Asset Selection with Estimation Risk

Apostolos Kourtis* and Raphael N. Markellos†

First version: December 20, 2009

This version: September 12, 2010

Abstract

Diversification across a large number of assets is likely to lead to poor portfolio performance when portfolio weights are estimated in finite samples. We tackle this problem in a new manner by contracting the asset space to an optimal dimension. A generic asset selection framework is developed for deriving the subset of the asset space that maximises the expected out-of-sample performance of a portfolio strategy. We adapt this framework to three popular strategies: the sample mean-variance portfolio, the sample global minimum variance portfolio and the $1/N$ rule. A comprehensive out-of-sample study using various datasets and criteria demonstrates that asset selection results in dramatic improvements in the performance of the strategies considered. For realistic portfolio problem configurations, we find that optimal portfolios are small and include between 1 and 5 assets. Finally, we discuss how asset selection can provide an alternative rational explanation to the underdiversification, home-bias and asset allocation puzzles, respectively.

Keywords: Portfolio Choice, Asset Selection, Estimation Risk

JEL Classification: C13, C51, C61, G11

*Norwich Business School, University of East Anglia, Norwich, Norfolk, NR4 7TJ, UK a.kourtis@uea.ac.uk. Tel. +44 (0)1603 591387

†Corresponding Author: Department of Management Science and Technology, Athens University of Economics and Business, Office 915, 47A Evelpidon Str. 113 62, Athens, Greece; Visiting Research Fellow, Centre for Research in International Economics and Finance (CIFER), Loughborough University, UK. E-mail: *rmarkel@aueb.gr* ; Tel. +30 210 8203671; Fax. +30 210 8828078. We are grateful to John Cochrane, George Constantinides, George Dotsis, Harry Markowitz and participants at the 2010 Symposium on Frontiers in Financial Research for useful comments and suggestions. All remaining errors are our own.

"*The more perfect a nature is,*
the fewer means it requires for its operation."
Aristotle (350 B.C.)

Introduction

Diversification is the only "free lunch" in finance and can be used to reduce risk by simply spreading out investments across as many assets as possible. This is the dogma we teach students for more than 50 years since Markowitz published his pathbreaking study in 1952 on optimal diversification using the Mean-Variance Theory (MVT). Diversification is a straightforward, intuitive and powerful idea that is deeply rooted in our profession and has heavily influenced thinking and analysis in all areas of finance. The so-called diversification principle has attained almost universal applicability by transcending across most business disciplines, including, for example, strategy, marketing and human resources. Despite its apparent importance, the implementation of diversification has not been particularly successful. For example, MVT dictates that investors should hold proportions of the mean-variance efficient portfolio that allocates across all assets. However, there is widespread evidence that professional investors and private households systematically ignore this advice and tend to concentrate their wealth on a handful of assets. At the same time, numerous studies demonstrate that efficient portfolios are hard to be exploited since data limitations lead to parameter estimation errors and defuse the possibility of realising diversification gains.[1] Although considerable effort has been devoted in developing techniques that reduce the adversities of estimation risk, success has not been impressive.[2] As characteristically concluded by DeMiguel et al. (2009a) in an extensive recent horserace between naive diversification and several sophisticated portfolio strategies: "*there are still many miles to go before the gains promised by optimal portfolio choice can actually be realized out of sample*".

The present study attempts to reconcile academic research with investment practice on the issues of diversification and portfolio choice. We reaffirm the practical superiority of mean-variance portfolio optimisation. However, when estimation risk is present, we show that "less is more" in the sense that diversification benefits can be maximised only if an optimally selected subset of the asset space is considered. We demonstrate that increasing the number of assets beyond some point is not only useless, as is the orthodox view and popular belief, but leads to an increasingly suboptimal out-of-sample performance. This is because after some point the negative effects of estimation risk gradually overcome the marginal benefits of diversification. The addition of a new asset expands the number of parameters that have to be estimated and this in turn deducts valuable degrees of freedom. Estimation risk is the price you pay for diversification.

