



UNIVERSITY OF
CAMBRIDGE

Cambridge Working Papers in Economics

Mean-Variance versus $1/N$:
What if we can forecast?

*David Allen, Colin Lizieri
and Stephen Satchell*

CWPE 1244

PRELIMINARY: NOT FOR QUOTATION PLEASE

Mean-Variance versus 1/N: What if we can forecast?

DAVID ALLEN[†], COLIN LIZIERI and STEPHEN SATCHELL

December 22nd, 2013

ABSTRACT

Mean-variance optimisation has been roundly criticised by financial economists and practitioners alike, leading many to advocate a simple 1/N weighting heuristic. We examine the performance of the mean-variance technique conditional on investor forecasting ability and in the presence of estimation error. Using a novel analytical approach, we demonstrate that investors with a modicum of forecasting ability can employ the mean-variance approach to significantly increase their ex ante utility, outperforming the 1/N rule. We also provide the critical amount of forecasting ability required to outperform 1/N. Our results call into question the finding that there are “many miles to go” before the benefits of optimal portfolio choice can be realised. Instead we show that we may have already arrived.

I. Introduction

We investigate the performance of the mean-variance approach when investors have forecasting ability and in the presence of estimation error. In the seminal *Portfolio Selection*, Markowitz (1952) proposes mean-variance optimisation as the normative method for allocating capital to risky assets by maximising the expected return for a given level of variance. The Markowitz approach was the basis for several important advances in financial economics including the Capital Asset Pricing Model (Sharpe, 1964 and Lintner, 1965) and the understanding of the dichotomy between systematic and diversifiable risk. A considerable controversy has raged over the efficacy of the technique ever since. Two key criticisms have been levelled at the mean-variance approach. First it is often incorrectly stated that the mean-variance approach rests on the assumption that either returns are normally distributed or that investors have quadratic utility, neither of which assumptions hold in practice.^{1,2} A close reading of Markowitz (1952, 1959) reveals that no such assumption was made³. The second criticism is the phenomenon known as error-maximisation where return and covariance forecast errors are magnified in the estimated portfolio weights leading to poor out-of-sample performance. We explore the second criticism, error maximisation and the role of forecasting

¹ It has been known since Mandelbrot (1963) that the return distributions of financial assets are fat tailed relative to the normal distribution (leptokurtic) and have more extreme negative values than positive values (negatively skewed)

² The quadratic utility function has an increasing level of absolute risk aversion. Hence in a two asset framework an increase in wealth will lead to a decrease in the allocation to the risky asset. In some circumstances it is also possible for the mean-variance investor to strictly prefer less to more.

³ At no stage did Markowitz (1959) ever maintain that agents have quadratic utility. Rather Markowitz (1959) and later Levy and Markowitz (1979) show that the quadratic approximation provides a reasonable working assumption for a surprisingly large range of utility functions. This finding has been widely replicated (Pulley, 1981, Kroll, Levy and Markowitz, 1984, and Simaan, 1993). For instance, in the case of the logarithmic utility function, Markowitz (1959) shows the quadratic approximation is quite robust in the range of -30% to 40%. Levy and Markowitz (1979) show that the correlation between logarithmic and mean-variance utility approaches unity for a sample of 149 mutual funds, while the correlation between exponential and mean-variance utility is very high for all but the highest level of risk aversion. Nor did Markowitz assume that asset returns follow a Gaussian distribution. On the one occasion where Markowitz investigated the return-generating distribution he concluded that the daily returns of the S&P 500 student-t distribution were most likely generated by the student-t distribution with between four and five degrees of freedom (Markowitz and Usmen, 1996a, 1996b).

ability. Estimation error and forecasting ability can be seen as two sides of the same coin. If we have perfect foresight, then we have no need for estimation, and hence no estimation error.

The analytical approach that we follow provides two advantages. First, unlike an empirical approach the conclusions of an analytical approach are not driven by the vagaries of the sample period. Second, an analytical approach lays bare the drivers of expected utility, whereas in an empirical study the key levers need to be inferred. In this paper, we make several contributions to the literature. First we unify the effect of forecasting and estimation error in a coherent utility maximisation framework. Second, we provide the critical amount of skill and estimation data required by a mean-variance investor to outperform $1/N$ ex-ante by incorporating the budget constraint. As we show, the Kan and Zhou (2007) conclusion that increasing the number of assets leads to a deterioration in performance does not necessarily hold once we allow for forecasting. Third, we show that for typical investment problems only a modicum of forecasting ability is required to outperform the much vaunted $1/N$ rule. Fourth, we show that estimation error in the covariance matrix does not need to detract from expected utility. Fifth, we show that the most basic application of the influential DeMiguel, Garlappi and Uppal (2009) model actually supports the use of mean-variance over the $1/N$ rule. Sixth we show in an out-of-sample dynamic portfolio rebalancing framework that mean-variance outperforms $1/N$ on average across a range of data sets. The contrasting DeMiguel et al. (2009) results appear to be driven by an idiosyncrasy in the way the authors estimated the optimal weights.

II. Literature Survey

By construction, the mean-variance approach will overweight those assets that have large estimated returns, low correlations, and small variances, i.e. precisely the assets that are most likely to have large estimation errors. Michaud (1989) coined the term “error maximisation” for this phenomenon. Best and Grauer (1991) show that a surprisingly small increase in the expected return of a single asset can drive half of the securities from the portfolio. Further, if two assets are highly correlated, with similar but not identical expected returns the optimisation algorithm will tend to take long positions in the asset with the higher return and short positions in the asset with the lower return even if the difference in expected return is within the range of statistical error (Scherer, 2002). The result is unintuitive highly leveraged portfolios. Indeed, numerous studies show that the mean-variance technique does not outperform the $1/N$ rule out-of-sample. Bloomfield, Leftwich and Long (1977) demonstrate that the mean-variance approach does not outperform $1/N$ out-of-sample for a selection of fifty equities. Jobson and Korkie (1981) demonstrate that the $1/N$ portfolio outperforms mean-variance using Monte Carlo simulation. Jorion (1991) shows that the $1/N$ and value weighted portfolios generate higher Sharpe ratios than the mean-variance portfolio using U.S. industry portfolios. Jagannathan and Ma (2003) show that the mean variance approach is inferior to both minimum variance and the $1/N$ portfolio for the S&P 500 investment universe. DeMiguel, Garlappi and Uppal (2009) utilising fourteen portfolio construction models and seven datasets, find that “none is consistently better than the $1/N$ rule in terms of Sharpe ratio, certainty equivalent, or turnover.” DeMiguel et al. (2009) conclude that there are “many miles” to go before the gains promised by optimal portfolio-choice can actually be realised out of sample.⁴ The DeMiguel (2009) et al. findings are all the more striking given that majority of the mean-variance models were designed to mitigate the effect of estimation error.

⁴ The strong performance of the $1/N$ rule provided the impetus for Tu and Zhou (2011) to investigate combining the weights derived through Bayesian approaches with the $1/N$ weights. Tu and Zhou (2011) find that combining portfolio rules outperforms both mean-variance and $1/N$ in the majority of scenarios. In a related piece of work, Paye (2010) argues that combining estimators of portfolio weights is itself equivalent to an asset allocation problem. Paye (2010)

Considering the prominent role the mean-variance model plays in the MBA programs, the empirical literature hardly provides a ringing endorsement of the approach. It is noteworthy that a unifying feature of the negative analyses⁵ is the exclusive reliance on sample estimates for forecasting expected returns. Further, short histories are almost always employed. Bloomfield, Leftwich and Long (1977) use a 30 month window to estimate expected returns; Jorion (1991), and Jagannathan and Ma (2003) use 60 month windows; Duchin and Levy (2009) use 60 months of data while DeMiguel et al. (2009) use 60 and 120 month windows. The arbitrary reliance on short time periods is surprising given that the predictive power of sample moments in particular is known to be poor. Jobson and Korkie (1980) establish the asymptotic properties of the sample mean and covariance matrix and conclude that the estimators of μ and σ are inappropriate for making inferences out of sample⁶. In the same vein, Jorion (1985) evaluates the predictive validity of sample return estimates and concludes that:

“estimation risk due to uncertain mean returns has a considerable impact on optimal portfolio selection and that alternative estimators for expected returns should be explored”

The use of short sample histories also creates problems for the estimation of the covariance matrix. When the number of assets, N , exceeds the number of periods in the estimation window, T , the covariance matrix is not invertible. If the ratio of T/N is less than one but not negligible, the covariance matrix is invertible, enabling the calculation of optimal portfolio weights, however the matrix is ill-conditioned leading to a dramatic amplification of estimation error (Ledoit, 2004).

In *Against the Gods*, Bernstein describes this tendency of individuals to employ patently dubious data, recounting a story from Kenneth Arrow who served as a statistician in the air force in World War II. Several officers had been assigned the task of forecasting weather conditions one month ahead. When Arrow found that their forecasts were worthless, the officers petitioned to be relieved of this duty. The response was: “The Commanding General is well aware that the forecasts are no good. However, he needs them for planning purposes”. In light of the long-standing evidence indicating that sample estimates are generally make poor forecasts a cynic might argue that much of the existing literature is guilty of a straw-man argument. The mean-variance model is no different from any other financial model. Sensible inputs are required to generate sensible output. We would not expect for the Black-Scholes model to give sensible prices if we were not using reasonable estimates of volatility.

Kritzman, Page and Turkington (2010) show that even with simple estimates for expected returns and covariance, such as long-term historical averages, mean-variance optimisation generates higher out-of-sample Sharpe ratios than the $1/N$ heuristic using thirteen data sets. Interestingly Kritzman (2006) does not take issue with the error-maximisation charge. Rather, Kritzman (2006) shows that even substantial misallocations induced by estimation error lead to only a small change in the portfolio ex ante return distribution and hence only a small reduction in expected utility.

The extrapolation of sample moments is particularly puzzling given Markowitz (1952) never advocated the use of sample estimates as inputs to the optimisation algorithm. Given the volume of work that almost universally employs noisy sample estimates it is easy to forget that this practice is a departure from the original theory. Indeed, Markowitz proposed the “E-V” rule where the “E” signifies expected return and not

finds this form of shrinkage applied to an ultra-conservative strategy and the non-parametric estimator of Brandt (1999), leads to an improvement over the Brandt (1999) approach used in isolation

⁵ Excluding Kritzman, Page and Turkington (2010)

⁶ Merton (1980) also demonstrated that we need a very long data history to estimate expected returns with any precision.

the “M-V” rule where “M” signifies the mean. Markowitz (1952) is explicit on this point from the opening paragraph to the last. He states:

“to use the M-V rule in the selection of securities, we must have procedures for finding reasonable μ and σ . These procedures, I believe should combine statistical techniques and the judgement of practical men.”⁷

Markowitz has remained steadfast in this position ever since he penned *Portfolio Selection*. Recently Markowitz (2010) attempted to clear up this misapplication of the model emphasising the role of forecasting.

“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.”

Allowing for forecasting is rooted in the voluminous “anomalies” literature. Indeed there is a growing acceptance, even amongst devout proponents of efficient markets⁸ that asset returns are not purely random. The predictive power of scaled price-ratios dates back to Graham (1949). Lakonishok, Schleifer and Vishny (1993) demonstrate that low price-earnings stocks outperform high price-earnings stocks by 10% per year. Banz (1981) shows that the smallest 50 stocks outperformed the largest 50 stocks by 12% per year. DeBondt and Thaler (1985) find that the long-term (three to five years) “loser” portfolio outperforms the long-term “winner” portfolio by 8% per year. Jegadeesh and Titman (1993) find that medium term (6-12 months) “winners” outperform medium term “losers” by 10% per year. Asness, Moskowitz and Pedersen (2009) provide compelling evidence of predictability for commodities, government bonds, and stock indices. The debate has largely shifted from whether or not predictability exists to the explanation behind the predictability. On the one hand behavioural theorists argue that cognitive and behavioural biases together with limits to arbitrage are behind the persistency of the so-called anomalies. On the other hand, efficient market theorists argue that the return premia represent compensation for bearing systematic risk. We do not attempt to resolve this debate here; rather, we assume that predictability may exist. The basis for this predictability in our models may be public information sources, such as those mentioned above, or private information.

Several empirical studies have looked at the performance of the mean-variance approach allowing for forecasting in line with Markowitz (1952). Under this approach, one-period ahead return forecasts are used as inputs to the mean-variance algorithm instead of the sample estimates. Solnik (1993) finds that the mean-variance approach using forecasts based on fundamental variables including the dividend yield, the risk-free-rate, and long-term bond yield outperforms passive market benchmarks for an international equity and bond allocation problem. Klemkosky and Bharati (1995) show that commonly used financial variables can be used as inputs to form profitable mean-variance portfolios. Fletcher (1997) finds that the mean-variance approach conditioned on lagged excess returns, the risk-free rate, and the dividend yield consistently outperforms passive benchmarks using U.K. industry portfolios. Marquering and Verbeek (2004) document economically significant profits from employing mean-variance with linear return prediction models for the S&P 500. Herold and Maurer (2006) show that conditioning on ten macroeconomic and valuation variables the mean-variance approach outperforms passive benchmarks using ten Euroland sector portfolios.

⁷ And of course women

⁸ for example Fama and French (1992)

There are three broad strands of research that examine the performance of the mean-variance approach in an analytical framework as we do here. This body of literature develops expressions for the expected utility of investors subject to a range of conditions. The first strand of literature examines the expected utility of the mean-variance approach assuming that agents cannot forecast and are subject to estimation error. Kan and Zhou (2007) provide the expected loss of using the sample mean return and covariance estimates as inputs for mean-variance analysis. The Kan and Zhou (2007) result forms the basis for the clearest antecedent of the current work, DeMiguel, Garlappi, and Uppal (2009). DeMiguel et al. (2009) consider three cases, where the mean is unknown and the covariance is known, where the mean is known and the covariance is unknown, and where the mean and the covariance are both unknown. The authors provide the conditions for mean-variance to outperform 1/N in each case.

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:

$$S_{mv}^2 - S_{ew}^2 - \frac{N}{T} > 0$$

where S_{mv}^2 and S_{ew}^2 are the squared in-sample Sharpe-ratios of the mean-variance and 1/N portfolios, N is the number of assets, and T is the length of the estimation window

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:

$$kS_{mv}^2 - S_{ew}^2 - \frac{N}{T} > 0$$

where

$$k = \left(\frac{T}{T - N - 2} \right) \left(2 - \frac{T(T - 2)}{(T - N - 1)(T - N - 4)} \right)$$

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:

$$kS_{mv}^2 - S_{ew}^2 - h > 0$$

where

$$h = \frac{NT(T - 2)}{(T - N - 1)(T - N - 2)(T - N - 4)}$$

Assuming the mean and covariance matrix are both unknown, and employing U.S. stock market data to calibrate the model, DeMiguel et al. (2009) show that for a portfolio containing 25 assets, 3000 months of estimation data is required for mean-variance to outperform 1/N. If the portfolio contains 50 assets, 6000 months are required. If the portfolio contains 100 assets, certainly not an unusual size for an institutional portfolio, over 12000 months, some 1000 years of data are required for mean-variance to outperform 1/N. Given that the origins of the New York Stock Exchange only date back to May 17th, 1792, this is something of a problem.

In the second strand of research, Grinold (1989), Grinold and Kahn (1999) and Campbell and Thompson (2008) allow for forecasting and ignore the effect of estimation error. Grinold (1989) relates the information ratio, defined as the active return divided by the active risk, to the level of forecasting ability as follows

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

where IR is the information ratio, IC is the information coefficient, defined as the correlation between forecast and realised returns, and N is the number of independent “bets” per year. The expected utility, U^* , of the active investor, is closely related to the information ratio and is given by $U^* = IR^2/4$.⁹ The so called “fundamental law of active management” is widely used by practitioners as both a conceptual tool and a means of evaluating potential investment strategies. Grinold and Kahn (1999) conclude:

“It takes only a modest amount of skill to win as long as that skill is deployed frequently and across a large number of stocks.”

The Grinold (1989) conclusion is analogous to the behaviour of a casino exploiting a small statistical edge numerous times. The result that increasing the number of assets increases investor welfare is diametrically opposed to the DeMiguel et al. (2009) finding. The reason lies in the different assumptions underlying the two models. DeMiguel et al. (2009) assume that there is no such thing as forecasting ability, and incorporate the effect of estimation error; hence, increasing the number of assets increases the amount of estimation error and decreases expected utility. In the case of the covariance matrix the number of terms that need to be estimated increases exponentially with the number of assets. Grinold and Kahn (1999) allow for forecasting ability, but do not account for estimation error; thus, increasing the number of assets increases the amount of breadth and expected utility. In the first chapter of the thesis we draw together these two strands of the literature, accounting for estimation error and forecasting ability in a single analytical framework. In this way we can reconcile the conclusions of the two schools.

A common thread in both the DeMiguel et al. (2009) and Grinold and Kahn (1999) models is the absence of a budget constraint. While keeping the maths simple, the two models are arguably only relevant for unconstrained agents that do not have a set level of gross exposure, such as a hedge fund manager. In contrast, a pension fund manager typically has a gross and net exposure of 100%. Further, comparing an unconstrained mean-variance portfolio with $1/N$ that by construction always satisfies a unit budget constraint creates an incongruity. The mean-variance approach tends to generate heavily leveraged portfolios with non-unit net exposure rendering any comparison with the $1/N$ rule questionable.

In the third strand of analytical research attempts have been made to characterise the expected utility of the mean-variance investor when returns depart from normality. Due to the difficulty of deriving closed form solutions for expected utility when returns depart from normality, there has not been a great deal of work done in this area. Simaan (1993) shows that the mean-variance assumption leads to a negligible reduction in investor welfare when returns are skewed and there is a riskless asset using a Pearson Type III distribution. Das and Uppal (2004) using a jump-diffusion approach find that the effect of systemic risk is small unless risk aversion is very low and leverage is high. The empirical work of Fleming, Kirby and Ost diek (2001) suggests that accounting for stochastic volatility leads to substantial gains in investor welfare. We utilise an analytical approach to trace the source of these gains and more generally to evaluate the gains from accounting for non-normality.

We proceed as follows. In section III we discuss the approach for estimating expected utility conditional on forecasting ability. In section IV, we derive expected mean-variance utility in the presence of forecasting ability and estimation error and in the absence of a budget constraint; in section V we incorporate the budget constraint in our calculations. In section VI we apply our propositions to a range of investment

⁹ See Grinold and Kahn (1999), chapter 5, p 124

problems, provide a comparison with the model of DeMiguel et al. (2009), and provide a dynamic portfolio rebalancing analysis. Section VII concludes.

III. Analytical Framework

The goal of our analytical approach is to provide a closed form expression for expected mean-variance utility conditional on forecasting ability and the length of the estimation window. Armed with this expression we can then provide the loss function for using mean-variance relative to 1/N. We can then estimate the required level of forecasting ability for mean-variance to outperform 1/N on average, and we can also estimate the probability that mean-variance outperforms 1/N for a given forecasting ability. We hope these insights prove valuable to practitioners in determining how to allocate capital and manage risk.

The 1/N benchmark is relevant for three reasons. First, the 1/N rule is not dependent on expected returns or covariances and is therefore devoid of estimation risk. Second, humans have an innate behavioural tendency to equal weight the options presented to them. Bernatzi and Thaler (2001) find that many investors equally weight the investment choices they are presented with¹⁰. It appears few are immune from this bias. Markowitz himself, when probed about how he allocated his retirement investments 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”¹¹.

Third, the 1/N rule is easy for investors to apply, and thus a viable alternative in practice.

We begin by formalising a metric for forecasting ability, d , as the correlation between forecast returns and realised returns. This of course is the familiar information coefficient. We use d to denote the information coefficient instead of IC , as the latter can create confusion with the identity matrix I , and the covariance vector C in our derivations. In the context of a regression of future returns on forecasts:

$$d = \frac{Cov(f_{t,t+1}, r_{t+1})}{Std(f_{t,t+1})Std(r_{t+1})} = \sqrt{\frac{SS_{exp}}{SS_{tot}}} = \sqrt{1 - \frac{SS_{err}}{SS_{tot}}} \quad (1)$$

where $f_{t,t+1}$ are the forecast returns, r_{t+1} are the realised returns, SS_{err} is the sum of the sum of squared errors, SS_{exp} , is the explained sum of squares, and SS_{tot} is the total sum of squares.

It can be seen that as forecasting ability increases, forecasting error, SS_{err} , decreases. When forecasting error is zero, $d=1$, the perfect foresight case. Grinold and Kahn (1999) show that for a typical investment problem a top-decile portfolio manager has an information coefficient of approximately 0.06. This lack of forecast accuracy is potentially problematic for the mean-variance approach because the optimisation algorithm takes the inputs as if they are known with certainty when in fact they are noisy estimates.

