation error to any arbitrary level. Cochran's (1934) theorem suggests that under normality, the sample variance based on *T* observations follows a scaled chi-squared distribution and the standard error of the sample error tends asymptotically to zero. Empirical support for this theorem can be found in the equity markets (Andersen, Bollerslev, Diebold, and Ebens 2001).

The theory that high-frequency data can be used to eliminate estimation error in the covariance matrix is constrained in practice because of microstructure issues, thin trading, and departures from normality, all of which place a limit on precision (Hansen and Lunde 2006). As a result, lower-frequency data tend to be used in conjunction with factor models to estimate the covariance matrix for investment problems involving large numbers of assets.

Mean–Variance Performance: A Simulation Approach

In this section, using simulation, we compare the expected performance of mean–variance and 1/*N* portfolios in the presence of forecasting ability. Simulation allows us to consider cases for which closed-form solutions for expected utility are not readily available. Consistent with Treynor and Black (1973) and Kane et al. (2010), for this analysis, we relaxed the assumption that estimation error in the covariance matrix can be eliminated and used the single-index model of Sharpe (1964). To help explain the poor performance of the mean–variance approach documented in DeMiguel et al. (2009), we also show here our examination of the optimal weight relationship used in the empirical work by these authors, together with the Treynor–Black and minimum-variance portfolios.

Replicating the research design of Kane et al. (2010), we simulated the activity of a hypothetical investment manager. In this framework, the manager uses *N* security analysts, each assigned to a single stock. Stock selection problems typically involve higher dimensions, shorter data histories, and lower cross-sectional dispersions than asset allocation problems because of the dominant equity market factor. For this reason, the stock selection problem should provide a more rigorous test of the benefits of the mean–variance approach.

The simulation was set up as follows. We generated 180 monthly market returns with a monthly excess return of 0.71% and a standard deviation of 4.33%. We randomly generated betas, β_i, for the *N* analyzed stocks from a normal distribution with a mean of 1.0 and a standard deviation of 0.3. We generated 180 abnormal returns, $z_{i,t}$, for each stock with a mean of zero and a cross-sectional variance of $\sigma^2_{z,i}$, where $\sigma_{z,i}$ was sampled from a lognormal cross-sectional distribution with μ = 1.98 and σ = 0.36.

Forecasts, z_i^f, of the abnormal returns were generated by using

$$z_i^f = r_{i,0} + r_{i,1} z_i + v_i, \tag{4}$$

where v_i is independent of z_i and is normally distributed with a mean of zero and a variance of $\sigma^2_{v,i}$. The variance is a function of the precision of the *i*th forecaster. We assumed that $r_{i,0} = 0$ and $r_{i,1} = 1$; these parameters are unknown to the investment manager, however, and must be estimated by regression. The level of error variance was set to give the desired level of forecasting ability as

$$\sigma_{v,i} = \sqrt{\frac{\left[\sigma^2_{z,i}\left(1 - R_i^2\right)\right]}{R_i^2}}. \tag{5}$$

We considered two skill levels: almost zero skill (*IC* = 0.001) and modest skill (*IC* = 0.07). We considered investment universe sizes of 10, 100, and 500 assets.

At the end of the 60th month, our investment manager uses the previous 60 months of data to calculate the market mean return and volatility, stock betas, realized abnormal returns, and expected abnormal returns. With these calculations, the manager computes the portfolio weights for, respectively, the mean–variance investor (we assume a budget-constrained investor with asset weights constrained to sum to 1.0); the Treynor–Black portfolio as in Kane et al. (2010), which is a constrained variant of the original Treynor–Black (1973) model; the DeMiguel et al. (2009) approach to determining optimal weighting for a mean–variance investor; the minimum-variance portfolio; and the 1/*N* portfolio. Then, the estimation window is shifted forward one month at a time and the exercise repeated until we had 120 months of out-of-sample performance for each model. The process was repeated 1,000 times for each level of forecasting ability and for each universe size, which enabled us to evaluate the statistical significance of the differences in relative performance.

We computed the average mean return, standard deviation, Sharpe ratio, and M^2 across simulations.[16]

In presenting the results, for consistency with prior research, we primarily focus on realized Sharpe ratios as a measure of risk-adjusted return. Panel A of **Table 4** presents the results for an investor with almost zero forecasting ability. The 1/*N* approach delivers a higher Sharpe ratio than mean–variance optimization whatever the size of the investment universe—a difference that is statistically significant at the 99% level in all cases. The 1/*N* rule also outperforms the other allocation models. Consistent,