Existing portfolio choice models generally derive estimates of the portfolio weights and apply them to the entire asset space. However, empirical evidence suggests that the accuracy of these

[1] For instance, see Michaud (1989), Best and Grauer (1991), and Litterman (2003).
[2] See Brandt (2009) for an excellent survey on the literature of optimal portfolio choice.

estimates decreases with the number of available assets. For example, Tu and Zhou (2009) find that their portfolio rules perform better in smaller datasets and wonder if there is an optimal number of assets that improves performance. Such empirical evidence motivates our approach in reducing the number of assets prior to computing the portfolio weights. Removing assets from a portfolio is equivalent to imposing the constraint that the weights on the eliminated assets are zero. On the one hand, this constraint introduces specification error since the *ex ante* portfolio weights are generally non-zero. On the other hand, it reduces sampling errors since there are fewer weights to be estimated and more degrees of freedom become available. If the performance gains from reducing the estimation errors exceed the losses due to the introduction of specification error, the constrained asset space should be preferred over the initial space. Our approach optimises the performance of sample-based portfolios by utilising the trade-off between specification and estimation error, respectively.

We formulate under general conditions the optimal asset selection problem for a mean-variance investor, i.e., the problem of choosing the set of assets that maximises the expected out-of-sample performance of his portfolio strategy. We employ the average expected utility as a measure of out-of-sample performance due to its analytical tractability and to the straightforward correspondence it has to the investor's objective function. When the investment universe contains a small number of assets then the optimal subset can be easily determined in an exact manner. For larger asset spaces we provide a simple stepwise procedure that can be used to obtain a near optimal subset. Our framework is particularly attractive since it can be adapted to most portfolio choice models. Different performance measures such as the Sharpe ratio can also be readily incorporated. Even though we adopt the normal distribution for the exposition of our approach, modifications for alternative distributions are straightforward.

To illustrate the potential of our methodology, we first apply the asset selection framework to the sample mean-variance portfolio strategy. This is the 'textbook' strategy which under the assumption of parameter certainty produces optimal results for investors. We also adapt our framework to the sample-based global minimum variance portfolio and the equal-weighted $1/N$ rule, respectively. The first portfolio strategy is commonly suggested in the portfolio choice literature as an alternative to the sample tangency portfolio (e.g., see Jagannathan and Ma, 2003). Even though this portfolio is not designed to maximise the mean-variance objective, it is well known that it generally outperforms the tangency portfolio since it is less sensitive to estimation risk.[3] The second strategy is widely used by investors (see, for example, Benartzi and Thaler, 2001, and Huberman and Jiang, 2006). DeMiguel et al. (2009a) study the efficiency of the $1/N$ rule and find that it surprisingly leads to superior out-of-sample performance in comparison to many other sophisticated approaches. Although the two alternative strategies we consider have little or no sensitivity to estimation risk, our asset selection approach can improve their performance. This is because our approach aims to adapt these strategies to the mean-variance objective function of the investor and determine those

[3] An empirical comparison between these two portfolios is provided by Jagannathan and Ma (2003) and DeMiguel et al. (2009a). An analytical comparison can be found in Kourtis et al. (2009).

assets that maximise it. An interesting result we find here is that for the global minimum variance portfolio and the $1/N$ allocation, the optimal set of assets turns out to be a function of the coefficient of relative risk aversion of the investor. This is in contrast to the two-fund separation theorem of Tobin (1958) which states that all investors should hold the same composition of risky assets. As will be discussed in a later section of the paper, this result allows an alternative explanation to the asset allocation puzzle.