Empirically, we can measure forecasting ability as the sample value of d .

¹⁰ See Huberman and Jiang (2006) for a more recent example.

¹¹ Zweig (1998)

$$\hat{d} = \sqrt{\frac{\sum_{i=1}^n (\hat{f}_{t,t+1} - \bar{r})^2}{\sum_{i=1}^n (r_{t+1} - \bar{r})^2}} = \sqrt{\frac{\overline{SS}_{exp}}{\overline{SS}_{tot}}} = \sqrt{1 - \frac{\overline{SS}_{err}}{\overline{SS}_{tot}}} \quad (2)$$

We use the mean-variance utility function for several reasons. Numerous studies have found that expected utility obtained through direct optimisation on an ex-post basis is very similar to the expected utility of the mean-variance approach¹². Further, the mean-variance approach is widely applied by practitioners¹³ and is arguably the standard approach for asset allocation¹⁴. Fabozzi, Focardi and Jonas (2007) survey 38 medium and large-sized equity investment managers in North America and Europe totalling \$4.3 trillion in assets under management. The authors find that 97% of managers use variance to measure risk and that 83% of managers employ mean-variance optimisation. The EDHEC European Practices Survey, 2008, examines the investment behaviour of 229 investment managers including fund management firms, pension funds, private banks, investment banks, family officers and consultants. The survey finds that, 57% of investment managers use the absolute return variance as the risk objective when performing portfolio optimisation. For managers that set relative risk objectives, 76% use the tracking error, which is, the square-root of the variance of excess returns. In addition, the mean-variance utility function is the transform of the highly tractable exponential utility function with a well-established analytical apparatus that has proved useful for the development of equilibrium arguments and statements regarding expected utility¹⁵.

The loss function is interpreted in the literature as a function of ex-post expected utility and corresponds to the idea that if we used the particular portfolio weights, we would get a random variable ex-post, $U(\hat{x})$ and its mean, which would correspond to the mean of some sample, based on an i.i.d. distribution and fixed weights would necessarily be less than the optimal non-stochastic quantity $U(x^*)$. However, in such an i.i.d. world, both values could be interpreted as ex-ante expected utilities as long as the assumed distribution of the investor corresponds to nature's measure. This distinction is important because when we come to compare equal weights with forecasts, we can think of the problem as the ex-ante expected utility of two separate prospects exactly as one would in formulating preferences over lotteries.

DeMiguel et al. (2009) focus on the critical value T_{mv}^* defined as the sample size that is necessary for mean-variance to outperform 1/N on average:

$$T_{mv}^* \equiv \inf\{T: L_{MV}(w^*, \hat{w}) < L_{1/N}(w^*, w^{ew})\} \quad (3)$$

The use of T_{mv}^* to facilitate comparison between 1/N and mean-variance is consistent with the use of the sample mean as the input to the optimisation problem. As discussed in section II, the use of sample estimators as optimisation inputs is a particular application of mean-variance analysis, and differs both from the original theory and the behaviour of practitioners. Thus, instead of a critical T_{mv}^* , which frames the problem in terms of sample estimates, we set up a general framework where forecasting accuracy depends on the information coefficient, d : the correlation between forecast returns and realised returns. Forecasting ability reflects either the ability to manipulate public information in a superior way, access to privileged information, or both¹⁶.

¹² Levy and Markowitz (1979), Pulley (1981), Kroll, Levy, and Markowitz (1984), Simaan (1993), Cremers, Kritman and Page (2005)

¹³ Amenc, Goltz and Lioui (2011)

¹⁴ Black and Litterman (1991)

¹⁵ Further, it is the mean-variance approach that is used by Grinold (1989), Kan and Zhou (2008), and DeMiguel, Garlappi and Uppal (2009) the clearest antecedents of the current work

¹⁶ We do not exclude the case where the investor has savant capabilities or mystical powers of prediction!

We define the critical level of forecasting ability, d^* as the level of forecasting ability where the expected utility of mean-variance is equal to the expected utility of the 1/N rule. If $d < d^*$, then the investor would be better off employing 1/N.

$$d^* \equiv \inf\{d: L_{MV/a}(w, \hat{w}) < L_{1/N}(w^*, w^{ew})\} \quad (4)$$

While the mean-variance approach underpins the CAPM, the CAPM is also based on the strong assumptions of perfect information and market efficiency. If we do not maintain that markets are perfectly efficient, employing forecasts is entirely consistent with the original Markowitz doctrine. To ensure that we are setting up a like for like comparison of mean-variance and 1/N we explicitly incorporate the budget constraint.

Our analytical approach consists of three stages. Firstly we obtain solutions for expected utility conditional on forecasting and estimation error. We then solve for d^* by setting the expected utility of the mean-variance investor to the expected utility of the 1/N investor, or equivalently the loss function to zero. Finally, we provide the probability of mean-variance outperforming 1/N conditional on forecasting ability and estimation error.

IV. Derivation of Propositions: The Unconstrained Case

A. Introducing Forecasting Ability

We assume we have forecasts a ($M \times 1$), of returns r_{t+1} ($N \times 1$), a mean vector, μ ($N \times 1$), a covariance matrix, Σ ($N \times N$) and a covariance vector, C ($N \times 1$) between forecast and realised returns. The joint density function of a and r_{t+1} is:

$$\begin{pmatrix} r_{t+1} \\ a \end{pmatrix} \sim N \left[\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C' & \Omega \end{bmatrix} \right] \quad (5)$$

In DeMiguel, Garlappi, and Uppal (2009), $\mu = \bar{r}$, $C = 0$ and $\Omega = T^{-1}\Sigma$, under the assumption of i.i.d. returns. We instead allow for forecasts a with $C \neq 0$. Whereas Kan and Zhou (2007) and DeMiguel et al. (2009) take expectations over the true distribution of returns, we take expectations with respect to the conditional distribution of returns. In common with other work in this area (Grinold, 1989), we treat forecasts a as i.i.d. normally distributed variables with a mean of zero and a standard-deviation of one.

$$\begin{pmatrix} r_{t+1} \\ a \end{pmatrix} \sim N \left[\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C' & I_N \end{bmatrix} \right] \quad (6)$$

The conditional pdf of $r_{t+1}|a$ under these assumptions is given by

$$N(\mu + Ca, \Sigma - CC') \quad (7)$$

We use the conditional covariance because we can calculate it but other authors use the unconditional covariance matrix because it provides succinct, results. Grinold and Kahn (1999) ignore the effect of the reduction in variance due to the knowledge embedded in a by ignoring terms higher than order d^2 leading to an approximation which is similar to using the unconditional covariance.

As we show in appendix E, if portfolio returns are normal and utility is exponential then after several manipulations, transformed expected utility¹⁷ can be written as a simple function quadratic in asset weights, w ¹⁸:

$$E[U] = \omega' \mu - \frac{\lambda}{2} \omega' \Sigma \omega \quad (8)$$

where λ is the coefficient of risk aversion.

The usual calculation of expected loss is to replace unknown parameters by estimators of the portfolio weights and compute expected utility, taking expectations over the sample distribution. If we ignore the information in the sample data, and equally weight the assets we have

$$\omega = \frac{1}{N} \mathbf{1}$$

where $\mathbf{1}$ is a vector of ones

Expected utility for this case is given by

$$V_{1/N} = E[U_{1/N}] = \frac{\mathbf{1}' \mu}{N} - \frac{\lambda}{2N^2} \mathbf{1}' \Sigma \mathbf{1} \quad (9)$$

which is non-stochastic.

We now derive the expected utility of the mean-variance investors by treating a as a forecast of r_{t+1} . An alternative approach, akin to Kan and Zhou (2007), could be developed which treats a as an estimate of μ . We consider seven cases. We build these cases up in order of complexity, beginning with a single forecasting variable, no estimation error, and no budget constraint, and progressing to multiple forecasting variables, with estimation error and a budget constraint. Estimation error arises because agents do not know the true inputs to the utility function. The seven cases we consider are as follows:

1. *A single forecasting variable, no estimation error, no budget constraint*
2. *Multiple forecasting variables, no estimation error, no budget constraint*
3. *Multiple forecasting variables, constant forecasting ability, no estimation error, no budget constraint*
4. *Multiple forecasting variables, constant forecasting ability, estimation error, no budget constraint*
5. *Multiple forecasting variables, constant forecasting ability, no estimation error, budget constraint*
6. *Multiple forecasting variables, constant forecasting ability, estimation error, budget constraint*
7. *Multiple forecasting variables, constant forecasting ability, estimation error, budget constraint, assuming constant pairwise correlation and constant volatility*

The base case is well-known and we report the results; returning to our utility function given by (8), the first order conditions can be solved to give:

$$w = \frac{\Sigma^{-1} \mu}{\lambda} \quad (10)$$

Substituting (10) into (8) we see that expected utility is:

¹⁷ or equivalently the certainty equivalent

¹⁸ For the interested reader, we produce the derivation in appendix F

$$V = E[U] = \frac{\alpha}{2\lambda} \quad (11)$$

where $\alpha = \mu' \Sigma^{-1} \mu$ refers to the squared generalised Sharpe ratio.

B. A single forecasting variable, no estimation error, no budget constraint

We begin with the case of a single forecasting variable. Of course this has many applications, for example the single-factor model of Sharpe (1964). In line with the CAPM, each asset has a sensitivity to the single factor, a , which equates to our normalised forecast over the next period. We define conditional moments as:

$$\mu^* = \mu + C a \quad (12)$$

$$\Sigma^* = (\Sigma - C C') \quad (13)$$

$$C = \Sigma^{-\frac{1}{2}} D \quad (14)$$

where μ ($N \times 1$) is the vector of population means, Σ ($N \times N$) is the population covariance matrix, C ($N \times 1$) is the covariance between forecast returns and realised returns, D ($N \times M$) is the matrix of information coefficients, and a ($M \times 1$) is a vector of $N(0,1)$ variables. Whilst in this first case we only have one forecast variable per period ($M=1$), the vector D leads to a different forecast for each asset.

By substituting the conditional expectations (12) to (14) into (11) we solve for expected utility.

Proposition 1

The Unconditional Expected Utility of the Mean-Variance investor under the assumptions of a Single Forecasting Variable, a ($M=1$), with forecasting ability given by the vector D , no Estimation Error and in the absence of a Budget Constraint is given by

$$E[U] = \frac{1}{2\lambda} \left(\alpha + \frac{D'D + (\mu' \Sigma^{-\frac{1}{2}} D)^2}{1 - D'D} \right)$$

Proof: See Appendix A

The conditional expected utility with a single forecasting variable a ($M \times 1$) ignoring estimation error and without a budget constraint is given by

$$E[U|a] = \frac{1}{2\lambda} \left(a^2 \left(D'D + \frac{(D'D)^2}{1 - D'D} \right) + a \left(2\mu' \Sigma^{-\frac{1}{2}} D + \frac{2D'D D' \Sigma^{-\frac{1}{2}} \mu}{1 - D'D} \right) + \left(\alpha + \frac{(\mu' \Sigma^{-\frac{1}{2}} D)^2}{1 - D'D} \right) \right) \quad (15)$$

Proof: See Appendix A

We can now derive the unconditional probability of the utility, given by (15), of mean-variance exceeding the utility of the 1/N rule. From proposition 1, we see that expected utility is quadratic in a .

Hence we have

$$E[U|a] = \phi_2 a^2 + \phi_1 a + \phi_0$$

where

$$\phi_2 = \frac{D'D}{1 - D'D}$$

This is strictly positive as long as $D'D$ is bounded by 1.

$$\phi_1 = \frac{2\mu'\Sigma^{-\frac{1}{2}}d}{1 - D'D}$$

$$\phi_0 = \alpha + \frac{(\mu'\Sigma^{-\frac{1}{2}}D)^2}{1 - D'D}$$

We are interested in the unconditional expected utility of the optimised strategy exceeding the $1/N$ case. Thus, the following analysis is valid, where the probability is taken over the a distribution.

$$\begin{aligned} \text{Prob}(E[U|a] < E[V_{1/N}]) &= \text{Prob}(\phi_2 a^2 + \phi_1 a + \phi_0 < V_{1/N}) \\ &= \text{Prob}\left(a^2 + \frac{\phi_1}{\phi_2} a + \frac{\phi_0}{\phi_2} - \frac{V_{1/N}}{\phi_2} < 0\right) \end{aligned}$$

we see that,

Corollary 1

The Probability of Mean-Variance Outperforming the $1/N$ rule under the assumptions of a Single Forecasting Variable, a ($M=1$), no Estimation Error and in the absence of a Budget Constraint is given by

$$\text{Prob}(E[U|a] < E[V_{1/N}]) = \Phi(\lambda_1) - \Phi(\lambda_2)$$

Supposing $\lambda_1 > \lambda_2$ without loss of generality, where

$$\lambda_i = \left(-\phi_1 \pm \sqrt{\phi_1^2 - 4(\phi_0 - V_{1/N})\phi_2}\right) \frac{1}{2\phi_2}$$

where $\Phi(\cdot)$ is the standardised normal distribution function. Solving for λ_i yields between zero and two real-roots enabling us to calculate the probability of $1/N$ outperforming mean-variance. If $\phi_1^2 > 4(\phi_0 - V_{1/N})\phi_2$ then we have two real roots λ_1 and λ_2 and there is a non-zero probability of $1/N$ outperforming mean-variance. If $\phi_1^2 \leq 4(\phi_0 - V_{1/N})\phi_2$ then there is a zero probability of mean-variance outperforming mean-variance.

C. Multiple forecasting variables, no estimation error, no budget constraint

We now consider multiple forecasting variables. We can draw a parallel here with the arbitrage pricing theory of Ross (1976) where there are multiple sources of systematic risk and return premia or a quantitative investor employing a multifactor model.

In all of the cases that follow:

$$\mu^* = \mu + Ca$$

$$C = \Sigma^{\frac{1}{2}} D$$

where μ ($N \times 1$), C ($N \times N$), a ($N \times 1$), and D ($N \times N$) is diagonal. The significance of D being diagonal is that forecast $a_{i,t}$ is correlated only with asset return $r_{i,t+1}$, and not with asset return $r_{j,t+1}$ where $j \neq i$.

Proposition 2

The Unconditional Expected Utility of the Mean-Variance investor under the assumptions of Multiple Forecasting Variables, a ($M=N$), with a diagonal forecasting ability matrix, D , no Estimation Error and in the absence of a Budget Constraint is approximated by¹⁹

$$E[U] \cong \frac{\alpha + \text{tr}(D^2) + \text{tr}(D^4)}{2\lambda} + O(D^6)$$

Proof: See Appendix A

Again, we see that as the squared Sharpe ratio, α , increases, expected utility increases. Expected utility is also positively related to forecasting ability as anticipated. The expected utility always exceeds the expected utility in the absence of forecasting.

D. Multiple forecasting variables, constant forecasting ability, no estimation error, no budget constraint

We now turn to the case where we have a constant expected information coefficient, d , for all assets. Setting the forecasting ability to a constant across all assets is consistent with the literature (Grinold, 1989 and Williams and Satchell, 2011).

We assume that the information sources are orthogonal as follows:

$$D = \begin{bmatrix} d_1 & 0 & \dots \\ 0 & \ddots & \ddots \\ \vdots & \dots & d_n \end{bmatrix} \quad (16)$$

where we set $d_1 = d_2 \dots = d_n = d$. This can be summarised as $D = dI_N$

The conditional moments of r_{t+1} given a are now:

$$\mu^* = \mu + \Sigma^{\frac{1}{2}} da \quad (17)$$

$$\Sigma^* = \Sigma(1 - d^2) \quad (18)$$

Substituting (17) and (18) into (10) we have

Proposition 3

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , no estimation error and in the absence of a budget constraint using the conditional covariance Matrix is given by

¹⁹To an order of approximation $O(d^4)$

$$E[U] = \frac{(\alpha + d^2 N)}{2\lambda(1 - d^2)} \quad (19)$$

Proof: See Appendix A

The numerator is quadratic in forecasting ability indicating that as forecasting ability increases, expected utility increases at an increasing rate. Expected utility is positively related to the number of assets N . This finding makes sense in that as the investor's opportunity set expands she is able to capture incremental uncorrelated excess return streams and is reminiscent of the fundamental law of active management (Grinold, 1989). Grinold's (1989) fundamental law refers to the information ratio. We focus on utility. Information ratios can give misleading results under certain circumstances. For example an investment may have a very high information ratio, comprised of a low return and a very low risk level. If we cannot apply leverage, the benefit to the investor will be limited, despite the ostensibly attractive information ratio. Utility functions do not suffer from this drawback. A further distinction between Grinold (1989) and the current work is that we assess the welfare of the investor in "total risk space" whereas Grinold (1989) focuses on active returns relative to a benchmark.

We can now solve for the skill level d , where the expected utility of the mean-variance investor equals the expected utility of the $1/N$ (equally-weighted) investor.

$$V_{ew} = E[U_{ew}] = \frac{\ell' \mu}{N} - \frac{\lambda}{2N^2} \ell' \Sigma \ell \quad (20)$$

The loss function for using mean-variance in preference to $1/N$ is therefore, as before, denoting $V_{ew} = E(U_{ew})$

$$E[L(w_{ew}, w_{mv} | \mu, \Sigma)] = V_{ew} - \frac{(\alpha + d^2 N)}{2\lambda(1 - d^2)} \quad (21)$$

Setting the expected loss function to zero and solving for d , yields,

Corollary 2

Critical forecasting ability, d^ , required for the mean-variance and $1/N$ investor to have the same expected utility under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, no estimation error and in the absence of a budget constraint using the conditional covariance is given by*

$$d^* = \sqrt{\frac{2\lambda V_{ew} - \alpha}{2\lambda V_{ew} + N}}$$

If $V_{ew} < \alpha/2\lambda$, then mean-variance with zero forecasting ability outperforms $1/N$ on average. We also note that as N increases, the required level of forecasting ability d^* to outperform $1/N$ decreases while the converse is true for the squared Sharpe ratio, α . In addition, because the numerator is strictly greater than the denominator, d^* is strictly increasing in the expected utility of the $1/N$ investor, V_{ew} .

E. Multiple forecasting variables, constant forecasting ability, estimation error, no budget constraint

Thus far, we have assumed that we know μ , and can forecast with skill d . We now introduce estimation error in μ . In practice investors do not know the true mean of the distribution and it must be estimated. We replace the population mean, μ , with the sample mean \bar{x} where \bar{x} is $N(\mu, \Sigma/T)$.

Assuming constant forecasting ability as before, the multivariate p.d.f. in this case is:

$$\begin{pmatrix} r_{t+1} \\ a \\ \bar{x} \end{pmatrix} \sim N \left[\begin{pmatrix} \mu \\ 0 \\ \mu \end{pmatrix}, \begin{bmatrix} \Sigma & C & 0 \\ C & I_N & 0 \\ 0 & 0 & \Sigma/T \end{bmatrix} \right] \quad (22)$$

The conditional moments of r_{t+1} given a and \bar{x} are now:

$$\mu^* = \bar{x} + \Sigma^{\frac{1}{2}} d a \quad (23)$$

$$\Sigma^* = \Sigma(1 - d^2) \quad (24)$$

However, the optimal weights are stochastic functions of a and \bar{x} since the investor does not know the true parameter values.

$$\hat{w} = \frac{\Sigma^{-1}(\bar{x} + d\Sigma^{\frac{1}{2}}a)}{\lambda(1 - d^2)} \quad (25)$$

Substituting (53) into (8) gives

Proposition 4

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the absence of a budget constraint using the conditional covariance matrix is given by

$$E[U] = \frac{\alpha - \frac{N}{T} + Nd^2}{2\lambda(1 - d^2)}$$

Proof: See Appendix A

Again this makes sense. The expected utility is positively related to the underlying squared Sharpe ratio, α , and the skill level, d . The expected utility is also positively related to the length of the estimation window, T : expected utility increases as estimation error decreases. The relationship between N and expected utility however is now more complex.

Corollary 3

The relationship between expected utility and the number of assets, N , under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , no estimation error and in the absence of a budget constraint using the conditional covariance matrix is given by

If

$$|d| > \sqrt{\frac{1}{T}} \text{ then } \frac{\partial E(U)}{\partial N} > 0$$

If

$$|d| < \sqrt{\frac{1}{T}} \text{ then } \frac{\partial E(U)}{\partial N} < 0$$

Note that this differs from the fundamental law of active management of Grinold (1989), $IR = d\sqrt{N}$ where the ex-ante information ratio, a transformation of expected utility, is a strictly positive function of the number of assets, N and the work of DeMiguel et al. (2009) where expected utility is a strictly negative function of the number of assets. The critical level of skill, d^* for mean-variance to outperform $1/N$ on average is given below:

Corollary 4

Critical Forecasting Ability, d^ , required for the mean-variance and $1/N$ investor to have the same expected utility under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting Ability level, no estimation error and in the absence of a budget constraint using the conditional covariance Matrix is given by*

$$d^* = \sqrt{\frac{2\lambda V_{\frac{1}{N}} - \alpha + \frac{N}{T}}{2\lambda V_{\frac{1}{N}} + N}}$$

If $V_{1/N} < (\alpha - N/T)/2\lambda$, then mean-variance with zero forecasting ability outperforms $1/N$ on average. Again this rings true. If the unconstrained squared Sharpe ratio, α is larger, we require less skill to outperform $1/N$. As T , the estimation window length increases, critical d^* decreases due to a reduction in estimation error, and converges to the result given by corollary 2. Conversely, if the expected utility of the $1/N$ portfolio is large, ceteris paribus, a higher skill level is required for mean-variance to outperform $1/N$.