Table 4. Simulation Summary: The Impact of Forecasting Ability

	Model	Mean	Standard Deviation	Sharpe Ratio	M^2
A. Almost zero forecasting ability					
Results for an investor with almost zero forecasting ability (R^2 = 0.000001, IC = 0.001)					
N = 10					
	Mean–variance: λ = 20	0.56**	4.91**	0.11**	–0.25**
	Mean–variance: λ = 100	0.54**	4.47**	0.12**	–0.22**
	Treynor–Black†	0.59**	10.64**	0.06**	–0.46**
	DeMiguel	8.64	406.00**	0.01**	–0.67**
	Minimum-variance	0.53**	4.45**	0.12**	–0.22**
	1/*N*	0.70	5.05	0.14	–0.14
N = 100					
	Mean–variance: λ = 20	0.47**	7.64**	0.06**	–0.48**
	Mean–variance: λ = 100	0.32**	2.9**	0.11**	–0.27**
	Treynor–Black†	0.67*	6.79**	0.10**	–0.31**
	DeMiguel	–41.47*	969.00**	0.01**	–0.68**
	Minimum-variance	0.28**	2.43**	0.11**	–0.24**
	1/*N*	0.71	4.4	0.16	–0.04
N = 500					
	Mean–variance: λ = 20	0.88**	16.46**	0.05**	–0.51**
	Mean–variance: λ = 100	0.35**	3.83**	0.09**	–0.35**
	Treynor–Black†	0.68**	4.93**	0.14**	–0.14**
	DeMiguel	–14.08*	494.00**	0.02**	–0.65**
	Minimum-variance	0.22**	1.61**	0.14**	–0.14**
	1/*N*	0.71	4.34	0.16	–0.03
B. Modest forecasting ability					
Results for an investor with medium forecasting ability (R^2 = 0.005, IC = 0.071)					
N = 10					
	Mean–variance: λ = 20	0.85**	5.25**	0.16**	–0.04**
	Mean–variance: λ = 100	0.6**	4.52**	0.14	–0.16
	Treynor–Black†	1.25**	10.81**	0.12**	–0.21**
	DeMiguel	32.27**	606.00**	0.05**	–0.50**
	Minimum-variance	0.54**	4.49**	0.12**	–0.21**
	1/*N*	0.70	5.05	0.14	–0.14
N = 100					
	Mean–variance: λ = 20	3.83**	9.69**	0.39**	0.95**
	Mean–variance: λ = 100	1.00**	3.17**	0.32**	0.63**
	Treynor–Black†	1.56**	6.35**	0.25**	0.33**
	DeMiguel	63.9**	855.00**	0.14**	–0.11**
	Minimum-variance	0.30**	2.45**	0.12**	–0.21**
	1/*N*	0.71	4.4	0.16	–0.04

(continued)

Table 4. Simulation Summary: The Impact of Forecasting Ability (continued)

	Model	Mean	Standard Deviation	Sharpe Ratio	M^2
N = 500					
	Mean–variance: λ = 20	17.98**	21.53**	0.84**	2.84**
	Mean–variance: λ = 100	3.78**	4.78**	0.79**	2.65**
	Treynor–Black†	1.86**	4.92**	0.38**	0.9**
	DeMiguel	125.00**	1,140.00**	0.28**	0.45**
	Minimum-variance	0.23**	1.62**	0.14**	–0.13**
	1/N	0.71	4.34	0.16	–0.03

Notes: This table presents results derived through simulation using the algorithm of Kane et al. (2010). Treynor–Black† refers to the Treynor–Black model with a margin requirement as in Kane et al. (2010), DeMiguel refers to the optimal weight relationship in DeMiguel et al. (2009), and 1/N refers to the equally weighted portfolio.

*Statistically significant at the 5% level.

**Statistically significant at the 1% level.

then, with the conclusions of DeMiguel et al. (2009), the findings imply that in the absence of any forecasting ability, investors are best off adopting the 1/N investment strategy.

The DeMiguel et al. (2009) mean–variance portfolios tend to perform poorly. As noted previously, the reason appears to relate to extreme weightings, which are likely to have resulted from the process by which the authors forced the asset weights to sum to unity.[17] Indeed, DeMiguel et al. noted that their mean–variance weight in the *international* data set ranged from –148,195% to +116,828%. Such extremes suggest that the empirical results of DeMiguel et al. could be driven by their portfolio construction method, as was also suggested by Kirby and Ostdiek (2012).

We also examined the relationship between Sharpe ratios and the size of the asset universe for the different strategies. For the 1/N strategy, the Sharpe ratio increases with N but plateaus quickly as diversification benefits are realized. While the Treynor–Black and minimum-variance portfolios also show improving Sharpe ratios, the 1/N strategy retains superior risk-adjusted performance. For the mean–variance strategy, the Sharpe ratio declines with N. Consistent with the analytical model, therefore, for low levels of forecasting ability, gains from forecasting are offset by increases in estimation error, and in the absence of skill, irrespective of universe size, investors are better off using the passive investment strategy.

We now consider an investor with modest forecasting ability, equating to an information coefficient of 0.07 (see Panel B of Table 4). Our mean–variance investor now performs at least as well as the 1/N investor for every universe size and level of risk aversion. The differences in Sharpe ratio are statistically significant in all but one case. Now, consistent with our analytical model, the Sharpe ratios of mean–variance optimization increase with the size of our investment universe. Even these low levels of forecasting ability are sufficient to overcome the additional estimation error induced by increasing the size of the universe. With 500 assets, the mean–variance Sharpe ratios are more than five times larger than those of the 1/N approach. The M^2s of the mean–variance portfolios are high, in excess of 2.5% per month. For context, note that Malkiel (2013) reported annual mutual fund fees averaging 0.9%.

Our results cast doubt on the conclusion of DeMiguel et al. (2009) that mean–variance is unworkable in higher dimensions. Consider that Grinold and Kahn (1999) developed a simple binary model that relates the information coefficient to the number of forecasts that are directionally correct (the "hit rate"). Our IC of 0.07 equates to a hit rate of just 53.5% a month, or an R^2 of 0.5%. That such modest levels of forecasting ability can generate meaningful gains in utility is remarkable. The uplift in utility is tangible, and it is a benefit that increases with the size of the asset universe.