Several useful links can be drawn between our paper and the existing literature. Firstly, although asset selection is a longstanding approach in finance, we are the first to consider the effect of estimation risk in this context. Previous studies account for factors such as transaction costs (e.g., Mao, 1970; Jacob, 1974; Brennan 1975) or financial constraints such as margin requirements and restrictions on borrowing (e.g., Liu, 2009; Roche et al., 2009). Secondly, the literature on mean-variance spanning, initiated by Huberman and Kandel (1987), also investigates if a smaller subset of assets is preferable to the whole asset space. However, estimation risk is not considered since mean-variance spanning tests whether the *ex ante* weights on the eliminated assets are zero. Our paper shows how this hypothesis can be relaxed in the presence of parameter uncertainty. We particulary provide an upper bound for the weights on the eliminated assets. Thirdly, there are several studies that result in a reduced asset space by imposing constraints on the portfolio weights, such as no-short-selling, in order to reduce the undesirable effects of sampling errors (see, for example, Frost and Savarino, 1988; Jagannathan and Ma, 2003; DeMiguel et al., 2009b). Contrary to our approach, these constraints do not improve the performance of sample tangency portfolios significantly as Jagannathan and Ma indicate (2003). Nevertheless, our analysis offers an alternative theoretical explanation on why such approaches lead to improved portfolios since they result in a smaller set of assets. Finally, our approach is related to Britten-Jones (1999) who shows that mean-variance efficient weights can be considered as regression coefficients with the excess returns on the risky assets as the independent variables. Intuitively, optimal asset selection is analogous to explanatory variable selection in a linear regression model. In fact, the stepwise procedure we propose for the solution of the asset selection problem in large asset spaces is inspired from this analogy.

In our empirical analysis, we evaluate the out-of-sample performance improvements obtained from optimal asset selection in four datasets of monthly actual returns corresponding to 5, 10, 25 and 50 assets, respectively. In each dataset, we solve the optimal asset selection problem for each portfolio strategy adopting four different sample sizes (120, 240, 360 and 480 months) and two risk aversion coefficients (1 and 3). We then investigate the performance of the strategies in the corresponding optimal set of assets in terms of expected utilities and their standard error as well as in terms of out-of-sample Sharpe ratios and certainty equivalent returns. Our results clearly illustrate that optimal asset selection can enhance considerably the out-of-sample performance of portfolio rules in the presence of estimation risk. This finding is consistent across both simulated and real datasets and across all considered measures. As expected, the improvement is more prominent when asset spaces are large and/or when sample sizes are small, because in these cases estimation errors are typically large.

Particularly noteworthy is the enhancement in the performance of the mean-variance strategy. For example, when the available risky assets are 50 and the estimation window contains 120 months, the application of the sample mean-variance strategy to the entire set of available assets produces an annualised expected utility of -1,539% and a Sharpe ratio of 14.67%. If the asset space is contracted to its optimal subset, the utility and the Sharpe ratio raise to 21.55% and 22.51%, respectively. We find that the improvement offered by our approach is equivalent to investing in the original asset space but with a sample size which is larger by a factor which ranges between 2 and 5. Performance comparisons between the sample tangency portfolio and the two alternatives considered allow another interesting conclusion to be drawn. In particular, optimal asset selection leads to sample mean-variance portfolios that outperform their competitors with a sample size that does not exceed 240 months. This result resolves the issue raised in the studies of Kourtis et. al. (2009) and DeMiguel et al. (2009a) who find that an unrealistically lengthy dataset is needed for the mean-variance rule to beat the minimum variance and $1/N$ portfolios, respectively (e.g, up to 6,000 monthly observations for 50 assets).

Our paper allows novel explanations to three longstanding puzzles in finance. First, it is well known in the literature that investors deviate from the MVT paradigm recommendations and concentrate their wealth in a small number of assets. Our analysis offers some justification to this practice. In particular, a striking recommendation that stems from our results is that the optimal application of mean-variance optimisation under estimation risk leads to underdiversification. This is because given that portfolio weights are typically computed using less than 20 years of monthly data, we find that the number of optimal assets is small for all strategies considered and varies from 1 to 5.[4] Second, several studies since French and Poterba (1991) demonstrate that U.S. investors defy the potential benefits of international diversification and exhibit what is known as 'home bias'. Our approach offers a rational explanation to this behaviour. We specifically apply our methodology to a dataset consisting of national market index returns for five major stock exchanges. Solving the optimal asset selection problem for the three portfolio strategies leads to a result surprisingly consistent with the existing empirical evidence. The optimal investment set for a U.S. investor who allocates using the mean-variance strategy contains only the domestic index if less than 45 years of monthly historical observations are used. The same result also applies to the $1/N$ strategy while the investor that adopts the minimum variance strategy places 74% of his wealth on the US index. Finally, an empirical inconsistency that can also be explained by optimal asset selection is the asset allocation puzzle introduced in Canner et al. (1997). They document that the more risk averse an investor is, the more he tends to allocate his wealth towards bonds rather than stocks. However, this finding contradicts the two-fund separation theorem. Using a simple example, we are able to demonstrate that optimal asset selection for the minimum variance portfolio and the naive diversification leads to results that are in line with this puzzle. Specifically,