V. Further Propositions: The Budget Constrained Case

In section IV, we did not consider the impact of the budget constraint. In practice the vast majority of investors are subject to such a constraint. Moreover, the $1/N$ approach that we are using as our benchmark always satisfies the budget constraint. We therefore incorporate a budget constraint directly into the optimisation problem.

Consider the constant absolute risk aversion utility function, subject to a budget constraint:

$$U = \omega' \mu - \frac{\lambda}{2} \omega' \Sigma \omega - \theta(\omega' \mathbf{1} - 1) \quad (26)$$

where θ is the Lagrange multiplier and the other notation is as before.

The optimal weight in the presence of a budget constraint is given by

Proposition 5

The optimal mean-variance weights in the presence of a budget constraint with known parameters is given by

$$\omega = \frac{1}{\lambda} \Sigma^{-1} \mu - \frac{(\beta - \lambda)}{\lambda \gamma} \Sigma^{-1} \mathbf{1}$$

where $\beta = \mu' \Sigma^{-1} \mu$, $\beta = \mu' \Sigma^{-1} \mathbf{1}$, $\gamma = \mathbf{1}' \Sigma^{-1} \mathbf{1}$

Proof: See Appendix B

The expected utility of the mean-variance investor subject to a budget constraint is then given by

$$V_c = E(U) = \frac{\alpha\gamma - (\beta - \lambda)^2}{2\lambda\gamma} \quad (27)$$

The loss in expected utility from the budget constraint is therefore given by equation (11) minus equation (28) as follows.

$$V - V_c = \frac{(\beta - \lambda)^2}{2\lambda\gamma} = \frac{\gamma\theta^2}{2\lambda} \quad (28)$$

So that if $\theta = 0$, ($\beta = \lambda$), there is no loss in constraining the portfolio, otherwise the loss is unambiguously positive. The condition ($\beta = \lambda$), can be re-expressed as $\lambda^{-1}i'\Sigma^{-1}\mu = 1$ which is simply the requirement that the unconstrained weights given by equation (10) add up to one.

A. Multiple forecasting variables, constant forecasting ability, no estimation error, budget constraint

We now derive the expected utility with a budget constraint and a constant skill level across all assets ignoring estimation error. We substitute the conditional moments (23) and (24) into proposition 5 to provide the optimal portfolio weights conditional upon a as follows

$$w = \frac{1}{\lambda(1 - d^2)} \left(\Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\beta + da' \Sigma^{-\frac{1}{2}} i - \lambda(1 - d^2) \right)}{\gamma} \Sigma^{-1} i \right) \quad (29)$$

We then substitute this relation into the mean-variance utility function (8) to yield the expected utility of the budget constrained mean-variance investor conditional on forecasting ability.

B. Multiple forecasting variables, constant forecasting ability, estimation error, budget constraint

To determine the joint impact of estimation error and forecasting on expected utility, we substitute optimal conditional weight relation (20) into the utility function (2). This is analogous to proposition 4, section IV.

Proposition 6 unifies the impact of estimation error and forecasting ability, two distinct areas of the literature. The proposition also incorporates “alpha” and “beta” returns, whereas the prior analytical literature focussed on one or the other, for example Grinold (1989) that focussed on active management or “alpha”. The theorem also incorporates the effect of a budget constraint facilitating like-for-like comparisons with the 1/N rule.

Proposition 6

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint using the conditional covariance matrix is given by

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

Proof: See Appendix C

We provide a proof in appendix C of the budget constrained case with forecasting, and without estimation error; however, it is more elegant to take the more general case that includes estimation error given by proposition 6, and let T become infinitely large. In doing so, we recover the case without estimation error.

Corollary 7

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , no estimation error and in the presence of a budget constraint using the conditional covariance matrix is given by

$$E[U] = \frac{\alpha + (N-1)d^2 - \frac{(\beta - \lambda(1-d^2))^2}{\gamma}}{2\lambda(1-d^2)}$$

We can now solve for the critical level of forecasting ability required to outperform the $1/N$ rule, d^* .

Corollary 8

The Critical Forecasting Ability, d^ , required for the mean-variance and $1/N$ investor to have the same Expected Utility under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the conditional covariance matrix is given by*

$$d^* = \sqrt{1 + \frac{\left(2V_{ew}\lambda + (N-1) - \frac{2\lambda\beta}{\gamma}\right) \pm \sqrt{\left(2V_{ew}\lambda + (N-1) - \frac{2\lambda\beta}{\gamma}\right)^2 - 4\frac{\lambda^2}{\gamma}\left(\frac{\beta^2}{\gamma} - \alpha - (N-1)\left(1 - \frac{1}{T}\right)\right)}}{\frac{2\lambda^2}{\gamma}}}$$

It is useful to compare the following four expected utility functions in table I.

Table I
Expected utility and Critical Forecasting Level using the Conditional Covariance Matrix

Table I shows the expected mean-variance utility given assumptions regarding the budget constraint, estimation error and forecasting ability. The covariance matrix conditional on forecasting ability is employed.

Case	Budget Constraint	Estimation Error	Forecasting Ability	Expected Utility Function: $E[U]$	Critical Forecasting Ability D^*
(a)	No	No	No	$\frac{\alpha}{2\lambda}$	N/A
(b)	No	No	Yes	$\frac{\alpha + Nd^2}{2\lambda(1-d^2)}$	$\sqrt{\frac{2\lambda V_{1/N} - \alpha}{2\lambda V_{1/N} + N}}$

(c)	Yes	No	Yes	$\frac{\alpha + (N-1)d^2 - \frac{(\beta - \lambda(1-d^2))^2}{\gamma}}{2\lambda(1-d^2)}$	See corollary 8 with $T \rightarrow \infty$
(d)	Yes	Yes	Yes	$\frac{\alpha + (N-1)\left(d^2 - \frac{1}{T}\right) - \frac{(\beta - \lambda(1-d^2))^2}{\gamma}}{2\lambda(1-d^2)}$	See corollary 8

From this table, we can see some clear conclusions. In the absence of a budget constraint and no estimation error, the addition of forecasting ability increases expected utility by

$$\frac{(N + \alpha)d^2}{2\lambda(1 - d^2)}$$

In the presence of a budget constraint, there is no tidy expression for the increase in utility due to forecasting ability. The addition of the budget constraint for an investor with forecasting ability leads to a loss in expected utility of

$$\frac{\gamma d^2 + (\beta - \lambda(1 - d^2))^2}{2\lambda\gamma(1 - d^2)}$$

For the budget constrained case with forecasting ability, the impact of estimation error in the historical mean leads to a reduction in expected utility of

$$\frac{N - 1}{2T\lambda(1 - d^2)}$$

As discussed in section IV, it is typical in the literature to employ an approximation that is akin to the unconditional covariance. Employing the unconditional covariance simplifies the relevant expressions without any meaningful loss of accuracy. For example, at the limit, a highly skilled investor with an information coefficient of 0.1 only needs to scale down the unconditional covariance matrix by 0.5%²⁰. Further, an argument can be made that using the unconditional covariance is consistent with the behaviour of practitioners; in reality, investors do not adjust their risk forecasts based on their ex-ante forecasting ability.

Proposition 7

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint using the unconditional covariance matrix is given by

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

Proof: See Appendix C

²⁰ $1 - \sqrt{1 - d^2} = 0.5\%$

The critical forecasting ability, d^* , is then more straightforward.

Corollary 9

The critical forecasting ability, d^ , required for the mean-variance and $1/N$ investor to have the same expected utility under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the unconditional covariance matrix is given by*

$$d^* = \sqrt{\frac{2\lambda V_{\frac{1}{N}} - \alpha + \frac{N-1}{T} + \frac{(\beta-\lambda)^2}{\gamma}}{N-1}}$$

Table II gives the expected utility and critical forecasting ability using the unconditional covariance matrix. Working from the bottom row, the expected utility of case (d) is derived in proposition 7. To attain the expected utility in case (c) we let T go to infinity, eliminating estimation error in case (d). To attain the expected utility in case (b) we let $\beta=\lambda$, removing the budget constraint. Finally, to derive the expected utility in case (a), we set $d=0$ eliminating forecasting ability. These expressions are far simpler to use, are consistent with the behaviour of practitioners and give almost identical numerical results.

Table II
Expected utility and Critical Forecasting Level using the Unconditional Covariance Matrix

Table II shows the expected mean-variance utility given assumptions regarding the budget constraint, estimation error and forecasting ability. The unconditional covariance matrix is employed.

Case	Budget Constraint	Estimation Error	Forecasting Ability	Expected Utility Function: $E[U]$	Critical Forecasting Ability D^*
(a)	No	No	No	$\frac{\alpha}{2\lambda}$	N/A
(b)	No	No	Yes	$\frac{\alpha + Nd^2}{2\lambda}$	$\sqrt{\frac{2\lambda V_{1/N} - \alpha}{2\lambda V_{1/N} + N}}$
(c)	Yes	No	Yes	$\frac{\alpha + (N-1)d^2 - \frac{(\beta-\lambda)^2}{\gamma}}{2\lambda}$	$\sqrt{\frac{2\lambda V_{ew} - \alpha + \frac{(\beta-\lambda)^2}{\gamma}}{N-1}}$
(d)	Yes	Yes	Yes	$\frac{\alpha + (N-1)\left(d^2 - \frac{1}{T}\right) - \frac{(\beta-\lambda)^2}{\gamma}}{2\lambda}$	$\sqrt{\frac{2\lambda V_{ew} - \alpha + \frac{(N-1)}{T} + \frac{(\beta-\lambda)^2}{\gamma}}{N-1}}$

In the absence of estimation error and a budget constraint, the addition of forecasting ability increases expected utility by

$$\frac{Nd^2}{2\lambda}$$

The addition of a budget constraint in the presence of forecasting ability leads to a loss in expected utility of

$$\frac{d^2 + \frac{(\beta-\lambda)^2}{\gamma}}{2\lambda}$$

Finally, for the budget constrained case with forecasting ability, estimation error in the historical mean reduces expected utility by

$$\frac{N-1}{2T\lambda}$$

C. Multiple forecasting variables, constant forecasting ability, estimation error, budget constraint, assuming constant pairwise correlation and constant volatility

We now provide the expression for expected utility for a simplified case where the volatility and pair-wise correlation is the same across all assets. The goal is to identify the key drivers of expected utility.

Corollary 10

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint with a constant pair-wise correlation, ρ , and a constant stock volatility, σ , across all assets using the unconditional covariance matrix is given by

$$E[U] = \frac{(N-1) \left(\frac{\sigma_u^2}{\sigma^2} (1-\rho) - \frac{1}{T} + IC^2 \right) + O(1)}{2\lambda}$$

Proof: See Appendix D

As asset specific volatility, σ , or asset correlation, ρ , increases, expected utility decreases. Conversely, as cross-sectional dispersion, σ_u , increases, expected utility also increases. It is comforting to see that the theory describes what rings true in investment practice. Further, this result reveals why active investors using optimisers like cross-sectional dispersion; it increases their expected utility.

VI. Empirical Application

We now investigate the expected utility of the mean-variance investor for allowing for forecasting ability in the presence of estimation error and subject to the budget constraint using proposition 6. This case aligns well with the problem of the institutional investment manager. We employ two data sets to replicate the asset allocation and the security selection problems. These problems are the two most common applications of mean-variance analysis in asset management.

A. Data

The first data set replicates the opportunity set of the institutional asset allocator. We use the key assets commonly employed by sophisticated institutional investors: domestic equities, foreign equities, real-estate, government bonds, corporate bonds, and commodities for the period 1975-2012. The summary statistics and are shown in table III. The FTSE/EPRA REIT index has the highest annualised return over the period followed by the S&P 500. The GSCI index has the lowest return and the highest volatility. All assets fail the Jarque-Bera test for normality at the 1% level. For each asset we also estimate the “bull” and “bear” correlations defined as the correlation between the asset and the S&P 500 index when the return of both the S&P 500 and the asset are in the upper (lower) quartile. All of the assets exhibit higher bear correlations than bull correlations. All of the assets exhibit positive autocorrelation in the absolute value of returns consistent with heteroskedasticity or “volatility clustering”.

Table III
Asset Classes: Summary Statistics

Table III provides the summary statistics for the S&P 500, MSCI EAFE, Barclays US Aggregate Bond, FTSE-NARIET, Barclays US Corporate Bond and Goldman Sachs Commodities Index for the period 12/1975-12/2011 using monthly data (N=6, T=433). The average annual return is calculated geometrically. The standard deviation is annualised and is calculated by multiplying the monthly standard deviation by $\sqrt{12}$. Kurtosis refers to excess kurtosis. The “bull” (“bear”) correlation refers to the correlation between the asset and the S&P 500 when the returns of the asset and the S&P 500 are in the upper (lower) quartile of returns. $corr(r_t, r_{t-1})$ refers to the autocorrelation at one lag. $corr(|r_t|, |r_{t-1}|)$ refers to the one-period autocorrelation of the absolute value of returns.

	S&P 500 Index	MSCI EAFE Index	FTSE/ EPRA REIT Index	Barclays US Govt Bond Index	Barclays Corporate Bond Index	GSCI Index
Annualised Return	10.9	10.1	12.1	9.6	9.5	7.1
Standard-deviation	15.4	17.3	17.0	10.6	9.6	19.3
Skewness	-0.55	-0.39	-0.85	0.33	0.18	-0.23
Kurtosis	2.01	0.94	7.84	1.78	3.96	2.36
Maximum	13.5	15.6	28.1	14.2	14.1	22.9
Minimum	-21.5	-20.2	-30.5	-9.2	-11.2	-28.2
Jarque-Bera stat	92.3	26.0	1132.0	63.0	276.1	100.6
Jarque-Bera p-value	0.00	0.00	0.00	0.00	0.00	0.00
"Bull" correlation	1.00	0.30	0.29	0.07	-0.04	-0.21
"Bear" correlation	1.00	0.75	0.65	0.29	0.47	0.59
$corr(r_t, r_{t-1})$	0.04	0.09	0.11	0.09	0.16	0.15
$corr(r_t , r_{t-1})$	0.07	0.21	0.30	0.23	0.28	0.25

The second data set includes the largest 50 U.S. equities²¹ and we refer to this as the stock selection problem. The summary statistics are shown in table II. The mean annualised return is 7.2%, the maximum of return of 47.8% belongs to Apple and the minimum return belongs to AIG. The median number of monthly data points available for each the largest 50 securities is T=265. 24 of the 50 stocks fail the Jarque-Bera tests, and 37 display positive autocorrelation in the absolute value of returns.

Table IV
Equities: Summary Statistics

Table II provides the summary statistics for the 50 largest stocks in the S&P 500 for the period 05/2002-05/2012 using monthly data (N=50, T=120). The average annual return is calculated geometrically. The standard deviation is annualised and is calculated by multiplying the monthly standard deviation by $\sqrt{12}$. Kurtosis refers to excess kurtosis. The “bull” (“bear”) correlation refers to the correlation between the asset and the S&P 500 when the returns of the asset and the S&P 500 are in the upper (lower) quartile of returns. $corr(r_t, r_{t-1})$ refers to the autocorrelation at one lag. $corr(|r_t|, |r_{t-1}|)$ refers to the one-period autocorrelation of the absolute value of returns.

	Mean	Minimum	Maximum
Annualised Return	7.2	-30.2	47.8
Standard-deviation	28.5	14.5	105.1
Skewness	0.09	-1.15	4.95
Kurtosis	3.35	-0.37	38.81
Maximum	31.7	10.1	245.0
Minimum	-25.3	-84.4	-11.6
Jarque-Bera stat	270.9	0.3	7381.9
Jarque-Bera p-value	0.12	0.00	0.50
Mean "bull" correlation	-0.08	-0.78	0.82
Mean "bear" correlation	0.17	-0.66	0.75
$corr(r_t, r_{t-1})$	-0.00	-0.21	0.20
$corr(r_t , r_{t-1})$	0.08	-0.12	0.48

B. *Expected Utility, Forecasting and Estimation Error*

Figure 1 shows the expected utility conditional on forecasting of the budget constrained investor in the presence of estimation error utilising proposition six for the asset allocation problem.

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

Note the hyperbolic relationship between the length of the estimation window and expected utility and the near quadratic relationship between forecasting ability and expected utility. It is evident that a large gain in welfare accompanies both an increase in forecasting ability and an increase in the estimation window.

²¹ As of 5/2012

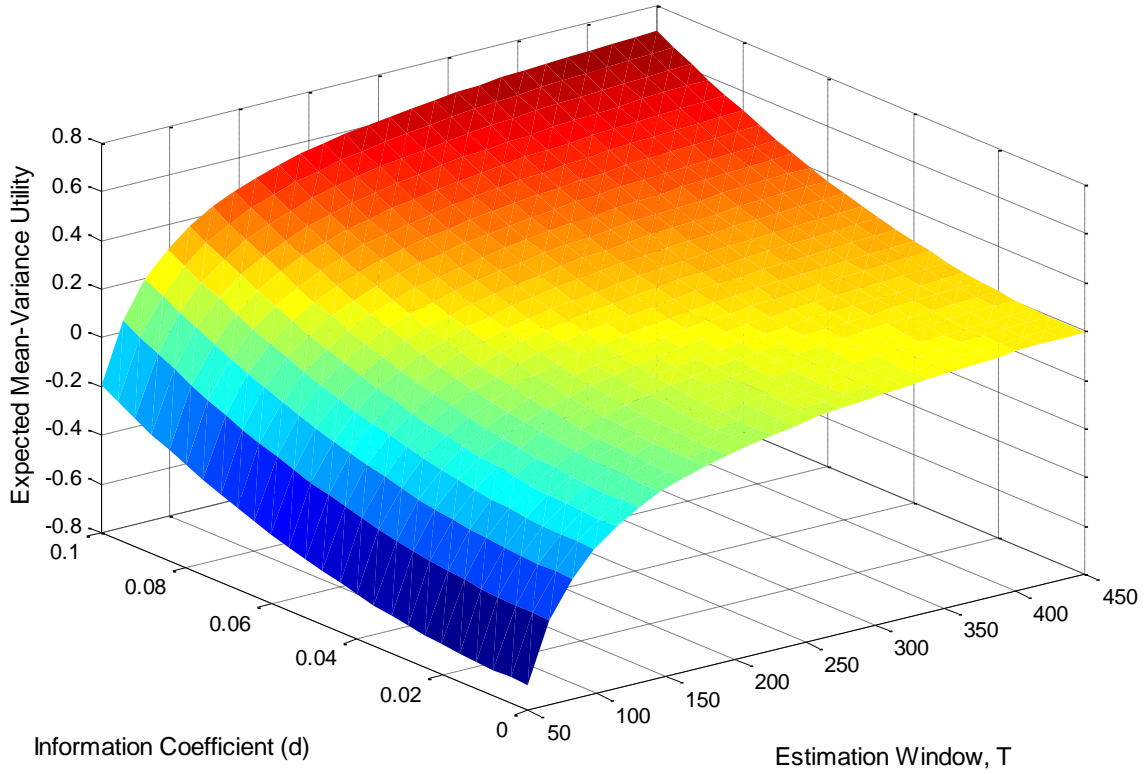


Figure 1 – Expected Utility vs. Forecasting ability and Estimation Window: Asset Allocation. Figure 2 shows the expected utility of the budget constrained investor derived using proposition 6 for the asset allocation problem. To calibrate the model we use the estimated mean and covariance for the S&P 500, MSCI EAFE, Barclays US Govt. Bond Index, Barclays US Corp. Bond Index, Commodities (GSCI), FTSE/NAREIT Index for the period 1975-2011.

Figure 2 shows the expected utility of the mean-variance investor conditional on forecasting and subject to the budget constraint and estimation error using proposition 6 for the asset allocation (panel A) and stock selection problems (panel B). In each case we use estimation windows T of 60, 120, 240, and 480 months. We use a risk aversion, λ , of 0.05 as derived in appendix E. When the estimation window is short, expected utility is low for both data sets irrespective of forecasting ability. In the empirical literature a 60 month estimation window is very common. It is therefore of no surprise that these studies are not complimentary of the mean-variance technique. This is particularly evident for the stock selection problem. As forecasting ability increases, expected utility increases at a near quadratic rate, substantially improving investor welfare. As the estimation window expands, expected utility also increases. The expected utility of the stock selection problem, driven by the higher number of assets, is more sensitive to both estimation error and the level of forecasting ability than the expected utility for the asset allocation problem.