To conclude this section, we revisit the impact of estimation error in the covariance matrix. We modeled two investors: One knows the true population covariance matrix, and the other, who has medium risk aversion and modest forecasting ability (IC = 0.07), must estimate it by using a single-index method. **Table 5** compares the results for these two investors. As can readily be seen, the Sharpe ratios, M^2s, and certainty equivalents are similar irrespective of whether the true covariance matrix is known. By implication, the adverse effects of estimation error are largely offset by the use of a reasonable factor model. This conclusion provides further support for our analytical results in which the covariance matrix was assumed to be known. The results make a strong case that mean–variance optimization performs better than the 1/N strategy in stock selection when the investor has modest forecasting ability.

The analytical and summary results presented so far suggest that substantial gains are available to an investor from using a mean–variance approach *provided that* the investor has some modest forecasting ability. These findings rest on some simplifying assumptions, however, that may be violated in practice. We have assumed that returns are independently and normally distributed, that forecasts are independent across stocks, and that trading costs are zero. Moreover, in the simulation results, our forecasts were generated by a single-index model that was aligned with our data generation process. Although these simplifications are valuable for modeling purposes, they may introduce a bias that favors the mean–variance approach. So, in the next section, we show results when we relaxed those assumptions by moving to an empirical setting where none of those properties could be assumed to hold. If the mean–variance approach still generates superior performance, the outcome will reinforce the benefits of this approach when investors do possess some forecasting ability.

Out-of-Sample Empirical Evaluation

We evaluated the performance of the mean–variance approach in an out-of-sample portfolio-rebalancing framework that was intended to replicate the problems faced by institutional fund managers. The most comprehensive work in this area is by DeMiguel et al. (2009), who evaluated 12 extensions to the mean–variance analysis designed to reduce estimation error for seven data sets. In five of those sets, they found that the 1/N approach generated higher Sharpe ratios and certainty equivalents than the mean–variance approach. Given their data span, these results cannot be dismissed lightly: DeMiguel et al. argued that some 6,000 months of history would be required for a mean–variance portfolio of 50 stocks to outperform the 1/N portfolio. Given our findings in the previous section, however, we contend that this result may have come from the particular portfolio allocation procedure DeMiguel et al. used.

In an international asset allocation context, Solnik (1993) has previously shown that the mean–variance

Table 5. Simulation Summary: Estimation Error and Forecasting

	Covariance Matrix	Mean	Standard Deviation	Sharpe Ratio	M^2	Certainty Equivalent	Maximum Gross Exposure
N = 10	Unknown	0.64	4.58	0.14	–0.13	0.11	1.48
	Known	0.63	4.38	0.14	–0.12	0.14	1.39
N = 50	Unknown	0.92	3.77	0.24	0.31	0.56	3.04
	Known	0.79	3.36	0.23	0.27	0.50	3.03
N = 100	Unknown	1.48	4.06	0.36	0.82	1.06	4.61
	Known	1.29	3.56	0.36	0.82	0.97	4.61
N = 500	Unknown	6.15	7.51	0.82	2.77	4.73	16.81
	Known	5.61	6.58	0.86	2.92	4.52	16.68

Notes: Simulation was based on the algorithm of Kane et al. (2010). The investor has medium forecasting ability (R^2 = 0.005, IC = 0.071). “Unknown” rows pertain to cases where the covariance was estimated from the single-index model with an estimation window of 60 months. “Known” rows pertain to cases where the investor knows the true population covariance matrix.

approach outperforms passive benchmarks if the forecasts are conditioned on fundamental variables. We extended this work by considering universes with some 1,500–3,000 stocks to evaluate per month and with covariance matrix estimation windows as short as 60 months (where estimation error issues are most acute). The aim was to reproduce the type of practical problems facing, for example, a small-capitalization manager with a Russell 2000 benchmark and securities with short trading histories. We used 25 years of data from across three regions—Asia ex Japan, Europe, and the United States.

In our setup, a hypothetical manager makes return forecasts at month-end, estimates the covariance matrix by using a single-index model, and derives optimal portfolio weights to rebalance the portfolio.[18] In previous sections, we endowed the investor with a given level of forecasting ability, but in this analysis, our investor develops forecasts by using a range of fundamental and price-based variables that have been documented as beneficial, prior to the out-of-sample tests. Because of prior results and the prominence of mean–variance as an active strategy in the finance industry, we focused this analysis on the mean–variance approach, comparing it with the 1/*N* approach and a minimum-variance strategy.

We estimated univariate models by regressing the local-currency excess returns against a range of variables that have been proposed in the literature. In this procedure, we followed Solnik (1993), who suggested that this approach is equivalent to using the currency-hedged risk premium with interest rate parity. We estimated separate models for the three regions and used an expanding window of data with an initial size of 120 months (1990–1999), which increased the difficulty of out-of-sample prediction. At the end of each month, we added a further month of data and reestimated the regression coefficients. As in Moskowitz, Ooi, and Pedersen (2012), data were stacked for all stocks and dates and a pooled panel regression was run in each region for each forecasting variable. The forecasting variables were winsorized at the 5% and 95% levels to mitigate the effect of erroneous data:

$$r_{i,t,local} = a_{T+1} + b_{T+1}X_{i,t} + e_{i,t}, \tag{6}$$

where $r_{i,t,local}$ is the excess stock return in local currency and $X_{i,t}$ is the forecasting variable.

We then used the coefficient from Equation 6 to generate a one-month-ahead forecast, $r^f_{i,T+1}$, for each stock in the investment universe. We followed Welch and Goyal (2008) in providing in- and out-of-sample correlations of our forecasting models. We estimated stock betas and idiosyncratic variances by using the trailing five years of daily local excess returns and used the FTSE World Europe, FTSE World Asia Pacific ex Japan, and FTSE USA indexes as our benchmark proxies.