[4] Interestingly, this result is in agreement with the largely forgotten views on asset allocation by Keynes who suggested investing heavily in only a few assets (Keynes, 1983). Recently, an insightful study by Boyle et al. (2009) also relates the views of Markowitz and Keynes on asset allocation relying on the concept of ambiguity aversion.

we consider two investors with different levels of risk aversion who invest in a market consisting of a bond and a stock. We then solve the optimal asset selection problem for the two aforementioned portfolios. Our results indicate that the more risk-tolerant investor should just hold the stock for both the minimum variance portfolio and the $1/N$ strategy. However, the optimal set of assets for the risk-averse investor also includes a large proportion of the bond.

Lastly, it is necessary to modify our asset selection approach in order to make it applicable to portfolio management applications, because the optimal set of assets is a function of unknown parameters. In accordance, we provide well-behaved estimates of these parameters and solve the (estimated) asset selection problem for the three portfolios in various simulated datasets. The solution also results in significant improvement in the portfolios' performance constituting optimal asset selection a powerful tool in the practice of investment management.

The remainder of the paper is organised as follows. Section I sets the asset selection framework for a mean-variance investor. Section II formulates the optimal asset selection problems for the three portfolio strategies under consideration. The solution to these problems along with implications for out-of-sample performance is investigated in various datasets and settings in Section III. Section IV demonstrates how our analysis and results can be used to explain the underdiversification puzzle, the home bias puzzle and the asset allocation puzzle, respectively. Section V provides directions for the practical implementation of our approach while Section VI concludes. The mathematical derivations are provided in the Appendix.

I The Optimal Asset Selection Problem

In this section, we develop the asset selection framework for a mean-variance investor that selects his portfolio at time T in a market which includes N risky assets along with a risk-free asset.[5] The vector of returns on the risky assets at time t is denoted by r_t, whereas the return on the risk-free is denoted by rf_t. Excess returns on the risky assets over the risk-free are given by $R_t = r_t - rf_t 1_N$, where 1_N is an N-dimensional vector of ones. In order to illustrate the implications of asset selection, we take on a standard convention in the literature and assume that R_t's are identically and independently distributed $(i.i.d.)$ according to a Gaussian density with mean μ and covariance matrix Σ.[6] Under the classical mean-variance framework of Markowitz (1952, 1959), the preferences of the investor are described by the quadratic expected utility function:[7]

$$U(w) = w'\mu - \frac{\gamma}{2} w'\Sigma w, \tag{1}$$

[5]When a risk-free asset is not present, one can replace it with a portfolio benchmark such as a stock index.

[6]Fama (1976) documents that the normal distribution is a good approximation for monthly returns data. Moreover, Tu and Zhou (2004) conclude that ignoring tails makes negligible difference to the mean-variance investor in terms of portfolio out-of-sample performance. Nevertheless, our framework can be modified to other distributional assumptions.

[7]See Constantinides and Malliaris (1995) for a review on the mean-variance theory and its implications.

where w are the weights assigned to the N risky assets and γ stands for the coefficient of the relative risk aversion for the investor. The weight on the risk-free is then $1 - w'1_N$. The portfolio weights which maximise the above objective function are given by

$$w^{\text{mv}} = \frac{1}{\gamma}\Sigma^{-1}\mu \tag{2}$$

and correspond to the maximum utility

$$U(w^{\text{mv}}) = \frac{1}{2\gamma}\mu'\Sigma^{-1}\mu = \frac{S_{\text{mv}}}{2\gamma}, \tag{3}$$

where $S_{\text{mv}} = \mu'\Sigma^{-1}\mu$ is the squared Sharpe ratio of the *ex ante* tangency portfolio w^{mv}.