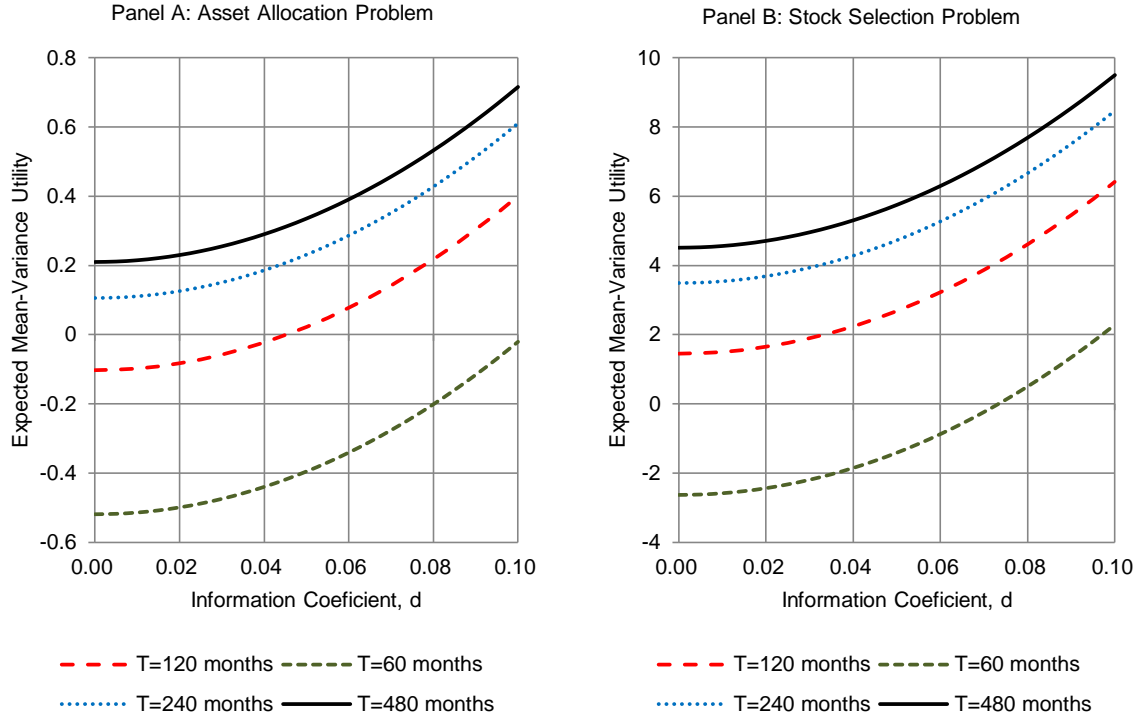


Figure 2 – Expected Utility vs. Forecasting ability: Asset Allocation and Stock Selection. Figure 1 provides the expected utility of the budget constrained investor using proposition 6 for estimation windows of 60, 120, 240, and 480 months. In panel A, we calibrate the model using the estimated mean and covariance for the S&P 500, MSCI EAFE, Barclays US Govt. Bond Index, Barclays US Corp. Bond Index, Commodities (GSCI), FTSE/NAREIT Index for the period 1975-2012. In panel B we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

C. How does utility change with N ?

Kan and Zhou (2008) and DeMiguel et al. (2009) show that in the presence of estimation error, expected utility declines as the number of assets N increases. Once we consider forecasting ability, we see that this conclusion does not necessarily hold. Figure 3 shows the expected utility of the budget constrained mean-variance investor as the number of assets increases for estimation windows of 60, 120, 240 and 480 months dependent on N . In line with the analysis in section IV, corollary 3, the effect of increasing the number of assets, N depends on the relationship between forecasting ability and estimation window.

If

$$d > \sqrt{\frac{1}{T}} \text{ then } \uparrow N \Rightarrow \uparrow E[U]$$

If

$$d < \sqrt{\frac{1}{T}} \text{ then } \uparrow N \Rightarrow \downarrow E[U]$$

This relation also allows practitioners to determine whether or not to add additional assets to the optimisation problem for a given forecasting ability and estimation window.

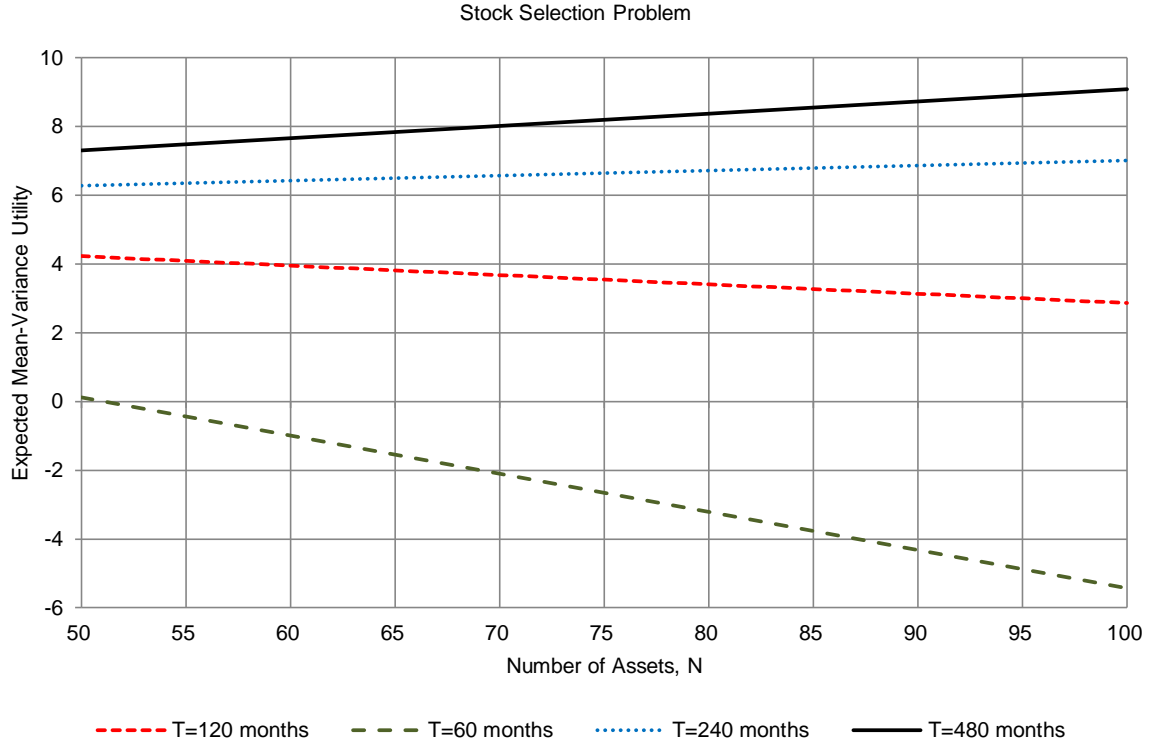


Figure 3 – Expected Utility vs. Assets, N given Estimation Window: 50 largest stocks in S&P 500. Figure 3 shows the expected utility of the budget constrained investor derived using proposition 6 for a forecasting ability, $d=0.075$. To calibrate the model we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

D. How much skill do investors need to outperform $1/N$?

Figure 4 shows the level of forecasting ability required for the budget constrained, mean-variance investor subject to estimation error to outperform $1/N$ using corollary 6 for the asset allocation (left-hand panel) and stock selection (right-hand panel) problems. The expected utility of the $1/N$ rule using equation (3) is 0.79. This leads to a critical skill level, d^* , of 0.0396 for the historical estimation window of $T=433$ months. Note that corollary 8, which uses the unconditional covariance forecast also gives us a critical skill level of 0.0396, with a more straightforward calculation. A correlation between forecast and realised returns of 0.04 is not considered high and is consistent with the levels of typical forecasting variables. For example, for the S&P 500 index, Campbell and Thompson (2008) report R^2 of 0.305%, 0.186%, and 0.271% equating to information coefficients of 0.055, 0.043, and 0.052 for inflation, the term-spread and the earnings-to price ratio. The expected utility of the $1/N$ rule for the stock selection data is 0.64. For the 60 month window commonly used in the mean-variance literature, a forecasting ability of 0.08 is required for mean-variance to outperform $1/N$. Given the average number of monthly data points for the constituents of the S&P 500 is 265, the 60-month assumption is conservative. With 265 months of data, the mean-variance investor does not require any forecasting ability to outperform $1/N$ for the stock selection problem. This implies that for the representative data sets we have presented here, the asset allocator requires a larger amount of skill to outperform $1/N$ than the stock selector. Moreover, we have demonstrated that mean-variance can be expected to outperform $1/N$ for reasonable estimation windows with little or no forecasting ability.

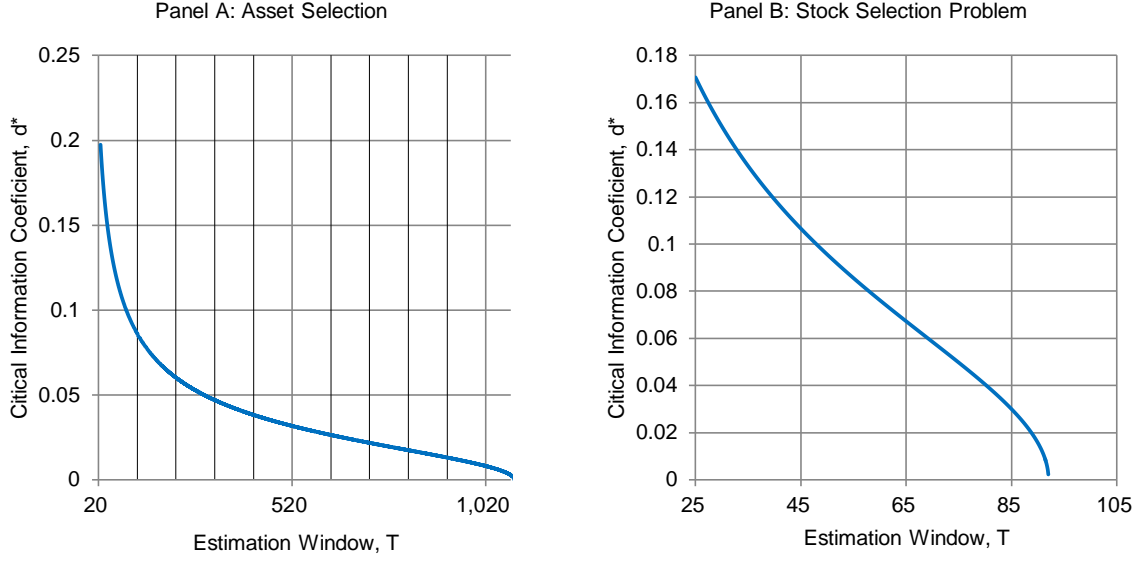


Figure 4 – Critical Information Coefficient, d^* : Asset Allocation and Stock Selection Figure 4 shows the required skill level IC^* for mean-variance to outperform $1/N$ using proposition 6. In panel A we calibrate the model using the estimated mean and covariance for the S&P 500, MSCI EAFE, Barclays US Govt. Bond Index, Barclays US Corp. Bond Index, Commodities (GSCI), FTSE/NAREIT Index for the period 1975-2011. In panel B we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012. In panel B we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

E. What is the probability of outperforming $1/N$?

To gain a richer understanding of the relative performance characteristics of mean-variance and $1/N$, we explore the probability of mean-variance outperforming $1/N$ in a given period. In section IV, we were able to derive the solution for the probability of mean-variance outperforming $1/N$ for the case of a single forecasting variable, $M=1$. In the case of multiple forecasting variables, $M \neq 1$, we estimate the probability through simulation as it is not possible as far as we are aware to express the probability in terms of simple functions as in corollary 1. Specifically, we investigate the probability of mean-variance outperforming $1/N$ conditional on forecasting ability, in the presence of estimation error and subject to a budget constraint, consistent with proposition 6. Our simulations are constructed as follows. We generate 10,000 return histories with 10,000 monthly periods for each forecasting level, d , estimation window length, T , combination using the following ten steps.

Algorithm 5.1

1. Generate information matrix: $N(0,1)$ variables, a (N x L)
2. Generate random matrix: $N(0,1)$ variables, z (N x L)
3. Generate estimation window matrix: $N(0,1)$ variables, e (N x K)
4. Calculate historic sample moments, $\hat{\mu}$ and $\hat{\Sigma}$
5. Generate estimation window: $r_e = \hat{\mu} + \hat{\Sigma}^{1/2}e$ (N x K)
6. Calculate sample mean of estimation window returns: $\bar{\mu}$
7. Combine random variables: $x_t = da_t + \sqrt{1-d^2}z_t$ (N x L)
8. Simulate measurement window: $r_t = \hat{\mu} + \hat{\Sigma}^{1/2}x_t$ (N x L)

9. Calculate covariance of estimation window returns: $\hat{\Sigma}$

10. Simulate expected returns: $\varphi_t = \bar{\mu} + \hat{\Sigma}^{-\frac{1}{2}} a_t$ (N x L)

Having simulated the expected and realised return histories, we then calculate the realised utility of the mean-variance and the 1/N investor each period using the derived weights of the budget constrained investor, as in proposition 5, and the mean-variance utility function. Figure 5 shows the probability of mean-variance outperforming 1/N for four estimation windows of 60, 120, 240 and 480 months using the asset allocation (panel A) and stock selection (panel B) data sets. We see here that the probability of mean-variance outperforming 1/N is greater than 50% for all estimation window lengths and skill levels for both the asset allocation and stock selection problems. Consistent with the finding that a lower level of skill is required to outperform 1/N for the stock selection problem, the probability of outperforming 1/N in a given month is universally higher than for the asset allocator. Even for low levels of skill and estimation data, the probability of mean-variance outperforming 1/N is above 60%.

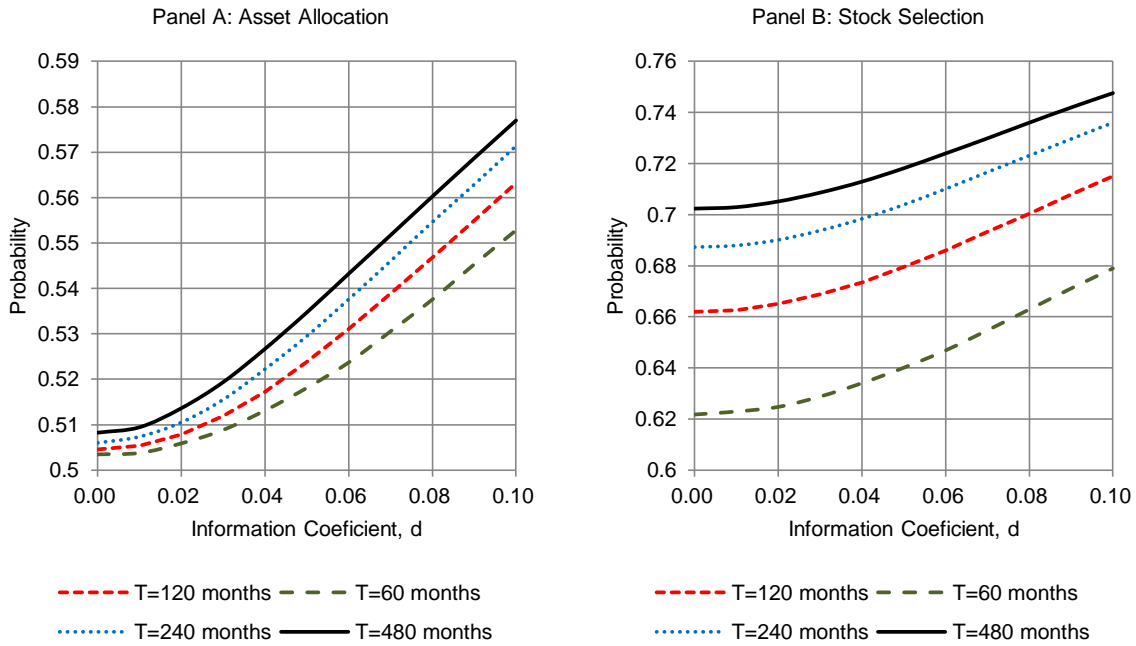


Figure 5 – Probability of Mean-Variance Outperforming 1/N: Asset Allocation and Stock Selection. Figure 5 shows the probability of mean-variance generating a higher utility than the 1/N rule in a monthly period. We attained the probabilities through simulating 10,000 sets of asset allocation data and stock selection data of 10,000 months, for each level of skill, D, and estimation window length, T. In panel A, we calibrate the model using the estimated mean and covariance for the S&P 500, MSCI EAFE, Barclays US Govt. Bond Index, Barclays US Corp. Bond Index, Commodities (GSCI), FTSE/NAREIT Index for the period 1975-2011. In panel B, we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

F. Expected Utility and Estimation Error in the Covariance Matrix

A potential shortcoming of the preceding analysis is that we assume the covariance matrix is known. Kan and Zhou (2007) and DeMiguel et al. (2009) for example consider the case when the covariance matrix is unknown. While conventional wisdom suggests that the covariance matrix is a key source of estimation risk, we argue that the estimation error in the covariance may not be particularly important in practice. We base this on two observations. First, factor model approaches are heavily employed in the finance industry. Second, increasing the sampling frequency of the covariance matrix leads to a reduction in estimation error

to any arbitrary level. Under the normal distribution, Cochran's theorem shows that the sample variance follows a scaled chi-square distribution (see Knight, 2000).

$$(n - 1) \frac{s^2}{\sigma^2} \sim \chi_{n-1}^2$$

where

$$E[s^2] = \sigma^2, \text{Var}(s^2) = \frac{2\sigma^4}{n - 1}$$

Hence the standard-error of the sample variance tends asymptotically towards zero. In an empirical setting, Andersen and Bollerslev (1998) show that realized volatility estimated from high-frequency returns is effectively an error-free volatility measure in the foreign exchange market. Andersen and Bollerslev, Diebold and Ebens (2001) extend this finding to the constituents of the DJIA. This conclusion has naturally led to the idea of observed or *realized* volatility and correlation. Conversely, the estimation error in expected returns is invariant to the sampling frequency and can only be reduced through expanding the estimation window.

In proposition 6, we assume that the true mean is unknown, whilst the true covariance is known. In the absence of a budget constraint and forecasting ability we can easily incorporate the effect of estimation error in the covariance matrix; however, once we incorporate both forecasting ability and the budget constraint, the additional terms lead to an unwieldy analytical result. We instead use simulation to understand the impact of estimation error in the covariance using the simulation framework set out above. Whereas in the preceding analysis we assumed the covariance matrix is known, we now assume that it must be estimated along with the mean.

Figure 6 shows the expected utility of the mean-variance investor assuming that the covariance matrix is known (panel A) and where the covariance matrix must be estimated (panel B) for the asset allocation problem. The additional estimation error in the covariance matrix leads to a downward shift and clockwise translation of expected utility. As one would expect, when the estimation window is short there is a significant reduction in expected utility. Given our historical period of 433 months, the inclusion of estimation error in the covariance matrix leads to only a slight increase in the critical skill level for mean-variance to outperform 1/N from 0.0396 to 0.0415.