Our data set begins January 1990 and ends December 2014. Our tests used the constituents of the S&P Broad Market indexes from the three regions. The tests included all publicly listed equities with float-adjusted market values in excess of $100 million and are free from survivorship bias. **Table 6** shows summary statistics for the out-of-sample test period from January 2000 to December 2014 for all stocks with at least 120 months of history. Returns were measured in local currency and include dividends; the values are consistent with prior literature.

We used seven individual forecasting variables, drawn from prior literature and conventionally defined.[19] Price momentum (PM) is defined as the 12-month price changes ending one period before estimation (see Moskowitz et al. 2012); earnings momentum (EM) is defined as the change in Institutional Brokers' Estimate System (I/B/E/S) consensus forecast earnings per share over the prior three months, divided by the current price; price reversal (PR) is captured from the price change one month previous; dividend yield (DY), earnings yield (EY), and book-to-market ratio (BM) are defined conventionally; finally, we defined return on equity (ROE) as the last reported earnings per share over the book value per share. All FactSet Fundamentals data were lagged by three months to eliminate any potential look-ahead bias.

For portfolio selection, the objective function seeks to maximize the expected return less risk aversion, multiplied by expected risk and trading costs with weights constrained:

$$\max_{w_t} \mathrm{U_p}(\mathbf{w}_t) = w_t'\mathbf{r}^f_{T+1} - \frac{\lambda}{2} w_t'\hat{\mathbf{\Sigma}}_{T+1}\mathbf{w}_t - \left|\mathbf{w}_t - w_{t,pre}\right|' \mathbf{TC} \tag{7}$$

subject to $\mathbf{w}_t' i = 1;\ \mathbf{w}_{min} \le \mathbf{w}_t \le \mathbf{w}_{max}$,

where

$\mathbf{w}_t$ = the vector of optimal weights

$\mathbf{r}^f_{T+1}$ = a vector of return forecasts

Table 6. Summary Statistics: Stock Returns, January 2000–December 2014

Area/Measure	Mean	Median	5th Percentile	95th Percentile
Asia Pacific ex Japan				
Mean	1.0	1.0	−0.1	2.4
Standard deviation	10.7	10.1	5.6	17.1
Skewness	0.4	0.2	−0.6	1.9
Kurtosis	7.1	5.1	3.0	15.5
Minimum	−33.4	−31.9	−55.3	−16.7
Maximum	45.4	37.3	17.5	92.8
Europe				
Mean	0.7	0.8	−0.7	1.8
Standard deviation	10.2	9.7	5.8	16.4
Skewness	0.2	0.1	−0.7	1.6
Kurtosis	6.6	5.1	3.3	14.1
Minimum	−33.7	−31.8	−59.2	−16.3
Maximum	42.5	35.8	17.4	90.1
United States				
Mean	1.0	1.0	−0.2	2.2
Standard deviation	12.4	11.3	6.1	21.9
Skewness	0.5	0.2	−0.6	2.3
Kurtosis	7.6	5.2	3.3	18.5
Minimum	−38.7	−36.7	−66.0	−18.2
Maximum	56.2	43.9	19.7	136.8

Notes: The table presents statistics based on monthly stock returns in local currency. The mean, standard deviation, minimum, and maximum are in percent.

λ = the coefficient of risk aversion

$\hat{\boldsymbol{\Sigma}}_{T+1}$ = the estimated single-index covariance matrix

$w_{t,pre}$ = the pre-rebalanced weights

TC = a vector of trading costs

$\mathbf{w}_{min}$ and $\mathbf{w}_{max}$ = minimum and maximum weight vectors

In fuller results, available from the authors, we used three levels of risk aversion to give a range of *ex post* portfolio variances. Here, we report the midrange risk aversion. We assumed a 0.5% transaction cost for all stocks both as an optimization input and to compute net returns—a choice consistent with DeMiguel et al. (2009), Balduzzi and Lynch (1999), and Kirby and Ostdiek (2012). Because we were using local-currency excess returns, we had no hedging costs to consider. To ensure that the optimal portfolios were feasible in practice, we constrained minimum and maximum weights to

$$w_{\min,i,t} = \frac{-ADV_{i,t}}{\text{Fund size}} \text{ and } \quad w_{\max,i,t} = \frac{ADV_{i,t}}{\text{Fund size}}, \tag{8}$$

where $ADV_{i,t}$ is the average daily volume of stock *i* over the previous 30 trading days at month *t*.

We used a fund size of $500 million at each point in time. This step ensured that the positions were feasible for a realistic institutional portfolio. Almost none of the existing literature accounts for market liquidity, but we believed this constraint was important to ensure that our results could be achievable in practice.

We used a number of metrics to evaluate portfolio performance. We calculated net-of-trading-costs mean excess return, the Sharpe ratio, the M^2, and the certainty equivalent. To test whether differences between the mean–variance and 1/*N* Sharpe ratios were statistically significant, we used the Jobson–Korkie (1981) *t*-statistic with Memmel's (2003) correction. Similarly, we used Greene (2002) to test for significant differences in the certainty equivalent.[20]

Table 7 provides coefficient estimates, *t*-statistics, and in- and out-of-sample correlations for the seven forecasting models for the three regions. With the exception of the book-to-market ratio in Europe and the price reversal variable in Asia, the signs of the coefficients are consistent with the literature and significant. The out-of-sample *R*-statistics are typically higher in Europe and Asia than in the United States, suggesting that the US equity market is more informationally efficient. The out-of-sample correlations between predicted and observed values are positive for 16 of the 21 cases. What is striking is how small the correlations are: If mean–variance optimization outperforms the 1/*N* strategy, that result would indicate that only a modest level of forecasting skill is required.