However, the maximum utility is unattainable in practice since the parameters μ and Σ defining the optimal weights are unknown. Commonly, the investor will estimate the portfolio weights using a sample of T historical excess returns on the risky assets, i.e., a matrix of observations $I_T = (R_1 R_2 ... R_T)$. This estimation is performed through the investor's portfolio strategy p, defined as a function from the set of information matrices Λ to the portfolio weights space:

$$p : \Lambda \subseteq \mathbb{R}^N \times \mathbb{R}^T \to \mathbb{R}^N, \text{s.t. } p\left(I_T\right) = \hat{w}^p \tag{4}$$

with $\hat{w}^p$ being the vector of weights that the portfolio strategy assigns to the risky assets given I_T.

Clearly, the optimal strategy is $p\left(I_T\right) = w^{\text{mv}}$ that results in the *ex ante* tangency portfolio independently of the realised sample I_T. Since this portfolio is not known, the investor needs to resort to an alternative sample-based portfolio strategy. Sample-based strategies introduce estimation errors to the portfolio choice process and result in suboptimal out-of-sample performance. We evaluate this performance using the average expected utility of a portfolio choice strategy:

$$M\left(p\right) := E\left(U\left(\hat{w}^p\right)\right) = E\left(\left(\hat{w}^p\right)'\mu - \frac{\gamma}{2}\left(\hat{w}^p\right)'\Sigma\hat{w}^p\right), \tag{5}$$

where the expectation is taken under the true distribution of I_T. Our choice of $M\left(p\right)$ is due to its analytical tractability for several portfolio strategies and its connection with the investor's objective function (1). Examples of studies that use the average expected utility as a portfolio performance measure are, among others, Kan and Zhou (2007), Kourtis et al. (2009) and DeMiguel et al. (2009a). For brevity, hereafter we refer to $M\left(p\right)$ as expected utility rather than average expected utility.

For expositional purposes, we drop the I_T argument and represent strategies and portfolio weights in terms of sets of assets. We attach a natural number to each risky asset and symbolise the asset space with $A = \{1, 2, ..., N\}$. If $B \subseteq A$, $p\left(B\right) = \hat{w}^p\left(B\right) \in \mathbb{R}^N$ denotes the portfolio weights that are obtained by the strategy p under the constraint that the weights on the assets outside B are zero ($\hat{w}_i^p = 0$ for $i \in A - B$). Note that if $B = \emptyset$, the investor places all his wealth on the riskless asset, since $p\left(\emptyset\right) = 0_N$, whereas $p\left(A\right) = \hat{w}^p$. We are now able to define the

Optimal Asset Selection Problem: *The optimal asset selection problem for the mean-variance investor is the problem of finding the set of assets that maximises the expected out-of-sample performance of his portfolio strategy*:

$$\max_{B\subseteq A} M\left(p\left(B\right)\right) \tag{6}$$

The optimal asset selection problem has always a solution $B_*^p = \arg\max_{B\subseteq A} M\left(p\left(B\right)\right)$ since the feasible set, i.e., the power set of A, is discrete and finite.[8] As a result, the investor constrains the asset space set to B_*^p in order to maximise the performance of his portfolio. We also observe that the investor places all his wealth in the risk-free asset if the expected out-of-sample performance is negative for any set of risky assets ($B \subseteq A$, $B \neq \emptyset$), since $M\left(p\left(\emptyset\right)\right) = 0$.

II Optimal Asset Selection for Three Portfolio Strategies

This section applies the asset selection framework developed in the previous section to the sample-based mean-variance strategy, the sample-based global minimum variance strategy and the $1/N$ strategy, respectively.

A Asset Selection for the Sample-Based Mean-Variance Strategy

The sample tangency portfolio is implied from the computation of the optimal portfolio weights in (2) using the common Maximum Likelihood (ML) estimates of the two unknown parameters:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t \tag{7}$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})', \tag{8}$$

Then, the sample-based mean-variance strategy is defined as

$$p^{\text{mv}}\left(I_T\right) = \hat{w}^{\text{mv}} := \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu} \tag{9}$$

for $T > N$ in order for the covariance matrix to be invertible. It is well known in the asset allocation literature that this portfolio rule is highly sensitive to estimation risk and leads to severe underperformance from an out-of-sample perspective (see, for example, Jobson and Korkie, 1980, and Michaud, 1989). This performance is elegantly expressed by Kan and Zhou (2007) in means of the average expected utility of the investor:

$$M\left(p^{\text{mv}}\right) = \frac{\kappa\left(N,T\right)S_{\text{mv}}}{2\gamma} - \frac{\lambda\left(N,T\right)}{2\gamma}, \tag{10}$$