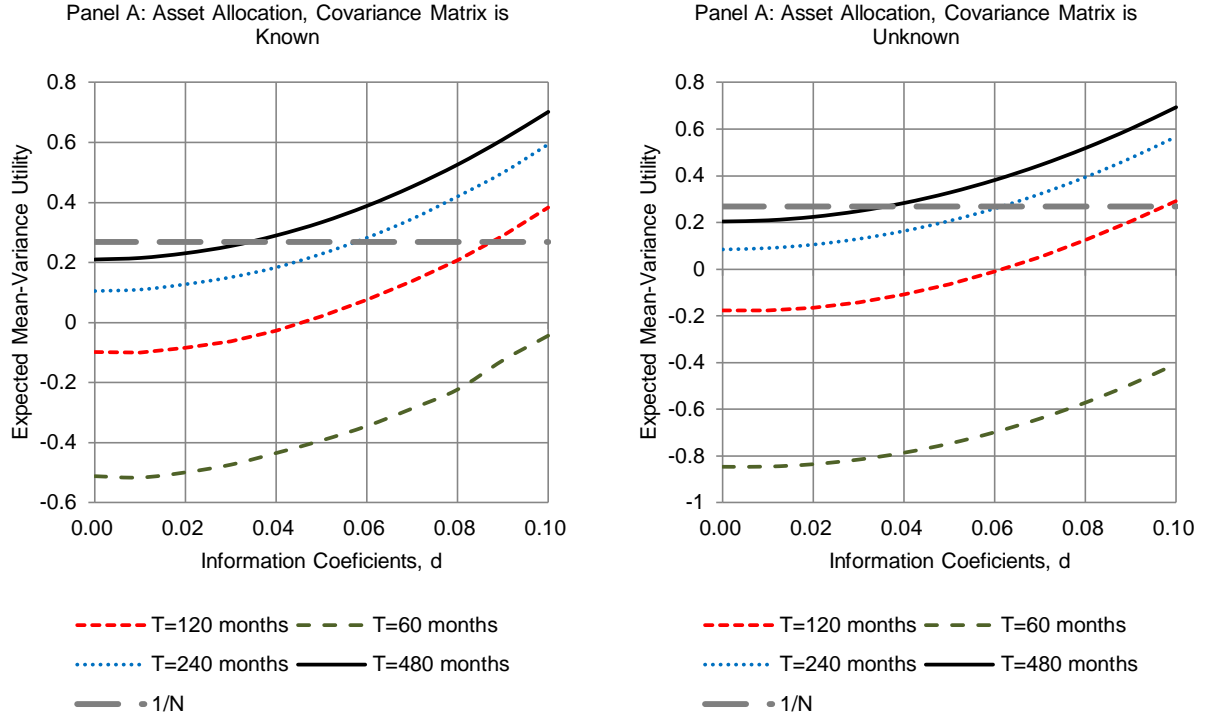


Figure 6 – Expected Utility vs. Forecasting ability: Asset Allocation. Figure 6 provides the expected utility generated through simulation for the budget constrained investor for each level of skill, D , and estimation window length, T . In the left-hand panel we show the result where the mean is unknown and the covariance is known. In the right-hand panel we show the result where both the mean and the covariance matrix are unknown and must be estimated. The horizontal line shows the expected utility of the $1/N$ investor. To calibrate the model we use the estimated mean and covariance for the S&P 500, MSCI EAFE, Barclays US Govt. Bond Index, Barclays US Corp. Bond Index, Commodities (GSCI), FTSE/NAREIT Index for the period 1975-2012.

Figure 7 shows the same analysis for the stock selection problem. In contrast to the asset allocation problem, the impact of estimation error is significant. This of course is driven by the large number of terms that must be estimated in the stock selection problem. Despite this, for estimation windows in line with the typical S&P 500 constituent and for moderate forecasting levels, mean-variance will still outperform $1/N$ on average. The large effect of estimation error in the covariance matrix for the stock selection problem is driven by scarce amount of data, NT , available to estimate each term in the covariance matrix, $N(N-1)/2$. In the stock selection problem for the 60 month window common in the literature, we have only 2.45 data-points per term, as against 24 data points per term for the asset allocation problem.

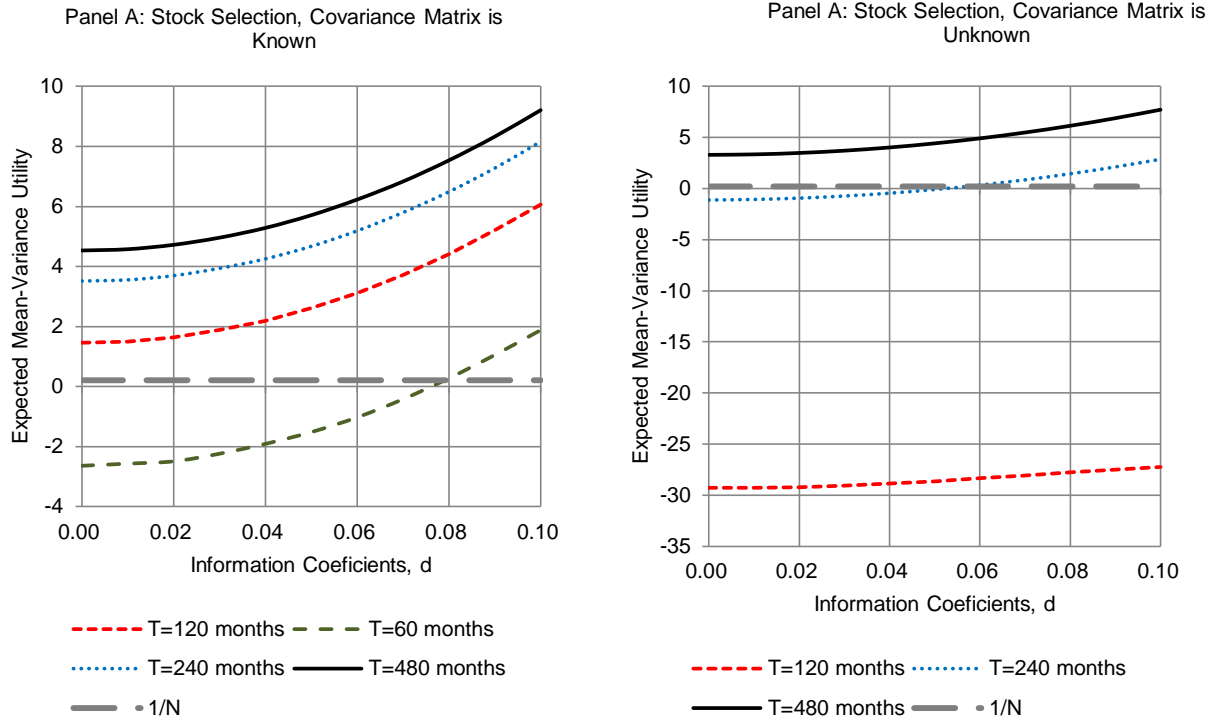


Figure 7 – Expected Utility vs. Forecasting ability: Stock Selection. Figure 7 provides the expected utility derived through simulation for the budget constrained investor for each level of skill, d , and estimation window length, T . In panel A we show the result where the mean is unknown and the covariance is known. In panel B we show the result where both the mean and the covariance matrix are unknown and must be estimated. The horizontal line shows the expected utility of the $1/N$ investor. To calibrate the model we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

In the preceding analysis, consistent with prior research in the area, we have used monthly observations to compute the covariance matrix. Sampling theory suggests however that we should use a high frequency to increase the precision of our covariance estimates. We can increase the level of precision to any arbitrary level by merely increasing the sampling frequency. Figure 8 shows the expected mean-variance utility when using fortnightly, weekly, daily, and 10 minute interval tick data for an estimation window of 120 months, assuming the mean vector and the covariance matrix are unknown. The expected utility levels are derived using simulation algorithm 5.1. We do not show the expected utility using monthly data because it is so low that it distorts the scale of the scale of the chart. It is of little wonder that the literature shows such poor results for the mean-variance technique given the almost blanket reliance on monthly data to estimate the covariance matrix. The solid line shows the expected utility of the mean-variance investor assuming the mean return is unknown and the covariance matrix is known. As we increase the sampling frequency, the covariance matrix is estimated with more and more precision and expected utility increases to the point where expected utility is virtually identical to that given by proposition 6. We do not show the expected utility using the one-minute frequency, as it is indistinguishable from the case when the true covariance matrix is known. We therefore maintain that as long as the covariance matrix is estimated using data of a sufficient frequency, the effect of estimation error in the matrix is not economically significant. An argument can be made that proposition 6 captures all of the practically relevant estimation error. This is fortuitous as incorporating estimation error and the budget constraint analytically, which we argue is essential given we are comparing the performance to of the mean-variance technique to $1/N$ which always satisfies the budget constraint, produces unwieldy expressions.

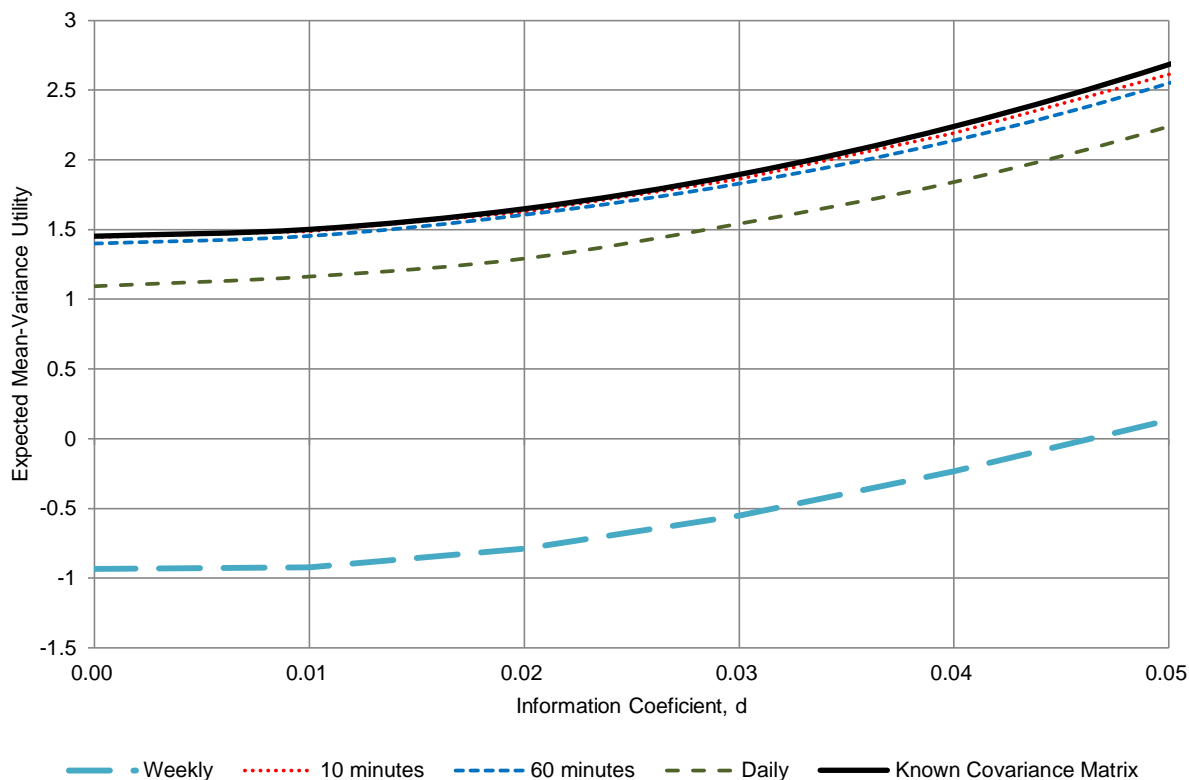


Figure 8 – Expected Utility vs. Forecasting ability: Stock Selection. Figure 8 provides the expected utility derived through simulation for the budget constrained investor for each level of skill, d , and an estimation window length of $T=120$ where the mean and covariance matrix are unknown. The figure shows the expected utility where the covariance matrix is estimated using fortnightly, weekly, daily and minute interval data, and where the covariance matrix is known. To calibrate the model we use the estimated mean and covariance for the largest 50 stocks (as at 5/2012) in the S&P 500 for the period 5/2002-5/2012.

G. Robustness Tests and Reconciliation with Extant Literature

Our key finding is that the mean-variance approach can be expected to outperform $1/N$ for typical investment problems. This runs counter to the sample-based empirical literature and the analytical findings of DeMiguel et al. (2009). The poor out-of-sample performance of the mean-variance technique in the empirical literature makes perfect sense given the almost universal reliance on short histories of monthly data to estimate the mean vector and covariance matrix. The problem is even more acute in high dimensions. In the case of the covariance matrix, this is an issue easily remedied by employing high-frequency data. Further, if we allow for forecasting we are not beholden to the estimation error in sample means.

What then of the analytical results that seem to so conclusively discount the mean-variance technique? Chief among these results is the model of DeMiguel, Garlappi and Uppal (2009) which follows from Kan and Zhou (2007). DeMiguel et al. (2009) find that for portfolios of 25 and 50 assets, prohibitively large estimation windows of 3000 (250 years) and 6000 months (500 years) are required respectively for mean-variance to outperform $1/N$. Zhou (2008) finds that over 10,000 months of data is required for mean-variance to achieve 90% of the true maximum utility level. Hence, the conclusion that there are “many miles to go” before the gains promised by optimal portfolio choice can actually be realised out of sample”. DeMiguel et al. (2009) consider seven data sets, so perhaps our analysis is too narrow in scope. We

therefore apply proposition 6 to the DeMiguel et al. (2009) data sets for five levels of forecasting ability²². The “International” data set includes eight developed market MSCI indices and the MSCI World index. The “Industries” data set includes ten U.S. value-weighted industry portfolios. The “MKT/SMB/HML” data set includes the market, size and book-to-market portfolios. The “FF-1” data set includes the 20 size and book-to-market portfolios and the market portfolio. The “FF-3” data set augments the FF-2 data set with the size and book-to-market portfolios. The “FF-4” data set augments FF-3 with the momentum factor, “UMD”. Table G describes the data sets in detail. We use the same inception points as in DeMiguel et al. (2009). The results are shown in table V. For every level of forecasting ability and for every investment problem the mean-variance approach outperforms 1/N with a single exception, the case of zero forecasting ability for the International data set. It is clear that our findings are not being driven by the use of different investment universes. It is noteworthy that the findings presented in table V are based on the truncated data sets used in DeMiguel et al. (2009). For example for five out of six of the data sets, DeMiguel et al. (2009) discard nearly 40 years’ worth of data that is available prior to 1963²³. If we incorporate the extra 40 years of data, estimation error is further reduced and the finding that mean-variance outperforms 1/N becomes even more emphatic. This makes the results in table V all the more striking.

Table V
Expected Utility of Mean-Variance vs 1/N: Proposition 6

Table V shows whether the expected utility of the mean-variance investor using proposition 6 exceeds the expected utility of the 1/N investor for each of the data sets. The estimation window, T, refers to the number of months used to estimate the expected return vector. A tick mark signifies that the expected utility of the mean-variance approach exceeds the 1/N approach using proposition 6, cross mark signifies the reverse.

	Internat.	Industry	MKT/SMB/ HML	FF-1- Factor	FF-3- factor	FF-4- factor
Assets, N	9	11	3	21	23	24
Estimation Window, T	526	601	603	603	603	603
<u>Information Coefficient, IC</u>						
IC = 0	✗	✓	✓	✓	✓	✓
IC = 0.025	✓	✓	✓	✓	✓	✓
IC = 0.05	✓	✓	✓	✓	✓	✓
IC = 0.075	✓	✓	✓	✓	✓	✓
IC = 0.10	✓	✓	✓	✓	✓	✓

DeMiguel et al. (2009) consider three cases, where the mean is unknown and the covariance is known, where the mean is known and the covariance is unknown, and where the mean and the covariance are both unknown. The authors provide the conditions for mean-variance to outperform 1/N in each case.

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:

$$S_{mv}^2 - S_{ew}^2 - \frac{N}{T} > 0$$

²² Note we did not have access to the S&P sector data set, however it is similar in construction to the value-weighted sectors data set that we do employ

²³ from the same K.R French source

where S_{mv}^2 and S_{ew}^2 are the squared in sample Sharpe-ratios of the mean-variance and 1/N portfolios, N is the number of assets, and T is the length of the estimation window

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:

$$kS_{mv}^2 - S_{ew}^2 - \frac{N}{T} > 0$$

where

$$k = \left(\frac{T}{T - N - 2} \right) \left(2 - \frac{T(T - 2)}{(T - N - 1)(T - N - 4)} \right)$$

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:

$$kS_{mv}^2 - S_{ew}^2 - h > 0$$

where

$$h = \frac{NT(T - 2)}{(T - N - 1)(T - N - 2)(T - N - 4)}$$

When DeMiguel et al. (2009) applied these equations²⁴ they used typical values of the mean-variance and 1/N Sharpe ratios from the seven data sets to calibrate their model. The data sets include between three and 24 assets. DeMiguel et al. (2009) then allowed the number of assets, N to increase to 100 to explore the impact of an increase in the number of assets. This implies that we can somehow massively increase the number of developed equity markets or industries while preserving the structure of the problem. In practice, if we were to add more countries or use a more granular definition of sectors, the Sharpe ratios of the mean-variance and 1/N portfolios will change, invalidating the analysis. We argue that the natural way to apply the DeMiguel et al. (2009) expressions is to simply substitute in the actual number of assets, the available estimation window, and the observed Sharpe ratios.

The results of this analysis are shown in table VI. The DeMiguel et al. (2009) model now shows that irrespective of whether the moments are known, mean-variance can be expected to outperform 1/N. This is true whether we use the same sample period as DeMiguel et al. (2009) or the extended sample including the most recent data. It is apparent, that the DeMiguel et al. (2009) analytical finding that mean-variance underperforms 1/N is driven by a peculiarity in the way they have applied the model and does not reflect the underlying dynamics of the problem. The goal of the authors was to investigate the performance of mean-variance when the number of assets is large. Rather than using the estimated Sharpe ratios of a “small” problem with a low number of assets and extrapolating the results to “large” problems, we argue that it makes much more sense to use the actual moments of the large problem as we have done with the stock selection data set in section VI D. For the stock selection problem, cross-sectional dispersion is high leading to high in-sample mean-variance Sharpe ratios. This is missed by the DeMiguel et al. (2009) extrapolation approach. Again the results from our application of the DeMiguel et al. (2009) model in table VI are conservative as we are not using the entire available data history to calibrate the model.

²⁴ p. 1940

Table VI
Expected Utility of Mean-Variance vs 1/N: DGU (2009)

Table VI shows whether the expected utility of the mean-variance investor using proposition 6 exceeds the expected utility of the 1/N investor for each of the data sets. The estimation window, T, refers to the number of months used to estimate the expected return vector. A tick mark signifies that the expected utility of the mean-variance approach exceeds the 1/N approach using the three conditions of DeMiguel et al (2009).

	Internat.	Industry	MKT/SMB/ HML	FF-1- factor	FF-3- factor	FF-4- factor
Assets, N	9	11	3	21	23	24
Estimation Window, T	526	601	603	603	603	603
S_{ew}^2	0.15	0.13	0.19	0.14	0.14	0.15
S_{mv}^2	0.48	0.23	0.22	0.45	0.45	0.48
<u>Conditions</u>						
μ unknown, Σ is known	✓	✓	✓	✓	✓	✓
μ known, Σ is unknown	✓	✓	✓	✓	✓	✓
μ unknown, Σ is unknown	✓	✓	✓	✓	✓	✓

H. Out-of-sample Tests: Dynamic Portfolio Rebalancing

For completeness, we now investigate the performance of the mean-variance approach in a dynamic portfolio rebalancing framework. This approach is complementary to our analytical model in that it steps away from the idealised Gaussian i.i.d. model and the assumption that the covariance matrix is known. Under the dynamic portfolio rebalancing framework, expected returns and the covariance matrix are estimated, the portfolio is rebalanced, and the returns of the optimal portfolio are recorded. The process is then repeated each month until the end of the data series. By far the most comprehensive work in this area is again DeMiguel et al. (2009). The authors evaluate 14 extensions to mean-variance designed to reduce estimation error across six data sets. In five out of the six of the data-sets that DeMiguel et al. (2009) consider the authors find that 1/N generates higher Sharpe ratios and certainty equivalents than the mean-variance approach. Given that four out of six of these data sets span more than 40 years' of time and the other two data sets span more than 20 years, this evidence is not to be dismissed lightly. The classical expression for the optimal unconstrained weights is given by:

$$w = \frac{\Sigma^{-1}\mu}{\lambda} \quad (30)$$

As discussed in section II, a potential issue with comparing the performance of equation (30) with 1/N is the classical mean-variance weights are not constrained to equal one creating an arguably unfair comparison. We dealt with this issue by imposing a budget constraint on the optimisation problem, yielding proposition 5 as follows

$$\omega = \frac{1}{\lambda} \left(\Sigma^{-1} \mu - \frac{(\beta - \lambda)}{\gamma} \Sigma^{-1} i \right) \quad (31)$$

where $\beta = \mu' \Sigma^{-1} \mu$, $\gamma = i' \Sigma^{-1} i$

In contrast, DeMiguel et al. (2009) solve the budget constraint problem by dividing the unconstrained mean-variance weights (30) by the absolute value of the sum of the weights as follows

$$w = \frac{\Sigma^{-1} \mu}{|i' \Sigma^{-1} \mu|} \quad (32)$$

The authors then apply (32) to “almost all of the models” that they consider. At first blush, although arguably heuristic in nature, the approach makes sense, rescaling the unconstrained weights to sum to one. On closer examination however the approach can lead to two absurdities. First, when the sum of the unconstrained weights (30) approaches zero, the absolute value of the weights given in (31) becomes very large. Second, and more worryingly, when the sum of the unconstrained weights is negative, the sum of the weights will equal negative one rather than positive one, violating the budget constraint. Perhaps these anomalies are more theoretical than practical concerns. If not they affect almost all of the empirical work of DeMiguel et al. (2009).