We show the results for a medium level of risk aversion for the three regions in **Table 8**. The Sharpe ratios in Panel A for the Asia Pacific ex Japan region are statistically larger for mean–variance optimization for all factors except the book-to-market ratio model, as are the M^2s and certainty equivalents. These results held for other levels of risk aversion

Table 7. Predictive Regression Summaries, January 2000–December 2014 (*t*-statistics in parentheses)

Area/Measure	DY	EY	BM	PM	EM	PR	ROE
Asia Pacific ex Japan							
Coefficient	0.43	0.37	0.00	0.45	0.55	0.01	0.51
	(18.44)	(15.37)	(0.15)	(19.68)	(22.55)	(0.38)	(21.23)
% Positive	100%	100%	27%	100%	100%	63%	100%
R in-sample	0.020	0.023	0.019	0.021	0.031	0.013	0.004
R out-of-sample	0.021	0.024	−0.022	0.024	0.030	−0.002	0.022
Europe							
Coefficient	0.13	0.35	−0.04	0.47	0.47	−0.17	0.36
	(10.35)	(26.07)	(−3.21)	(37.19)	(34.57)	(−13.88)	(28.42)
% Positive	98%	100%	66%	100%	100%	0%	100%
R in-sample	0.016	0.025	−0.008	0.033	0.028	−0.002	0.005
R out-of-sample	0.017	0.028	−0.008	0.030	0.030	−0.005	0.023
United States							
Coefficient	0.20	0.59	0.03	0.35	0.44	−0.23	0.53
	(16.69)	(48.46)	(2.53)	(29.81)	(34.93)	(−19.72)	(44.03)
% Positive	100%	100%	93%	100%	100%	0%	100%
R in-sample	0.004	0.013	0.004	0.009	0.008	−0.012	0.004
R out-of-sample	0.006	0.016	0.012	−0.005	0.004	0.010	0.010

Notes: DY = dividend yield; EY = earnings yield; BM = book-to-market ratio; PM = price momentum; EM = earnings momentum; PR = price reversal; ROE = return on equity. The in-sample *R* was given by the final regression for the period January 1990–December 2014. The out-of-sample *R* refers to the out-of-sample correlation of the forecasts with excess returns for the period January 2000–December 2014.

Table 8. Mean–Variance Approach vs. 1/*N* Approach: Summary Out-of-Sample Monthly Performance Statistics, January 2000–December 2014

Portfolio/Model	Gross Return	Net Return	Standard Deviation	Sharpe Ratio	M^2	Certainty Equivalent
A. Asia Pacific ex Japan						
Mean–variance						
EY	1.12	1.00	3.88	0.29*	0.50*	0.97*
DY	1.16	1.03	4.04	0.29*	0.50*	1.00*
BM	1.01	0.89	3.91	0.26	0.37	0.86*
PM	1.16	1.02	4.09	0.28*	0.48*	0.99*
PR	1.06	0.94	3.91	0.27*	0.42*	0.91*
EM	1.30	1.13	3.78	0.34**	0.74**	1.15**
ROE	1.12	0.99	4.02	0.28*	0.46*	0.96*
Minimum-variance	1.06	0.92	3.91	0.27*	0.43*	0.91*
1/*N*	0.55	0.49	5.48	0.09	−0.30	0.25
B. Europe						
Mean–variance						
EY	1.16	1.06	3.60	0.32**	0.64**	1.03**
DY	0.91	0.84	3.31	0.27*	0.44*	0.8*
BM	0.79	0.72	3.35	0.23	0.27*	0.67*
PM	1.11	0.97	4.96	0.22	0.23	0.86
PR	0.93	0.86	3.44	0.27*	0.42*	0.81
EM	1.51	1.33	4.18	0.36*	0.81*	1.34**
ROE	1.00	0.92	3.39	0.30*	0.53*	0.89*
Minimum-variance	0.9**	0.82**	3.33	0.27*	0.42*	0.79*
1/*N*	0.41	0.36	5.32	0.07	−0.40	0.13
C. United States						
Mean–variance						
EY	2.13	1.98	7.05	0.30	0.56	1.64
DY	1.10	1.05	4.70	0.23	0.26	0.87
BM	1.07	1.04	3.53	0.30	0.57	0.95*
PM	0.43	0.24	8.88	0.05	−0.52	−0.36
PR	1.33	1.22	4.70	0.28	0.48	1.11
EM	1.55	1.21	8.07	0.19	0.09	0.89
ROE	1.62	1.49	7.11	0.23	0.24	1.11
Minimum-variance	1.13	1.08	3.62	0.31	0.6	0.99*
1/*N*	0.73	0.67	6.22	0.11	−0.23	0.34

Notes: Results are for medium risk aversion (λ = 60); results for other levels of risk aversion are given in the appendixes in the online material, available at www.tandfonline.com/doi/suppl/10.1080/0015198X.2019.1600958. See definitions of terms in the notes to Table 7.

*Statistically significant at the 5% level.

**Statistically significant at the 1% level.

(unreported). The average Sharpe ratio for the mean–variance models is substantially higher than that for the 1/*N* strategy, and the standard deviation is lower, despite the higher gross exposure. Note also that the minimum-variance portfolio (with no forecast involved) outperformed the 1/*N* portfolio.

These findings are echoed in Panel B for the European region: significantly larger Sharpe ratios for mean–variance optimization for all models (and all levels of risk aversion). Typically, the Sharpe ratios are three to four times larger than that for the 1/*N* portfolio, and the other metrics confirm the outperformance of the forecast-based mean–variance strategies. As before, the 1/*N* strategy has a higher variance than any of the model-based mean–variance approaches.

The results in Panel C for the US models are similar in form, with the mean–variance models exhibiting higher Sharpe ratios (on average, double that of the 1/*N* portfolio), but few differences are statistically significant. Closer inspection reveals that this outcome is driven, in part, by the low correlation between the mean–variance and 1/*N* strategies, which led to a higher standard error for the difference in Sharpe ratios. Because the 1/*N* strategy maps to the market portfolio, the low correlation provides diversification benefits for the mean–variance strategies. As noted previously, the out-of-sample correlations of the US model forecasts with observed returns are substantially lower than those for the European and Asian models, and consistent with our analytical and simulation results, lower levels of forecasting ability translate into lower levels of utility. Nevertheless, across all seven models and for three different levels of risk aversion, mean–variance optimization delivered higher Sharpe ratios in 18 of 21 cases (with only the price momentum model delivering a lower figure). These results, confirming the analytical and simulation findings, stand in marked contrast to the DeMiguel et al. (2009) contention that 1/*N* strategies are superior to mean–variance optimization and point to substantial benefits from a forecast-driven mean–variance approach.

Conclusion

Much recent finance literature has contended that equal-weighted investment strategies are preferable to mean–variance optimization. The basis for this conclusion seems to be two strong assumptions: first, that the investor has no forecasting ability and, second, that estimation error in the covariance matrix is irreducible. Allowing for forecasting ability, however, is consistent with Markowitz's (1952, 1959) original formulation of portfolio theory and the extensive literature on capital market anomalies.

Our analytical results present a compelling case for the mean–variance portfolio approach and stand in contrast to De Miguel et al. (2009), who argued that vast amounts of data are required for mean–variance optimization to outperform an equal-weighting, 1/*N* approach. We showed here that only a modicum of forecasting ability is required for mean–variance optimization to outperform the 1/*N* approach. Indeed, in many of the investment problems we considered, mean–variance optimization was favored even in the absence of forecasting ability. We also showed that applying the DeMiguel et al. model in the most basic way produces results that favor mean–variance optimization.

In our simulation study of the stock selection problem, we showed that in the absence of forecasting ability, investor welfare decreases as the size of the investment universe increases and a 1/*N* approach should be preferred to mean–variance optimization. However, when we assumed only a modest level of forecasting ability, this conclusion was reversed: Investor welfare increases as the size of the investment universe grows, and mean–variance optimization tends to outperform the 1/*N* approach.

Our empirical findings support the analytical and simulation-based results. Even with short covariance matrix estimation windows of 60 months and using simple univariate forecasting models, mean–variance optimization outperforms 1/*N* across almost all models and regions, with the superior performance being statistically significant in many instances.

DeMiguel et al. (2009) concluded that we have "many miles to go" (p. 1) before the promised benefits of optimal portfolio choice can be realized out of sample. Our findings suggest that we may have already arrived.

Editor's Note

This article was externally reviewed using our double-blind peer-review process. When the article was accepted for publication, the authors thanked the reviewers in their acknowledgments. Nick Baltas was one of the reviewers for this article.

Submitted 8 June 2018

Accepted 20 March 2019 by Stephen J. Brown

Notes

1. See, for example, Amenc, Goltz, Le Sourd, and Martellini (2008).

2. Markowitz (1959) and Levy and Markowitz (1979) showed that quadratic approximation provides a reasonable and robust working assumption for a broad range of utility functions and return distributions. But in neither the 1959 nor the 1979 research is the assertion made that normality or quadratic utility holds, nor is either one a requirement for the model. See also Kritzman and Markowitz (2017). An alternative approach broadly within this framework would be to have regime-dependent risk aversion (see Chow, Jacquier, Kritzman, and Lowry 1999).

3. Note also that there are accepted Bayesian approaches to deal with parameter uncertainty.

4. Certainty equivalent is the amount of payoff that an agent would have to receive to be indifferent between that payoff and a given gamble.

5. Kritzman (2006) also demonstrated that when the assets are close substitutes, substantial misallocation induced by estimation error results in only relatively small changes in portfolio *ex ante* distributions and, hence, only a limited reduction in expected utility.

6. Significant contributions include Solnik (1993), Ulf and Maurer (2006), Campbell and Shiller (1988), Fama and French (1988), Lintner (1975), Fama and Schwert (1977), Campbell (1987), Poterba and Summers (1988), and Lo and MacKinlay (1988).

7. Among authors using a myopic agent (one that optimizes one-period-ahead utility) approach are Fleming, Kirby, and Ostdiek (2001, 2003); Jagannathan and Ma (2003); DeMiguel et al. (2009); and Kirby and Ostdiek (2012).

8. DeMiguel et al. (2009) divided the unconstrained mean–variance weights by their sum to ensure that the weights satisfied the budget constraint. Our results show that this seemingly innocuous approach has some unusual effects. We, instead, incorporated the budget constraint directly into the optimization problem by using the result given by Ingersoll (1987).

9. The online supplemental material is available at www.tandfonline.com/doi/suppl/10.1080/0015198X.2019.1600958.

10. We should sound a note of caution: These results hold when alpha, beta, and gamma are held constant. Adding assets to the investment universe will change the return and covariance matrixes, however, and marginal assets may become more prone to forecast errors.