In figure 9, we show the time-series of the range of the monthly optimal weights using (30), (31) and (32), defined as the maximum weight minus the minimum weight, for the FF-4-factor data set²⁵. It is evident that expressions (30) and (31) yield smooth time-series of maxima and minima. The DeMiguel (2009) et al. expression however leads to spikes in the optimal weights. Rather than being the response to sudden decreases in risk, these spikes are an artefact of the normalisation procedure. Further, we find that for the FF-4-factor data set, the budget constraint is violated in 73% of months with the sum of the unconstrained weights summing to negative one.

²⁵ For the purpose of the figure we use a 120 month estimation window to be consistent with DeMiguel (2009) et al. the test begins at 7/1973 and ends at 9/2013

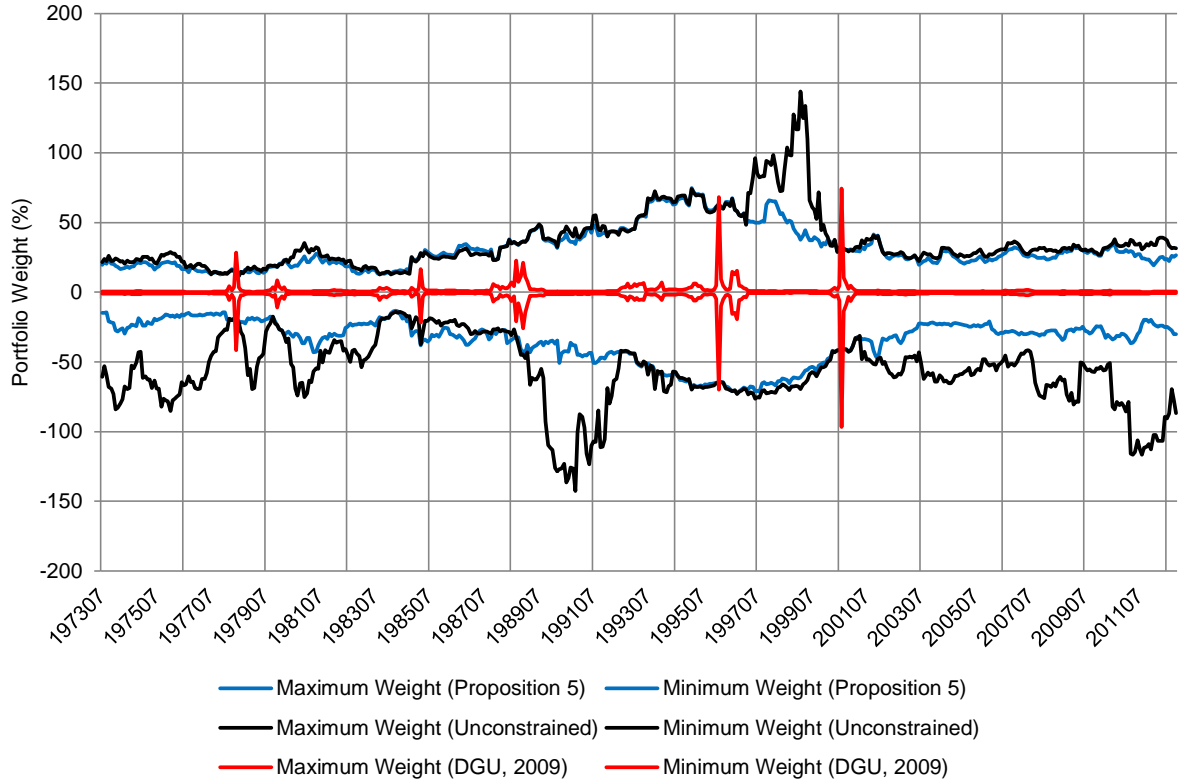


Figure 9 – Maximum and Minimum Optimal Weights. Figure 9 shows the time series of maximum and minimum portfolio weights using expressions (29), (30), and (31) for the FF-4-Factor data set with $\lambda = 1$

To see if the DeMiguel et al. normalisation is driving the poor performance of the mean-variance approach we replicate the dynamic portfolio monthly rebalancing analysis using the three weight expressions (30), (31) and (32). We consider the same out-of-sample period and the same data sets²⁶ as DeMiguel et al. (2009) and begin each test at the same point in time²⁷ to ensure that our results are comparable. To be conservative, we do not allow for forecasting or the use of high frequency covariance estimators. Rather than arbitrarily using a 60 or 120 month estimation window and discarding the majority of the data we use an expanding estimation window. We use three levels of risk aversion, λ , of 1, 5, and 10.

Practitioners tend to focus on Sharpe ratios rather than certainty equivalents. We therefore provide the out-of-sample Sharpe ratios for the 1/N rule, the DeMiguel et al. (2009) optimal weights (31), the classical unconstrained weights (30), and the budget constrained optimal weights (31) strategies for the six data sets in table VII. The optimal weight expression used by DeMiguel et al. (2009) yields an average Sharpe ratio that is 20% lower than the 1/N strategy²⁸, mirroring the original DeMiguel et al. (2009) conclusion. The correct budget constrained weights given in proposition 5 generate an average Sharpe ratio that is 10% lower ($\lambda=1$) and 27% and 29% higher ($\lambda=5, \lambda=10$).

Table VII
Sharpe Ratios of Mean-Variance vs 1/N: DGU (2009)

Table VII shows the annualised out-of-sample Sharpe ratios for the 1/N rule, and for weights derived using the DGU (2009) expression for the optimal mean-variance weights (31), the classical expression for mean-variance weights (29), and the optimal budget constrained optimal weights given (30) and proposition 5 for three levels of risk aversion, λ .

²⁶ 120 months after the data inception dates of DeMiguel et al. (2009)

²⁸ The DGU (2009) optimal weights and the Sharpe ratio are invariant to risk aversion. Thus, the Sharpe ratios are the same for each level of risk aversion.

	Internat.	Indust.	MKT/ SMB/ HML	FF-1- factor	FF-3- factor	FF-4- factor	Average	% Gain vs. 1/N
<u>$\lambda=1$</u>								
1/N	0.43	0.47	0.69	0.53	0.54	0.57	0.54	
DGU (2009), (31)	0.19	0.54	0.68	0.82	-0.01	0.73	0.49	-20%
Proposition 5, (30)	-0.15	0.24	0.30	0.72	0.79	1.01	0.49	-10%
Classical (29)	0.01	0.50	0.67	0.80	0.73	0.97	0.61	26%
<u>$\lambda=5$</u>								
1/N	0.43	0.47	0.69	0.53	0.54	0.57	0.54	
DGU (2009), (31)	0.19	0.54	0.68	0.82	-0.01	0.73	0.49	-20%
Proposition 5, (30)	0.21	0.55	0.71	0.81	0.80	1.01	0.68	27%
Classical (29)	0.01	0.50	0.67	0.80	0.73	0.97	0.61	-10%
<u>$\lambda=10$</u>								
1/N	0.43	0.47	0.69	0.53	0.54	0.57	0.54	
DGU (2009), (31)	0.19	0.54	0.68	0.82	-0.01	0.73	0.49	-20%
Proposition 5, (30)	0.35	0.64	0.70	0.68	0.80	1.02	0.70	29%
Classical (29)	0.01	0.50	0.67	0.80	0.73	0.97	0.61	-12%

The certainty equivalent is the theoretically correct means of evaluating ex-post utility. Table VIII shows the out-of-sample certainty equivalents for the monthly rebalanced portfolios of the four strategies. The average certainty equivalent of the DeMiguel et al. (2009) optimal weights (32) is negative and large for each level of risk. The negative certainty equivalents are the result of the spikes in volatility driven by the normalisation step. In contrast, the average certainty equivalent of the correct budget constrained optimal weights (31) is positive and exceeds the certainty equivalent of the 1/N rule for all three levels of risk aversion. The absolute gain in average certainty equivalent ranges between 4.8% and 8.4%. Given that the annual management fees on a typical mutual fund are in the region of 1%, this gain is economically meaningful. Table VIII also shows that the gain from using the unconstrained optimal weights is also economically significant for all three levels of risk aversion. It is noteworthy that the “International” data set where the mean-variance approach performs poorly relative to 1/N is the data set with the shortest history. For example, only 120 months are used to estimate the returns and covariance matrix for the first monthly optimisation. The lack of predictive power of such short windows is well known to be poor (Jorion, 1985).

It appears that by using the correct expression for the optimal weights and an expanding estimation window we have turned the DeMiguel et al. (2009) conclusion on its head. It would seem the DeMiguel et al. (2009) result is driven by a step that appears innocuous, the normalisation of weights. As the baseball great Yogi Berra quips, “in theory there is no difference between theory and practice. But in practice, there is”. We reiterate that our results may be conservative in that we do not allow for forecasting ability, factor-based risk models, high-frequency covariance estimators or employ the full data history.

Table VIII

Certainty Equivalents of Mean-Variance vs 1/N

Table VIII shows the annualised out-of-sample certainty equivalents for the 1/N rule using weights derived using the DGU (2009) expression for the optimal mean-variance weights (31), the classical expression for mean-variance weights (29), and the optimal budget constrained optimal weights given (30) and proposition 5 for three levels of risk aversion, λ .

	Internat.	Indust.	MKT/ SMB/ HML	FF-1- factor	FF-3- factor	FF-4- factor	Average	CE Gain vs. 1/N
<u>$\lambda=1$</u>								
1/N	5.7%	6.0%	4.5%	8.3%	8.0%	8.1%	6.8%	
DGU (2009), (31)	-23.6%	9.4%	4.1%	18.4%	-3708%	20.4%	-613%	-620%
Proposition 5, (30)	-26.0%	-11.1%	4.3%	25.6%	31.3%	43.4%	11.3%	4.8%
Classical (29)	-22.9%	2.6%	15.2%	32.2%	26.4%	37.7%	15.2%	8.7%
<u>$\lambda=5$</u>								
1/N	0.5%	1.4%	3.6%	1.1%	1.9%	2.5%	1.8%	
DGU (2009), (31)	-188.0%	-0.2%	3.3%	3.9%	-18494%	-7.1%	-3114%	-3109%
Proposition 5, (30)	-4.2%	0.9%	3.5%	6.3%	6.4%	8.8%	3.6%	8.4%
Classical (29)	-4.6%	0.5%	3.0%	6.4%	5.3%	7.5%	3.0%	7.8%
<u>$\lambda=10$</u>								
1/N	-6.1%	-4.5%	2.4%	-7.9%	-5.8%	-4.4%	-4.4%	
DGU (2009), (31)	-393.5%	-12.1%	2.3%	-14.4%	-36978%	-41.5%	-6240%	-6235%
Proposition 5, (30)	-6.3%	-2.3%	2.4%	-0.9%	3.2%	4.4%	0.1%	4.8%
Classical (29)	-2.3%	0.3%	1.5%	3.2%	2.6%	3.8%	1.5%	6.3%

VII. Conclusions

The academic literature would appear to be virtually unanimous in promoting equal weighted investment as preferable to mean-variance optimisation. This literature tends to make three strong assumptions, namely, that the investor has no forecasting ability, that there is no budget-constraint, and that estimation error in the covariance matrix is irreducible and highly destructive. We depart from previous work by allowing for forecasting and imposing a budget constraint. Allowing for forecasting is rooted in the burgeoning anomalies literature on the cross-section of expected returns, the behaviour of practitioners and the original Markowitz (1952, 1959) theory. Incorporating the budget constraint is consistent with the structure of a mutual fund and we assert is necessary for a fair comparison with the 1/N rule. Further, imposing the budget constraint leads to a strictly non-positive change in expected utility, and therefore provides a more rigorous test of mean-variance performance. By incorporating forecasting ability and estimation error we get a number of new results for ex-ante expected utility.

A key goal of this thesis is to provide a decision rule for practitioners to choose between the mean-variance approach and 1/N. DeMiguel et al. (2009) provide such a rule, but ignore the possibility of forecasting ability and the budget constraint. By incorporating these important effects we argue that we have provided a more accurate rule. We have also provided a closed form expression for the critical amount of forecasting ability that is required for mean-variance to outperform 1/N. We show that for the majority of the investment problems that we consider, zero forecasting ability is required for mean-variance to outperform 1/N. Further, modest levels of forecasting ability translate into significant increases in expected utility.

We have also shown that once we allow for forecasting the idea that increasing the number of assets in the portfolio strictly decreases expected utility due to the effect of estimation error needs revision. If the level

of forecasting ability exceeds the reciprocal of the length of the estimation window then increasing the number of assets in the portfolio leads to a gain in expected utility. This runs counter to the analytical models of Kan and Zhou (2007) and DeMiguel, Garlappi, and Uppal (2009). We find that for plausible skill levels, increasing the number of assets increases expected utility; a result first shown by Grinold (1989), but neglected by the academic literature.

A potential shortcoming of our analytical analysis is that we ignore the effect of estimation error in the covariance matrix. The estimation error inherent in the covariance matrix increases exponentially with the number of assets and is a potential hurdle to the successful implementation of mean-variance optimisation in high dimensions. We have shown that in the absence of microstructure effects that it is possible to all but eliminate the negative effect of covariance matrix estimation error on expected utility by using high-frequency estimators. This result lends support for the use of high-frequency covariance matrices by institutional asset managers.

We attempt to reconcile our results with the generally poor mean-variance performance reported in the influential analytical and empirical work of DeMiguel et al. (2009). We show that the most natural application of the DeMiguel et al. (2009) model itself shows that mean-variance can be expected to outperform $1/N$. The DeMiguel et al. (2009) result is due to a mismatch in model parameters. We also show empirically that constrained mean-variance outperforms $1/N$ on average across a range of investment problems. This serves as a robustness check to our analytical model in that we no longer assume that the covariance matrix is known or that returns are i.i.d. Gaussian. The empirical findings of DeMiguel et al. (2009) appear to be driven by an idiosyncrasy in the way the authors calculate optimal weights that leads to extreme weights and violations of the budget constraint.

Our results present a much more compelling case for optimised portfolios. The analytical results of Kan and Zhou (2007) and DeMiguel et al. (2009) on the performance of mean-variance suggest that vast amounts of data are required for mean-variance to outperform $1/N$. We show that even for modest levels of forecasting ability and realistic estimation windows we can expect mean-variance to outperform $1/N$. DeMiguel et al. (2009) conclude that there are “many miles to go” before the promised benefits of optimal portfolio choice can be realised out of sample. Our results suggest that we may have already arrived.

Appendix A – Derivation of Propositions: The Unconstrained Case

A. Proof of Proposition 1

The unconditional expected utility of the mean-variance investor under the assumptions of a single forecasting variable, a ($M=1$), no estimation error and in the absence of a budget constraint:

Given equation (1), the conditional density of $r_{t+1}|a$:

$$N(\mu + Ca, \Sigma - CC')$$

We let the conditional squared Sharpe ratio equal:

$$\alpha^* = \mu^{*'} \Sigma^{*-1} \mu^{*'} \quad A1$$

where

$$\mu^* = \mu + Ca \quad A2$$

$$C = \Sigma^{\frac{1}{2}} D \quad A3$$

where D is $(N \times 1)$

Using the relation:

$$(\Sigma - CC')^{-1} = \Sigma^{-1} + \frac{\Sigma^{-1} CC' \Sigma^{-1}}{1 - C' \Sigma^{-1} C}$$

which holds since C is $N \times 1$. Substituting equation A3 into the above, we have:

$$(\Sigma - CC')^{-1} = \Sigma^{-1} + \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D}$$

So

$$\begin{aligned} \alpha^* &= (\mu + Ca)' \left(\Sigma^{-1} + \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \right) (\mu + Ca) \\ &= \left(\mu + \Sigma^{\frac{1}{2}} D a \right)' \left(\Sigma^{-1} + \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \right) \left(\mu + \Sigma^{\frac{1}{2}} D a \right) \\ &= \alpha + \mu' \Sigma^{-\frac{1}{2}} D a + \left(\Sigma^{\frac{1}{2}} D a \right)' \Sigma^{-1} \mu + \left(\Sigma^{\frac{1}{2}} D a \right)' \Sigma^{-1} \left(\Sigma^{\frac{1}{2}} D a \right) + \\ &\quad \left(\Sigma^{\frac{1}{2}} D a \right)' \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \mu + \left(\Sigma^{\frac{1}{2}} D a \right)' \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \mu + \left(\Sigma^{\frac{1}{2}} D a \right)' \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \left(\Sigma^{\frac{1}{2}} D a \right) + \mu' \frac{\Sigma^{-\frac{1}{2}} D D' \Sigma^{-\frac{1}{2}}}{1 - D' D} \mu \end{aligned}$$

So

$$\begin{aligned}
E[U|a] &= \frac{1}{2\lambda} \left(\alpha + 2\mu' \Sigma^{-\frac{1}{2}} D a + D' D a^2 + \frac{2a D' D D' \Sigma^{-\frac{1}{2}} \mu}{1 - D' D} + \frac{a^2 (D' D)^2}{1 - D' D} + \frac{(\mu' \Sigma^{-\frac{1}{2}} D)^2}{1 - D' D} \right) \\
&= \frac{1}{2\lambda} \left(a^2 \left(D' D + \frac{(D' D)^2}{1 - D' D} \right) + a \left(2\mu' \Sigma^{-\frac{1}{2}} D + \frac{2D' D D' \Sigma^{-\frac{1}{2}} \mu}{1 - D' D} \right) + \left(\alpha + \frac{(\mu' \Sigma^{-\frac{1}{2}} D)^2}{1 - D' D} \right) \right) \quad A4
\end{aligned}$$

which is quadratic in a .

Taking expectations over a , using $E[a] = 0$, $E[a^2] = 1$, we arrive at the formulae below:

Proposition 1

The unconditional expected utility of the mean-variance investor under the assumptions of a single forecasting variable, a ($M=1$), no estimation error and in the absence of a budget constraint is given by

$$E[U] = \frac{1}{2\lambda} \left(\alpha + \frac{D' D + (\mu' \Sigma^{-\frac{1}{2}} D)^2}{1 - D' D} \right)$$

B. Proof of Proposition 2

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), no estimation error and in the absence of a budget constraint

Because C is a matrix in this case and D is symmetric we have:

$$(\Sigma - C C')^{-1} \cong \Sigma^{-\frac{1}{2}} (I - D^2)^{-1} \Sigma^{-\frac{1}{2}}$$

and

$$(I - D^2)^{-1} = \sum_{j=0}^{\infty} (D^2)^j$$

If D is a matrix with eigenvalues less than 1. Since D is assumed to be a diagonal matrix with correlations on the diagonal, this is satisfied.

Using $(\Sigma - C C')^{-1} \cong \Sigma^{-\frac{1}{2}} (I + D^2) \Sigma^{-\frac{1}{2}}$ which is an approximation to $O(D^4)$.