11. Formally,

$$IC^* \equiv \inf\{IC : L_{mv}(w^*, \hat{w}) = L_{1/N}(w^*, w^{1/N})\},$$

where $L_y(w^*, \hat{w})$ is the expected utility loss of using weights $\hat{w}$ instead of the optimal weight strategy. Here, we set the expected utility loss of strategy y using weights $\hat{w}$ instead of optimal weights $\hat{w}^*$ equal for the different portfolios. We could equally well define IC^* in terms of expected utility.

12. Provided $V_{1/N}$ is positive.

13. See Appendix B and Appendix C in the online supplemental material, available at www.tandfonline.com/doi/suppl/10.1080/0015198X.2019.1600958, for more details and the full proof, respectively.

14. Kenneth French's data library is available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

15. As an illustration, the correlation between the "soda" and "gold" industries in the 48-industry classification is just 0.06; their parent sectors in the 5-industry classification have a correlation of 0.88.

16. M^2 converts the Sharpe ratio to a measure of the risk-adjusted return premium to the market, as proposed by Modigliani and Modigliani (1997). Mathematically, $M^2 = SR_p\sigma_m - \mu_m$, where SR_p is the portfolio Sharpe ratio, μ_m is the mean market return, and σ_m is the market standard deviation. We set μ_m = 0.73% and σ_m = 4.27%, as in Kane et al. (2010).

17. DeMiguel et al. (2009) forced the asset weights to sum to unity by dividing the unconstrained mean–variance weights by the absolute value of the sum of the weights, as follows:

$$w = \frac{\Sigma^{*-1}\mu^*}{\left[\left|i'\Sigma^{*-1}\mu^*\right|\right]}.$$

18. We acknowledge that superior estimates of risk can be attained by using factor models, such as the Fama and French (1993) three-factor model (rarely used by practitioners), macroeconomic risk models, or statistical factor models. Fundamental factor models, such as those developed by Barra and Axioma, are routinely used by practitioners to reduce estimation error. The results that we present based on a one-factor model (if they support the mean–variance approach) should, therefore, be conservative. We note also that our models involve calculating a matrix inverse and that the work of Fan, Fan, and Lv (2008) implies that, relative to a *k*-factor model, we have incurred large losses in accuracy in this context. Rather than choosing a *k*-factor model among many possibilities, we chose the simplest and most parsimonious model to demonstrate the impact of forecasting ability. A reviewer has pointed out, however, that our single-factor specification may be helping improve performance of the mean–variance approach by reducing estimation error in the covariance matrix; we acknowledge this possibility.

19. Full data sources and prior literature supporting the inclusion of these variables are available in the appendixes in the online material (available at www.tandfonline.com/doi/suppl/10.1080/0015198X.2019.1600958).

20. We acknowledge that improved tests of Sharpe ratios could be accomplished by using bootstrapping procedures, as in Ledoit and Wolf (2008), but these procedures involve some implementation costs.

References

Amenc, N., F. Goltz, V. Le Sourd, and L. Martellini. 2008. *EDHEC European Investment Practices Survey*. Nice, France: EDHEC Business School.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens. 2001. "The Distribution of Realised Stock Return Volatility." *Journal of Financial Economics* 61 (1): 43–76.

Balduzzi, P., and A.W. Lynch. 1999. "Transaction Costs and Predictability: Some Utility Cost Calculations." *Journal of Financial Economics* 52 (1): 47–78.

Ball, R. 1978. "Anomalies in Relationships between Securities' Yields and Yield-Surrogates." *Journal of Financial Economics* 6 (2–3): 103–26.

Benartzi, S., and R.H. Thaler. 2001. "Naive Diversification Strategies in Defined Contribution Saving Plans." *American Economic Review* 91 (1): 79–98.

Black, F., and R. Litterman. 1991. "Asset Allocation Combining Investor Views with Market Equilibrium." *Journal of Fixed Income* 1 (2): 7–18.

Campbell, J.Y. 1987. "Stock Returns and the Term Structure." *Journal of Financial Economics* 18 (2): 373–99.

Campbell, J.Y., and R.J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1 (3): 195–228.

Chow, G., E. Jacquier, M. Kritzman, and K. Lowry. 1999. "Optimal Portfolios in Good Times and Bad." *Financial Analysts Journal* 55 (3): 65–73.

Cochran, W.G. 1934. "The Distribution of Quadratic Forms in a Normal System, with Applications to the Analysis of Covariance." *Mathematical Proceedings of the Cambridge Philosophical Society* 30 (2): 178–84.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. "Optimal versus Naive Diversification: How Inefficient Is the 1/*N* Portfolio Strategy?" *Review of Financial Studies* 22 (5): 1915–53. http://faculty.london.edu/avmiguel/DeMiguel-Garlappi-Uppal-RFS.pdf.

Fama, E.F., and K.R. French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics* 22 (1): 3–25.

———. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.

Fama, E.F., and G. Schwert. 1977. "Asset Returns and Inflation." *Journal of Financial Economics* 5 (2): 115–46.

Fan, J., Y. Fan, and J. Lv. 2008. "High Dimensional Covariance Matrix Estimation Using a Factor Model." *Journal of Econometrics* 147 (1): 186–97.

Fisher, K., and M. Statman. 1997. "The Mean-Optimization Puzzle: Security Portfolios and Food Portfolios." *Financial Analysts Journal* 53 (4): 41–50.

Fleming, J., C. Kirby, and B. Ostdiek. 2001. "The Economic Value of Volatility Timing." *Journal of Finance* 56 (1): 329–52.

———. 2003. "The Economic Value of Volatility Timing Using 'Realised' Volatility." *Journal of Financial Economics* 67 (3): 473–509.

Frost, P.A., and J.E. Savarino. 1988. "For Better Performance—Constrain Portfolio Weights." *Journal of Portfolio Management* 15 (1): 29–34.

Greene, W.H. 2002. *Econometric Analysis*. Fifth edition. Upper Saddle River, NJ: Prentice Hall.

Grinold, R.C. 1989. "The Fundamental Law of Active Management." *Journal of Portfolio Management* 15 (3): 30–37.

Grinold, R.C., and R. Kahn. 1999. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Selecting Superior Money Managers*. Second edition. New York: McGraw-Hill.

Hansen, P.R., and A. Lunde. 2006. "Realized Variance and Market Microstructure Noise." *Journal of Business & Economic Statistics* 24 (2): 127–61.

Ingersoll, J.E. 1987. *Theory of Financial Decision Making*. Totowa, NJ: Rowman & Littlefield.

Jagannathan, R., and T.S. Ma. 2003. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *Journal of Finance* 58 (4): 1651–83.

Jobson, J.D., and B.M. Korkie. 1981. "Performance Hypothesis Testing with the Sharpe and Treynor Measures." *Journal of Finance* 36 (4): 889–908.

Kan, R., and G. Zhou. 2007. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis* 42 (3): 621–56.

Kane, A., T. Kim, and H. White. 2010. "Forecast Precision and Portfolio Performance." *Journal of Financial Econometrics* 8 (3): 265–304.

Kirby, C., and B. Ostdiek. 2012. "It's All in the Timing: Simple Active Portfolio Strategies That Outperform Naive Diversification." *Journal of Financial and Quantitative Analysis* 47 (2): 437–67.

Kritzman, M. 2006. "Are Optimizers Error Maximizers?" *Journal of Portfolio Management* 32 (4): 66–69.

Kritzman, M., and H. Markowitz. 2017. "An Interview with Nobel Laureate Harry M. Markowitz." *Financial Analysts Journal* 73 (4): 16–21.

Kritzman, M., S. Page, and D. Turkington. 2010. "In Defense of Optimization: The Fallacy of 1/*N*." *Financial Analysts Journal* 66 (2): 31–39.

Ledoit, O., and M. Wolf. 2008. "Robust Performance Hypothesis Testing with the Sharpe Ratio." *Journal of Empirical Finance* 15 (5): 850–59.

Levy, H., and H.M. Markowitz. 1979. “Approximating Expected Utility by a Function of Mean and Variance.” *American Economic Review* 69 (3): 308–17.

Lintner, J. 1975. “Inflation and Security Returns.” *Journal of Finance* 30 (2): 259–80.

Lo, A.W., and A.C. MacKinlay. 1988. “Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test.” *Review of Financial Studies* 1 (1): 41–66.

Malkiel, B.G. 2013. “Asset Management Fees and the Growth of Finance.” *Journal of Economic Perspectives* 27 (2): 97–108.

Markowitz, H.M. 1952. “Portfolio Selection.” *Journal of Finance* 7 (1): 77–91.

———. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.

———. 2010. “Portfolio Theory: As I Still See It.” *Annual Review of Financial Economics* 2: 1–23.

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

Michaud, R. 1989. “The Markowitz Optimization Enigma: Is ‘Optimized’ Optimal?” *Financial Analysts Journal* 45 (1): 31–42.

———. 1998. *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. Boston: Harvard Business School Press.

Modigliani, F., and L. Modigliani. 1997. “Risk-Adjusted Performance.” *Journal of Portfolio Management* 23 (2): 45–54.

Moskowitz, T.J., Y.H. Ooi, and L.H. Pedersen. 2012. “Time Series Momentum.” *Journal of Financial Economics* 104 (2): 228–50.

Petajisto, A. 2013. “Active Share and Mutual Fund Performance.” *Financial Analysts Journal* 69 (4): 73–93.

Poterba, J.M., and L.H. Summers. 1988. “Mean Reversion in Stock Prices: Evidence and Implications.” *Journal of Financial Economics* 22 (1): 27–59.

Rozeff, M.S. 1984. “Dividend Yields Are Equity Risk Premiums.” *Journal of Portfolio Management* 11 (1): 68–75.

Scherer, B. 2002. “Portfolio Resampling: Review and Critique.” *Financial Analysts Journal* 58 (6): 98–109.

Sharpe, W.F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance* 19 (3): 425–42.

Shiller, R.J. 1984. “Stock Prices and Social Dynamics.” *Brookings Papers on Economic Activity* 1984 (2): 457–510.

Solnik, B. 1993. “The Performance of International Asset Allocation Strategies Using Conditioning Information.” *Journal of Empirical Finance* 1 (1): 33–55.

Treynor, J.L., and F. Black. 1973. “How to Use Security Analysis to Improve Portfolio Selection.” *Journal of Business* 46 (1): 66–86.

Ulf, Herold, and R. Maurer. 2006. “Portfolio Choice and Estimation Risk. A Comparison of Bayesian to Heuristic Approaches.” *ASTIN Bulletin* 36 (1): 135–60.

Welch, I., and A. Goyal. 2008. “A Comprehensive Look at the Empirical Performance of Equity Premium Prediction.” *Review of Financial Studies* 21 (4): 1455–1508.

Zhou, G. 2008. “On the Fundamental Law of Active Portfolio Management: What Happens If Our Estimates Are Wrong?” *Journal of Portfolio Management* 34 (4): 26–33.

Zweig, J. 1998. “Five Investing Lessons from America’s Top Pension Fund.” *Money Magazine* (January).