So

$$\begin{aligned}
\alpha^* &= (\mu + \Sigma^{\frac{1}{2}} D a)' \Sigma^{-\frac{1}{2}} (I + D^2) \Sigma^{-\frac{1}{2}} (\mu + \Sigma^{\frac{1}{2}} D a) \\
&= \alpha + 2\mu' \Sigma^{\frac{1}{2}} \Sigma^{-1} D a + (\Sigma^{\frac{1}{2}} D a)' \Sigma^{-1} \Sigma^{\frac{1}{2}} D a + \\
&\quad \mu' \Sigma^{-\frac{1}{2}} D^2 \Sigma^{-\frac{1}{2}} \mu + 2\mu' \Sigma^{-\frac{1}{2}} D^2 \Sigma^{-\frac{1}{2}} \Sigma^{\frac{1}{2}} D a + (\Sigma^{\frac{1}{2}} D a)' \Sigma^{-\frac{1}{2}} D^2 \Sigma^{-\frac{1}{2}} (\Sigma^{\frac{1}{2}} D a)
\end{aligned}$$

$$\begin{aligned}
&= \alpha + 2\mu' \Sigma^{-\frac{1}{2}} D a + a' D^2 a + 2\mu' \Sigma^{-\frac{1}{2}} D^3 a + a' D^4 a \\
\alpha^* &= \alpha + a \left(2\mu' \Sigma^{-\frac{1}{2}} D + 2\mu' \Sigma^{-\frac{1}{2}} D^3 \right) + a' (D^2 + D^4) a
\end{aligned} \tag{A5}$$

Taking expectations over a we have $E[a] = 0$, $E[a' X a] = \text{tr}(X)$

Proposition 2

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), no estimation error and in the absence of a budget constraint is approximated by

$$E[U] \cong \frac{\alpha + \text{tr}(D^2) + \text{tr}(D^4)}{2\lambda} + O(D^6)$$

C. Proof of Proposition 3

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , no estimation error and in the absence of a budget constraint using the conditional covariance

In this case:

$$C = \Sigma^{\frac{1}{2}} D = \Sigma^{\frac{1}{2}} d$$

as

$$\begin{aligned}
D &= d I_N \\
\Sigma^{*-1} &= (\Sigma - C C')^{-1} \\
\left(\Sigma - d^2 \Sigma^{\frac{1}{2}} I \Sigma^{\frac{1}{2}} \right)^{-1} &= \frac{\Sigma^{-1}}{1 - d^2}
\end{aligned}$$

Thus we have:

$$\begin{aligned}
\alpha^* &= (\mu + C a)' \Sigma^{*-1} (\mu + C a) \\
&= \frac{1}{1 - d^2} \left(\mu + \Sigma^{\frac{1}{2}} d a \right)' \Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} d a \right) \\
&= \frac{1}{1 - d^2} \left(\alpha + 2\mu' \Sigma^{-\frac{1}{2}} d a + \left(\Sigma^{\frac{1}{2}} d a \right)' \Sigma^{-1} \left(\Sigma^{\frac{1}{2}} d a \right) \right) \\
\alpha^* &= \frac{1}{1 - d^2} \left(\alpha + 2d\mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right)
\end{aligned} \tag{A6}$$

Taking expectations we have $E[a] = 0$, $E[a' X a] = \text{tr}(X)$:

Proposition 3

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, no estimation error and in the absence of a budget constraint using the conditional covariance is given by

$$E[U] = \frac{(\alpha + d^2 N)}{2\lambda(1 - d^2)}$$

D. Proof of Proposition 4

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the absence of a budget constraint using the unconditional covariance

The estimated weights are:

$$w = \frac{\Sigma^{-1}\mu}{\lambda}$$

The conditional moments are:

$$\begin{aligned}\mu^* &= \bar{x} + \Sigma^{\frac{1}{2}}da \\ \Sigma^* &= \frac{\Sigma}{(1 - d^2)} \\ \hat{w} &= \frac{\Sigma^{-1}(\bar{x} + \Sigma^{\frac{1}{2}}da)}{\lambda(1 - d^2)}\end{aligned}$$

The expected return of the mean-variance investor conditional on a is thus:

$$\begin{aligned}E[r_p|a] &= \frac{(\bar{x} + \Sigma^{\frac{1}{2}}da)' \Sigma^{-1} (\mu + \Sigma^{\frac{1}{2}}da)}{\lambda(1 - d^2)} \\ &= \frac{\bar{x}' \Sigma^{-1} \mu + d \bar{x}' \Sigma^{-\frac{1}{2}} a + d \mu' \Sigma^{-\frac{1}{2}} a + a' a d^2}{\lambda(1 - d^2)}\end{aligned}$$

Averaging over a we have:

$$E[r_p] = \frac{\alpha + Nd^2}{\lambda(1 - d^2)} \tag{A7}$$

Expected risk in this case is:

$$\begin{aligned}E[\sigma_p^2|a, \bar{x}] &= E(\omega' \Sigma^* \omega|a) = \frac{(\bar{x} + \Sigma^{\frac{1}{2}}da)' \Sigma^{-1} (\bar{x} + \Sigma^{\frac{1}{2}}da)}{\lambda^2(1 - d^2)} \\ &= \frac{\bar{x}' \Sigma^{-1} \bar{x} + 2 \bar{x}' \Sigma^{-\frac{1}{2}} a d + a' a d^2}{\lambda^2(1 - d^2)}\end{aligned}$$

$$E[\sigma_p^2] = \frac{\alpha + \frac{N}{T} + Nd^2}{\lambda^2(1 - d^2)} \quad A8$$

Proposition 4

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the absence of a budget constraint using the conditional covariance is given by

$$E[U] = \frac{\alpha - \frac{N}{T} + Nd^2}{2\lambda(1 - d^2)}$$

Appendix B – Utility, Optimal Weights and the Budget Constraint

A. *Proof of Proposition 5*

Optimal mean-variance weights in the presence of a budget constraint

The Lagrangian of the budget constrained constant absolute risk aversion investor is as follows:

$$U = \omega' \mu - \frac{\lambda}{2} \omega' \Sigma \omega - \theta(\omega' i - 1) \quad A9$$

where θ is the Lagrange multiplier.

Our first order conditions are:

$$\mu - \lambda \Sigma \omega - \theta i = 0$$

So

$$\omega = \frac{1}{\lambda} \Sigma^{-1} (\mu - \theta i)$$

$$i' \omega = \frac{1}{\lambda} (i' \Sigma^{-1} \mu - \theta i' \Sigma^{-1} i)$$

Let

$$\alpha = \mu' \Sigma^{-1} \mu \quad A10$$

$$\beta = \mu' \Sigma^{-1} i \quad A11$$

$$\gamma = i' \Sigma^{-1} i \quad A12$$

In the literature, α is often referred to as the squared Sharpe ratio.

$\alpha > 0$, $\beta > 0$ and $\alpha\gamma - \beta^2 > 0$ for μ not proportional to i by the Cauchy-Schwarz inequality.

So

$$1 = \frac{1}{\lambda} (\beta - \theta\gamma)$$

$$\beta = \lambda + \theta\gamma$$

Thus

$$\theta = \frac{(\beta - \lambda)}{\gamma}$$

Finally, we have the optimal weights in the presence of a budget constraint

Proposition 5

The optimal mean-variance weights in the presence of a budget constraint with known parameters is given by

$$\omega = \frac{1}{\lambda} \Sigma^{-1} \mu - \frac{(\beta - \lambda)}{\lambda\gamma} \Sigma^{-1} \mathbf{1} \tag{A13}$$

It has been brought to our attention that this result was also given by Jorion (1985) and in Ingersoll (1987)²⁹

We can compute our expected utility V_c , as follows:

$$\begin{aligned} V_c &= \omega' \mu - \frac{\lambda}{2} \omega' \Sigma \omega \\ &= \frac{\alpha}{\lambda} - \frac{\beta(\beta - \lambda)}{\lambda\gamma} - \frac{\lambda}{2\lambda^2} \left(\mu' - \frac{(\beta - \lambda)}{\gamma} \mathbf{1}' \right) \Sigma^{-1} \left(\mu - \frac{(\beta - \lambda)}{\gamma} \mathbf{1} \right) \\ &= \frac{\beta}{\gamma} - \frac{\beta(\beta - \lambda)}{\lambda\gamma} - \frac{1}{2\lambda} \left(\alpha - \frac{2\beta(\beta - \lambda)}{\gamma} + \frac{(\beta - \lambda)^2}{\gamma} \right) \\ &= \frac{\alpha\gamma - \beta^2}{\lambda\gamma} + \frac{\beta}{\gamma} - \frac{1}{2\lambda\gamma} (\alpha\gamma - (\beta - \lambda)(\beta + \lambda)) \\ &= \frac{\alpha\gamma - \beta^2}{\lambda\gamma} + \frac{\beta}{\gamma} - \frac{\alpha\gamma - \beta^2}{2\lambda\gamma} - \frac{\lambda}{2\lambda} \\ V_c &= \frac{\alpha\gamma - (\beta - \lambda)^2}{2\lambda\gamma} \tag{A14} \end{aligned}$$

Appendix C – Derivation of Propositions: The Budget Constrained Case

A. Proof of Corollary 7

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, no estimation error and in the presence of a budget constraint using the conditional covariance matrix.

We begin by restating the conditional moments:

²⁹ See chapter 5, appendix A, expression A.5.

$$\mu^* = \mu + Ca$$

$$\Sigma^* = (\Sigma - CC')$$

where

$$C = d\Sigma^{\frac{1}{2}}$$

and the optimal weight relation in the presence of a budget constraint:

$$\omega = \frac{1}{\lambda} \Sigma^{-1} \mu - \frac{(\beta - \lambda)}{\lambda \gamma} \Sigma^{-1} \mathbf{i}$$

where

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

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

Substituting the conditional moments into the optimal weight relation yields:

$$\omega = \frac{1}{\lambda(1-d^2)} \left(\Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\beta + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2) \right)}{\gamma} \Sigma^{-1} \mathbf{i} \right)$$

The expected return of the mean-variance portfolio conditional on a is:

$$\begin{aligned} E[r_p|a] &= \frac{1}{\lambda(1-d^2)} \left(\Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\beta + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2) \right)}{\gamma} \Sigma^{-1} \mathbf{i} \right)' \left(\mu + \Sigma^{\frac{1}{2}} da \right) \\ &= \frac{1}{\lambda(1-d^2)} \left(\left(\mu + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} da \right) - \frac{(\beta + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2))}{\gamma} \left(\mu + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} \mathbf{i} \right) \\ &= \frac{1}{\lambda(1-d^2)} \left(\left(\mu' \Sigma^{-1} \mu + d\mu' \Sigma^{-\frac{1}{2}} a + d\mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right) \right. \\ &\quad \left. - \frac{\left(\beta + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2) \right)}{\gamma} \left(\beta + da' \Sigma^{-\frac{1}{2}} \mathbf{i} \right) \right) \end{aligned}$$

$$E[r_p|a] = \frac{1}{\lambda(1-d^2)} \left(\left(\mu' \Sigma^{-1} \mu + d\mu' \Sigma^{-\frac{1}{2}} a + d\mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right) \right. \\ \left. - \frac{\beta^2 + 2\beta da' \Sigma^{-\frac{1}{2}} i + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a - \lambda(1-d^2)\beta - \lambda(1-d^2)da' \Sigma^{-\frac{1}{2}} i}{\gamma} \right)$$

Taking expectations over a , we have:

$$E[r_p] = \frac{1}{\lambda(1-d^2)} \left((\alpha + Nd^2) - \frac{\beta(\beta - \lambda(1-d^2)) + d^2 \text{tr}(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}})}{\gamma} \right)$$

$$\text{Now } \text{tr}(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}}) = \text{tr}(i \Sigma^{-1} i)$$

$$\text{tr}(\gamma) = \gamma, \text{ so}$$

$$E[r_p] = \frac{1}{\lambda(1-d^2)} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))}{\gamma} \right)$$

The expected risk of the portfolio conditional on a is:

$$E[\sigma_p^2|a] = \frac{1}{\lambda^2(1-d^2)} \left(\left(\left(\mu + \Sigma^{\frac{1}{2}} da \right) - \frac{(\beta + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2))}{\gamma} i \right)' \Sigma^{-1} \left(\left(\mu + \Sigma^{\frac{1}{2}} da \right) \right. \right. \\ \left. \left. - \frac{(\beta + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2))}{\gamma} i \right) \right) \\ = \frac{1}{\lambda^2(1-d^2)} \left(\left(\mu + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} \left(\mu + \Sigma^{\frac{1}{2}} da \right) - 2 \frac{(\beta + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2))}{\gamma} \left(\mu + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} i \right. \\ \left. + \frac{(\beta + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2))^2}{\gamma^2} i' \Sigma^{-1} i \right)$$

$$\begin{aligned}
&= \frac{1}{\lambda^2(1-d^2)} \left(\mu' \Sigma^{-1} \mu + 2d\mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right. \\
&\quad - 2 \frac{(\beta^2 + 2\beta d a' \Sigma^{-\frac{1}{2}} i + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a - \lambda(1-d^2)\beta - \lambda(1-d^2) d a' \Sigma^{-\frac{1}{2}} i)}{\gamma} \\
&\quad \left. + \frac{\beta^2 + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a + \lambda^2(1-d^2)^2 + 2\beta d a' \Sigma^{-\frac{1}{2}} i - 2\beta \lambda(1-d^2) - 2\lambda(1-d^2) d a' \Sigma^{-\frac{1}{2}} i}{\gamma} \right)
\end{aligned}$$

Taking expectations over a , we have

$$\begin{aligned}
&= \frac{1}{\lambda^2(1-d^2)} \left(\alpha + Nd^2 - 2 \frac{\beta(\beta - \lambda(1-d^2)) + d^2 \text{tr}(a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a)}{\gamma} \right. \\
&\quad \left. + \frac{(\beta - \lambda(1-d^2))^2 - d^2 \text{tr}(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}})}{\gamma} \right)
\end{aligned}$$

$$E[\sigma_p^2] = \frac{1}{\lambda^2(1-d^2)} \left(\alpha + (N-1)d^2 + \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right)$$

Thus,

$$\begin{aligned}
E[U] &= \frac{1}{\lambda(1-d^2)} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))}{\gamma} \right) \\
&\quad - \frac{\lambda}{2} \left(\frac{1}{\lambda^2(1-d^2)} \left(\alpha + (N-1)d^2 + \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \right) \\
&= \frac{1}{\lambda(1-d^2)} \left(\left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))d^2}{\gamma} \right) - \frac{1}{2} \left(\alpha + (N-1)d^2 + \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \right) \\
&= \frac{1}{2\lambda(1-d^2)} \left(\alpha + (N-1)d^2 - \frac{2\beta(\beta - \lambda(1-d^2))}{\gamma} - \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \\
E[U] &= \frac{\alpha\gamma + (N-1)d^2\gamma - (\beta - \lambda(1-d^2))^2}{2\lambda(1-d^2)\gamma}
\end{aligned}$$

Corollary 7

The unconditional expected utility of the mean-Variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint using the conditional covariance matrix is given by

$$E[U] = \frac{\alpha + (N-1)d^2 - \frac{(\beta - \lambda(1-d^2))^2}{\gamma}}{2\lambda(1-d^2)}$$

B. Proof of Proposition 6

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the conditional covariance matrix

To determine the joint impact of estimation error and forecasting on expected utility in the presence of a budget constraint, we substitute the conditional moments (14) and (15) into the optimal weight relation:

$$\omega = \frac{1}{\lambda} \Sigma^{-1} \mu - \frac{(\beta - \lambda)}{\lambda \gamma} \Sigma^{-1} \mathbf{i}$$

So this becomes

$$w = \frac{1}{\lambda(1-d^2)} \left(\Sigma^{-1} \left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\bar{x}' \Sigma^{-1} \mathbf{i} + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2) \right)}{\gamma} \Sigma^{-1} \mathbf{i} \right)$$

The expected return of the mean-variance portfolio conditional on a and \bar{x} is:

$$\begin{aligned} E[r_p | a, \bar{x}] &= \frac{1}{\lambda(1-d^2)} \left(\Sigma^{-1} \left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\bar{x}' \Sigma^{-1} \mathbf{i} + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2) \right)}{\gamma} \Sigma^{-1} \mathbf{i} \right)' (\mu + \Sigma^{\frac{1}{2}} da) \\ &= \frac{1}{\lambda(1-d^2)} \left(\left(\bar{x} + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} (\mu + \Sigma^{\frac{1}{2}} da) - \frac{(\bar{x}' \Sigma^{-1} \mathbf{i} + da' \Sigma^{-\frac{1}{2}} \mathbf{i} - \lambda(1-d^2))}{\gamma} (\mu + \Sigma^{\frac{1}{2}} da)' \Sigma^{-1} \mathbf{i} \right) \\ &= \frac{1}{\lambda(1-d^2)} \left(\left(\bar{x}' \Sigma^{-1} \mu + d \bar{x}' \Sigma^{-\frac{1}{2}} a + d \mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right) \right. \\ &\quad \left. - \frac{\bar{x}' \Sigma^{-1} \mathbf{i} \beta + (\beta + \bar{x}' \Sigma^{-1} \mu) da' \Sigma^{-\frac{1}{2}} \mathbf{i} + d^2 a' \Sigma^{-\frac{1}{2}} \mathbf{i}' \Sigma^{-\frac{1}{2}} a - \lambda(1-d^2) \beta - \lambda(1-d^2) da' \Sigma^{-\frac{1}{2}} \mathbf{i}}{\gamma} \right) \end{aligned}$$

Taking expectations over both conditioning variables, we have:

$$E[r_p] = \frac{1}{\lambda(1-d^2)} \left((\alpha + Nd^2) - \frac{\beta(\beta - \lambda(1-d^2)) + d^2 \text{tr} \left(\Sigma^{-\frac{1}{2}} \mathbf{i} \mathbf{i}' \Sigma^{-\frac{1}{2}} \right)}{\gamma} \right)$$

$$\text{Now } \text{tr} \left(\Sigma^{-\frac{1}{2}} \mathbf{i} \mathbf{i}' \Sigma^{-\frac{1}{2}} \right) = \text{tr}(\mathbf{i}' \Sigma^{-1} \mathbf{i})$$

$$= \text{tr}(\gamma) = \gamma, \text{ so}$$

$$E[r_p] = \frac{1}{\lambda(1-d^2)} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))}{\gamma} \right)$$

The expected risk of the portfolio conditional on a is:

$$\begin{aligned} E[\sigma_p^2|a, \bar{x}] &= \frac{1}{\lambda^2(1-d^2)} \left(\left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) - \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2) \right)}{\gamma} i \right)' \Sigma^{-1} \left(\left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) \right. \\ &\quad \left. - \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2) \right)}{\gamma} i \right) \\ &= \frac{1}{\lambda^2(1-d^2)} \left(\left(\bar{x} + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} \left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) - 2 \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2) \right) \left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i \right)}{\gamma} \right. \\ &\quad \left. + \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda(1-d^2) \right)^2}{\gamma^2} i' \Sigma^{-1} i \right) \\ &= \frac{1}{\lambda^2(1-d^2)} \left(\bar{x}' \Sigma^{-1} \bar{x} + 2d \bar{x}' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right. \\ &\quad - 2 \frac{\left(\bar{x}' \Sigma^{-1} i i' \Sigma^{-1} \bar{x} + d \bar{x}' \Sigma^{-1} i a' \Sigma^{-\frac{1}{2}} i + d \bar{x}' \Sigma^{-1} i a' \Sigma^{-\frac{1}{2}} i + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a - \lambda(1-d^2) \bar{x}' \Sigma^{-1} i - \lambda(1-d^2) da' \Sigma^{-\frac{1}{2}} i \right)}{\gamma} \\ &\quad \left. + \frac{\bar{x}' \Sigma^{-1} i i' \Sigma^{-1} \bar{x} + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a + \lambda^2(1-d^2)^2 + 2 \bar{x}' \Sigma^{-1} i da' \Sigma^{-\frac{1}{2}} i - 2 \bar{x}' \Sigma^{-1} i \lambda(1-d^2) - 2 \lambda(1-d^2) da' \Sigma^{-\frac{1}{2}} i}{\gamma} \right) \end{aligned}$$

Taking expectations over both conditioning variables, we have

$$\begin{aligned} &= \frac{1}{\lambda^2(1-d^2)} \left(\alpha + \frac{N}{T} + Nd^2 - 2 \frac{\left(\beta^2 + \frac{\text{tr}(i \Sigma^{-1} i)}{T} - \beta \lambda(1-d^2) + d^2 \text{tr} \left(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} \right) \right)}{\gamma} \right. \\ &\quad \left. + \frac{\left(\beta - \lambda(1-d^2) \right)^2 + \frac{\text{tr}(i \Sigma^{-1} i)}{T} - d^2 \text{tr} \left(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} \right)}{\gamma} \right) \end{aligned}$$

Now $\text{tr}(i'\Sigma^{-1}i)$ and $\text{tr}\left(\Sigma^{-\frac{1}{2}}ii'\Sigma^{-\frac{1}{2}}\right)$ both equal γ , thus:

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

and,

$$\begin{aligned} E[U] &= \frac{1}{\lambda(1-d^2)} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))}{\gamma} \right) \\ &\quad - \frac{\lambda}{2} \left(\frac{1}{\lambda^2(1-d^2)} \left(\alpha + (N-1) \left(d^2 - \frac{1}{T} \right) + \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \right) \\ &= \frac{1}{\lambda(1-d^2)} \left(\left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda(1-d^2))d^2}{\gamma} \right) \right. \\ &\quad \left. - \frac{1}{2} \left(\alpha + (N-1) \left(d^2 - \frac{1}{T} \right) + \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \right) \\ &= \frac{1}{2\lambda(1-d^2)} \left(\alpha - (N-1) \left(d^2 - \frac{1}{T} \right) - \frac{2\beta(\beta - \lambda(1-d^2))}{\gamma} - \frac{\lambda^2(1-d^2)^2 - \beta^2}{\gamma} \right) \\ E[U] &= \frac{\alpha + (N-1) \left(d^2 - \frac{1}{T} \right) - \frac{(\beta - \lambda(1-d^2))^2}{\gamma}}{2\lambda(1-d^2)} \end{aligned}$$

Proposition 6

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the conditional covariance matrix is given by

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

B. Proof of Proposition 7

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the unconditional covariance Matrix.

Instead of substituting the conditional covariance matrix into the optimal weight expression (A13), we now use the unconditional covariance matrix.

The optimal weight expression, conditional on a and \bar{x} becomes

$$w = \frac{1}{\lambda} \left(\Sigma^{-1} (\bar{x} + \Sigma^{\frac{1}{2}} da) - \frac{(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda)}{\gamma} \Sigma^{-1} i \right)$$

The expected return of the mean-variance portfolio conditional on a and \bar{x} is:

$$\begin{aligned} E[r_p|a, \bar{x}] &= \frac{1}{\lambda} \left(\Sigma^{-1} (\bar{x} + \Sigma^{\frac{1}{2}} da) - \frac{(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda)}{\gamma} \Sigma^{-1} i \right)' (\mu + \Sigma^{\frac{1}{2}} da) \\ &= \frac{1}{\lambda} \left((\bar{x} + \Sigma^{\frac{1}{2}} da)' \Sigma^{-1} (\mu + \Sigma^{\frac{1}{2}} da) - \frac{(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda)}{\gamma} (\mu + \Sigma^{\frac{1}{2}} da)' \Sigma^{-1} i \right) \\ &= \frac{1}{\lambda} \left((\bar{x}' \Sigma^{-1} \mu + d\bar{x}' \Sigma^{-\frac{1}{2}} a + d\mu' \Sigma^{-\frac{1}{2}} a + d^2 a' a) \right. \\ &\quad \left. - \frac{\bar{x}' \Sigma^{-1} i \beta + (\beta + \bar{x}' \Sigma^{-1} i) da' \Sigma^{-\frac{1}{2}} i + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a - \lambda \beta - \lambda da' \Sigma^{-\frac{1}{2}} i}{\gamma} \right) \end{aligned}$$

Taking expectations over both conditioning variables, we have:

$$E[r_p] = \frac{1}{\lambda(1-d^2)} \left((\alpha + Nd^2) - \frac{\beta(\beta - \lambda) + d^2 \text{tr}(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}})}{\gamma} \right)$$

$$\text{Now } \text{tr}(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}}) = \text{tr}(i' \Sigma^{-1} i)$$

$$= \text{tr}(\gamma) = \gamma, \text{ so}$$

$$E[r_p] = \frac{1}{\lambda} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda)}{\gamma} \right)$$

The expected risk of the portfolio conditional on a is:

$$\begin{aligned} E[\sigma_p^2|a, \bar{x}] &= \frac{1}{\lambda^2} \left((\bar{x} + \Sigma^{\frac{1}{2}} da) - \frac{(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda)}{\gamma} \Sigma^{-1} i \right)' \left((\bar{x} + \Sigma^{\frac{1}{2}} da) \right. \\ &\quad \left. - \frac{(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda)}{\gamma} \Sigma^{-1} i \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\lambda^2} \left(\left(\bar{x} + \Sigma^{\frac{1}{2}} da \right)' \Sigma^{-1} \left(\bar{x} + \Sigma^{\frac{1}{2}} da \right) - 2 \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda \right) \left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i \right)}{\gamma} \right. \\
&\quad \left. + \frac{\left(\bar{x}' \Sigma^{-1} i + da' \Sigma^{-\frac{1}{2}} i - \lambda \right)^2}{\gamma^2} i' \Sigma^{-1} i \right) \\
&= \frac{1}{\lambda^2} \left(\bar{x}' \Sigma^{-1} \bar{x} + 2d \bar{x}' \Sigma^{-\frac{1}{2}} a + d^2 a' a \right. \\
&\quad - 2 \frac{\left(\bar{x}' \Sigma^{-1} i i' \Sigma^{-1} \bar{x} + d \bar{x}' \Sigma^{-1} i a' \Sigma^{-\frac{1}{2}} i + d \bar{x}' \Sigma^{-1} i a' \Sigma^{-\frac{1}{2}} i + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a - \lambda \bar{x}' \Sigma^{-1} i - \lambda da' \Sigma^{-\frac{1}{2}} i \right)}{\gamma} \\
&\quad \left. + \frac{\bar{x}' \Sigma^{-1} i i' \Sigma^{-1} \bar{x} + d^2 a' \Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} a + \lambda^2 + 2 \bar{x}' \Sigma^{-1} i da' \Sigma^{-\frac{1}{2}} i - 2 \bar{x}' \Sigma^{-1} i \lambda - 2 \lambda da' \Sigma^{-\frac{1}{2}} i}{\gamma} \right)
\end{aligned}$$

Taking expectations over both conditioning variables, we have

$$\begin{aligned}
&= \frac{1}{\lambda^2(1-d^2)} \left(\alpha + \frac{N}{T} + Nd^2 - 2 \frac{\left(\beta^2 + \frac{tr(i' \Sigma^{-1} i)}{T} - \beta \lambda + d^2 tr \left(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} \right) \right)}{\gamma} \right. \\
&\quad \left. + \frac{(\beta - \lambda)^2 + \frac{tr(i' \Sigma^{-1} i)}{T} - d^2 tr \left(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} \right)}{\gamma} \right)
\end{aligned}$$

Now $tr(i' \Sigma^{-1} i)$ and $tr \left(\Sigma^{-\frac{1}{2}} i i' \Sigma^{-\frac{1}{2}} \right)$ both equal γ , thus:

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

and,

$$\begin{aligned}
E[U] &= \frac{1}{\lambda} \left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda)}{\gamma} \right) - \frac{\lambda}{2} \left(\frac{1}{\lambda^2} \left(\alpha + (N-1) \left(d^2 - \frac{1}{T} \right) + \frac{\lambda^2 - \beta^2}{\gamma} \right) \right) \\
&= \frac{1}{\lambda} \left(\left((\alpha + (N-1)d^2) - \frac{\beta(\beta - \lambda)d^2}{\gamma} \right) - \frac{1}{2} \left(\alpha + (N-1) \left(d^2 - \frac{1}{T} \right) + \frac{\lambda^2 - \beta^2}{\gamma} \right) \right) \\
&= \frac{1}{2\lambda(1-d^2)} \left(\alpha - (N-1) \left(d^2 - \frac{1}{T} \right) - \frac{2\beta(\beta - \lambda)}{\gamma} - \frac{\lambda^2 - \beta^2}{\gamma} \right)
\end{aligned}$$

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

Proposition 7

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, estimation error and in the presence of a budget constraint using the unconditional covariance Matrix is given by

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

Appendix D – Proposition 6 under a Simplified Model of Return Generation

A. Proof of Corollary 9

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint with a constant pair-wise correlation, ρ , and a constant stock volatility, σ , across all assets using the conditional covariance matrix

Proposition 6 as derived in appendix C is given by

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

For the $1/N$ investor we have

$$V_{\frac{1}{N}} = \hat{\mu}_1 - \frac{\lambda}{2N^2} i' \Sigma i$$

where

$$\hat{\mu}_1 = \sum \frac{\mu_i}{N}, \hat{\mu}_2 = \sum \frac{\mu_i^2}{N}, \hat{\sigma}_\mu^2 = \hat{\mu}_2 - \hat{\mu}_1^2$$

we assume that $\text{plim}_{N \rightarrow \infty} \hat{\mu}_1 = \mu_1$ and so forth.

To explicate the key drivers of expected utility, we assume a constant correlation, ρ , and constant variance, σ^2 , structure such that

$$\Sigma = \sigma^2(1 - \rho)I_N + \sigma^2 p i i'$$

where p is an $(N \times 1)$ vector of ones.

Using the standard inversion result, we can show that

$$\Sigma^{-1} = \frac{1}{\sigma^2(1 - \rho)} \left(I_N - \frac{\rho i i'}{1 + \rho(N - 1)} \right)$$

In what follows we drop the sample operator, $\hat{\cdot}$.

Now

$$\begin{aligned}
\alpha &= \mu' \Sigma^{-1} \mu \\
&= \frac{1}{\sigma^2(1-\rho)} \left(N\mu_2 - \frac{\rho N^2 \mu_1^2}{1+\rho(N-1)} \right) \\
&= \frac{1}{\sigma^2(1-\rho)} \left(\frac{(N^2 - N)\mu_2 \rho - N^2 \mu_1^2 \rho}{1+\rho(N-1)} \right) \\
&= \frac{1}{\sigma^2(1-\rho)} \left(\frac{N^2 \rho \sigma_\mu^2 - N \rho \mu_2}{1+\rho(N-1)} \right) \\
&= N^2 \frac{\sigma_\mu^2}{\sigma^2} (1-\rho) + O(1)
\end{aligned}$$

By similar arguments

$$\begin{aligned}
\gamma &= i' \Sigma^{-1} i \\
&= \frac{N(1+\rho(N-1)) - \rho N^2}{\sigma^2(1-\rho)(1+\rho(N-1))} \\
&= \frac{N(1-\rho)}{\sigma^2(1-\rho)(1+\rho(N-1))} \\
&= \frac{1}{\rho \sigma^2} + O\left(\frac{1}{N}\right)
\end{aligned}$$

Likewise,

$$\begin{aligned}
\beta &= \frac{1}{\sigma^2(1-\rho)} \left(\frac{N\mu_1(1-\rho)}{1+\rho(N-1)} \right) \\
&= \frac{\mu_1}{\sigma^2 \rho}
\end{aligned}$$

Thus,

Corollary 9

The unconditional expected utility of the mean-variance investor under the assumptions of multiple forecasting variables, a ($M=N$), with a constant forecasting ability level, d , estimation error and in the presence of a budget constraint with a constant pair-wise correlation, ρ , and a constant stock volatility, σ , across all assets using the conditional covariance matrix

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

Now

$$\begin{aligned}
V_{\frac{1}{N}} &= \mu_1 - \frac{\lambda}{2N^2} i' \Sigma i \\
&= \mu_1 - \frac{\lambda}{2N^2} (N\sigma^2 + (N^2 - N)\rho\sigma^2) \\
&= \mu_1 - \frac{\lambda\sigma^2}{2N} - \frac{\lambda\rho\sigma^2}{2} + \frac{\lambda\rho\sigma^2}{2N} \\
&= \mu_1 - \frac{\lambda\rho\sigma^2}{2} + O\left(\frac{1}{N}\right)
\end{aligned}$$

Thus, utility is increasing in N , for

$$T > \frac{\sigma^2}{d^2\sigma^2 + \sigma_u^2(1 - \rho)}$$

We can also conclude that as σ and ρ increase, expected utility decreases. Further, as cross-sectional dispersion, σ_u increases, expected utility increases.

Appendix E – Derivation of the Coefficient of Risk Aversion

To ensure that our results are robust across the risk spectrum, we use three levels of risk aversion, representing a conservative, a balanced and an aggressive investor. Kritzman (2011) for a different purpose, shows how we can infer the level of risk aversion using actual investor allocations. For an investor that allocates to bonds and equities the constant absolute risk aversion utility function can be expressed as:

$$E(U) = w_e\mu_e + w_b\mu_b - \frac{\lambda}{2}(\sigma_e^2w_e^2 + \sigma_b^2w_b^2 + 2\rho\sigma_e\sigma_bw_e w_b)$$

The marginal utility is maximised by equating the partial derivatives:

$$\frac{\partial E[U]}{\partial x_e} = \mu_e - \lambda(\sigma_e^2w_e + \rho\sigma_e\sigma_bw_b)$$

$$\frac{\partial E[U]}{\partial x_b} = \mu_b - \lambda(\sigma_b^2w_b + \rho\sigma_e\sigma_bw_e)$$

$$\lambda = (\mu_e - \mu_b)/(\sigma_e^2w_e + \rho\sigma_e\sigma_bw_b - \sigma_b^2w_b - \rho\sigma_e\sigma_bw_e)$$

To calibrate equation x we use long-dated US government bonds and the S&P 500 since 1927. The average allocation for balanced US mutual funds is 50% equities and 50% fixed income and cash³⁰. Morningstar, the fund rating service, define a conservative (aggressive) fund as a 20-50% (70-90%) allocation to equities, and a 50-80% (10-30%) allocation to fixed income and cash. We use the mid-points of these ranges to give the allocations of the conservative and aggressive investors. The resultant levels of λ are 0.015, 0.025, and 0.045 respectively.

³⁰ Ibbotson and Kaplan (2000)

Appendix F – Derivation of the Practitioner’s Utility Function

In this appendix we show that maximising the constant absolute risk aversion (CARA) utility function is equivalent to maximising the quadratic utility function.

The CARA utility function, is defined as:

$$U = -e^{-\lambda w}$$

where $\lambda > 0$ equals the Arrow-Pratt coefficient of absolute risk aversion.

The utility function is positively sloped and concave indicating risk aversion.

$$U'(C) = \lambda e^{-\lambda w} > 0$$

$$U''(C) = -\lambda^2 e^{-\lambda w} < 0$$

If wealth, w , is normally distributed:

$$\tilde{w} \sim N(\mu_w, \sigma_w^2)$$

Then the certainty equivalent, CE , can be derived as follows:

$$-e^{-\lambda CE} = -E[e^{-\lambda \tilde{w}}]$$

$$= -e^{-\lambda \mu_w + \frac{\lambda^2}{2} \sigma_w^2}$$

Therefore:

$$-\lambda CE = -\lambda \mu_w + \frac{\lambda^2}{2} \sigma_w^2$$

$$CE = \mu_w - \frac{\lambda}{2} \sigma_w^2$$

Appendix G – DeMiguel Data Sets

Table G
DeMiguel, Garlappi and Uppal (2009) data sets

Table G summarised the data sets that we have used to replicated the DeMiguel, Garlappi and Uppal (2009) findings.

Data set	Description	Source	Period
1. International	The “International” data set consists of the excess total monthly returns in USD of eight international equity indices: Canada, France, Germany, Italy, Japan, Switzerland, the UK, and the U.S. The data set also includes a Global developed market index.	MSCI total return indices Bloomberg	1/1970-10/2013
2. Industries	The “Industries” data set consists of the excess total monthly returns in USD of 10 value-weighted U.S sectors: Non-durable, Durable, Manufacturing, Energy, Tech, Telecommunications, Shops, Health, Utilities and Other. The data set also includes the Fama-French “MKT” factor.	K.R. French website	7/1926-7/2013
3. MKT/SMB/HML	The “MKT/SMB/HML” data set consists of the excess total monthly returns in USD of 10 value-weighted U.S sectors: Non-durable, Durable, Manufacturing, Energy, Tech, Telecommunications, Shops, Health, Utilities and Other.	K.R. French	7/1926-9/2013
4. FF-1-Factor	The “FF-1-Factor” data set consists of the excess total monthly returns in USD of the 20 value-weighted portfolios sorted by market capitalisation and book-to-market. The data set also includes the Fama-French “MKT” factor.	K.R. French	7/1926-9/2013
5. FF-3-Factor	The “FF-1-Factor” data set consists of the excess total monthly returns in USD of the 20 value-weighted portfolios sorted by market capitalisation and book-to-market. The data set also includes the Fama-French “MKT”, “SMB”, and “HML” factors.	Factset, Ex-share database	7/1926-9/2013
6. FF-4-Factor	The “FF-1-Factor” data set consists of the excess total monthly returns in USD of the 20 value-weighted portfolios sorted by	K.R. French	1/1927-7/2013

market capitalisation and book-to-market. The data set also includes the Fama-French “MKT”, “SMB”, and “HML” factors, and the “UMD” momentum factor of Carhart (1997)

	Internat.	Industry	MKT/SMB/ HML	FF-1- Factor	FF-3- factor	FF-4- factor
Assets, N	9	11	3	21	23	24
Estimation Window, T	526	601	603	603	603	603
<u>Information Coefficient, IC</u>						
$IC = 0$	✓	✓	✓	✓	✓	✓
$IC = 0.025$	✓	✓	✓	✓	✓	✓
$IC = 0.05$	✓	✓	✓	✓	✓	✓
$IC = 0.075$	✓	✓	✓	✓	✓	✓
$IC = 0.10$	✓	✓	✓	✓	✓	✓

References

- Allen, D, C Lizieri, and S Satchell, 2013, Quantifying the Non-Gaussian Gain, Cambridge University Working Paper in Economics
- Amenc, N., F. Goltz, and A. Lioui, 2011, Practitioner portfolio construction and performance measurement: Evidence from europe, *Financial Analysts Journal* 67, 39-50.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43-76.
- Asness, Cliff, Tobias Moskowitz, and Lasse Pedersen, 2010, Value and momentum everywhere, *University of Chicago and AQR Working Paper*.
- Banz, R. W., 1981, The relationship between return and market value of common-stocks, *Journal of Financial Economics* 9, 3-18.
- Black, F, and R Litterman, 1992, Global portfolio optimization, *Financial Analysts Journal*.
- Bloomfield, T, R Leftwich, and JB Long, 1977, Portfolio strategies and performance, *Journal of Financial Economics* 5, 201-218.
- Bollerslev, T, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307-327.
- Chan, L. K. C., J. Karceski, and J. Lakonishok, 1999, On portfolio optimization: Forecasting covariances and choosing the risk model, *Review of Financial Studies* 12, 937-974.

- Clarke, R., H. de Silva, and S. Thorley, 2002, Portfolio constraints and the fundamental law of active management, *Financial Analysts Journal* 58, 48-66.
- Cremers, J. H., M. Kritzman, and S. Page, 2005, Optimal hedge fund allocations - do higher moments matter?, *Journal of Portfolio Management* 31, 70
- Das, S. R., and R. Uppal, 2004, Systemic risk and international portfolio choice, *Journal of Finance* 59, 2809-2834.
- Debondt, W. F. M., and R. Thaler, 1985, Does the stock-market overreact, *Journal of Finance* 40, 793-805.
- Debondt, W. F. M., and R. H. Thaler, 1987, Further evidence on investor overreaction and stock-market seasonality, *Journal of Finance* 42, 557-581.
- DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915-1953.
- Ding, Z, 2010, The fundamental law of active management: Time series dynamics and cross-sectional properties (Fuller and Thaler Asset management, Working Paper).
- Duchin, Ran, and Haim Levy, 2009, Markowitz versus the Talmudic portfolio diversification strategies, *Journal of Portfolio Management* 35, 71
- Engle, R. F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United-Kingdom inflation, *Econometrica* 50, 987-1007.
- Fama, E. F., and K. R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.
- Frost, P. A., and J. E. Savarino, 1988, For better performance - constrain portfolio weights, *Journal of Portfolio Management* 15, 29
- Graham, Benjamin, and David Dodd, 1934. *Security analysis* (McGraw-Hill).
- Grauer, R. R., and F. C. Shen, 2000, Do constraints improve portfolio performance?, *Journal of Banking & Finance* 24, 1253-1274.
- Grinold, R. C., 1989, The fundamental law of active management, *Journal of Portfolio Management* 15, 30-37.
- Ibbotson, Roger, and Paul Kaplan, 2000, Does asset allocation policy explain 40, 90, or 100 of performance?, *Financial Analysts Journal* Jan/Feb, 2000.
- Jagannathan, R., and T. S. Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651-1683.
- Jegadeesh, N., 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.
- Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers - implications for stock-market efficiency, *Journal of Finance* 48, 65-91.
- Jobson, J. D., and B. Korkie, 1980, Estimation for Markowitz efficient portfolios, *Journal of the American Statistical Association* 75, 544-554.
- Jobson, J. D., and B. M. Korkie, 1981, Performance hypothesis-testing with the sharpe and treynor measures, *Journal of Finance* 36, 889-908.
- Jorion, P., 1985, International portfolio diversification with estimation risk, *Journal of Business* 58.
- Jorion, P., 1991, Bayesian and capm estimators of the means - implications for portfolio selection, *Journal of Banking & Finance* 15, 717-727.
- Kan, Raymond, and Guofu Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621-656.
- Kroll, Y., H. Levy, and H. M. Markowitz, 1984, Mean-variance versus direct utility maximization, *Journal of Finance* 39, 47-61.
- Lakonishok, J., A. Shleifer, and R. W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

- Lehmann, B. N., 1990, Fads, martingales, and market-efficiency, *Quarterly Journal of Economics* 105, 1-28.
- Levy, H., and H. M. Markowitz, 1979, Approximating expected utility by a function of mean and variance, *American Economic Review* 69, 308-317.
- Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.
- Mandelbrot, B., 1963, The variation of certain speculative prices, *Journal of Business* 36, 394-419.
- Markowitz, Harry, 1952, Portfolio selection, *Journal of Finance* 7, 77-91.
- Mommel, Christoph, 2003, Performance hypothesis testing with the Sharpe ratio, *Finance Letters* 1.
- Merton, R. C., 1980, On estimating the expected return on the market - an exploratory investigation, *Journal of Financial Economics* 8, 323-361.
- Ross, S, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 20.
- Sharpe, W. F., 1964, Capital-asset prices - a theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.
- Simaan, Y., 1993, What is the opportunity cost of mean-variance investment strategies, *Management Science* 39, 578-587.
- Tu, Jun, and Guofu Zhou, 2011, Markowitz meets talmud: A combination of sophisticated and naive diversification strategies, *Journal of Financial Economics* 99, 204-215.
- Ye, Jia, 2008, How variation in signal quality affects performance, *Financial Analysts Journal* 64, 48-61.
- Zhou, Guofu, 2008, On the fundamental law of active portfolio management: What happens if our estimates are wrong?, *Journal of Portfolio Management* 34, 26
- Zweig, J, 1998, Five lessons from America's top pension fund, *Money*.