[8] Without loss of generality, we assume that this solution is unique. However, if the maximum is attained in more than one subsets of A, it is simple to impose additional conditions to ensure the uniqueness of B_*, for example, in terms of its cardinality (the number of assets contained in B_*) or the order of the asset indices.

where

$$\begin{aligned}\kappa(N,T) &= \frac{T}{T-N-2}\left(2-\frac{T(T-2)}{(T-N-1)(T-N-4)}\right)\\ \lambda(N,T) &= \frac{NT(T-2)}{(T-N-1)(T-N-2)(T-N-4)}.\end{aligned} \tag{11}$$

and $T > N+4$.

We adopt this expression to derive the asset selection problem for the sample-based mean variance strategy. If $B \subseteq A, B \neq \emptyset$, then the sample mean and covariance matrix of the excess returns $R_t(B)$ on the assets contained in B are

$$\begin{aligned}\hat{\mu}(B) &= \frac{1}{T}\sum_{t=1}^{T} R_t(B)\\ \hat{\Sigma}(B) &= \frac{1}{T}\sum_{t=1}^{T}\left(R_t(B)-\hat{\mu}(B)\right)\left(R_t(B)-\hat{\mu}(B)\right)'.\end{aligned} \tag{12}$$

The population counterparts of the two parameters will be denoted without the hat. It is also convenient to present our analysis by assuming that $B = \{1, 2, ..., |B|\}$ where $|B|$ is the cardinality of B, i.e., the number of assets in B.[9] The application of the sample mean-variance strategy to B results in the portfolio weights

$$\hat{w}^{\mathrm{mv}}(B) = \left(\left(\frac{1}{\gamma}\left(\hat{\Sigma}(B)\right)^{-1}\hat{\mu}(B)\right)', 0'_{N-|B|}\right)', \tag{13}$$

where $0_{N-|B|}$ is a vector of $N-|B|$ zeros. The average out-of-sample performance of the above portfolio is

$$M\left(p^{\mathrm{mv}}(B)\right) = \frac{\kappa(|B|,T)\,S_{\mathrm{mv}}(B)}{2\gamma} - \frac{\lambda(|B|,T)}{2\gamma}, \tag{14}$$

where $S_{\mathrm{mv}}(B) = \mu(B)'\left(\Sigma(B)\right)^{-1}\mu(B)$ is the squared Sharpe ratio of the *ex ante* tangency portfolio applied to B. We accept the convention that $S_{\mathrm{mv}}(\emptyset) = 0$ in order to account for the portfolio that consists of the riskless asset. Therefore, the optimal asset selection problem for the sample-based mean-variance strategy is $\max_{B\subseteq A} M\left(p^{\mathrm{mv}}(B)\right)$, or, equivalently

$$\max_{B\subseteq A}\left(\kappa(|B|,T)\,S_{\mathrm{mv}}(B) - \lambda(|B|,T)\right) \tag{15}$$

The above asset selection problem can be solved by computing the expected utility in (14) for all subsets of A. However, this approach is not suitable for large values of N, since it requires $2^N - 1$ computations of the inverse of $\Sigma(B)$ which determines $S_{\mathrm{mv}}(B)$. For example, when $N = 50$ the number of subsets is 10^{15}.To address this issue, we propose a stepwise solution of (15) in Appendix A. This procedure is based on the stepwise regression method of selecting independent variables

[9] We can always rearrange the indices which characterise the assets to express B in this form.

in a linear regression, first introduced by Efroymson (1960) and modified in the related literature thereafter. Interestingly, it turns out that the solution $B_*^{\mathrm{mv}} = \arg\max_{B \subseteq A} M\left(p^{\mathrm{mv}}\left(B\right)\right)$ is independent of γ. Therefore, the optimal assets are the same for all investors who allocate their contributions using the sample Markowitz portfolio regardless of their attitude towards risk. This outcome is in direct agreement with the two-fund separation theorem of Tobin (1958) which states that the relative proportions invested in each risky asset are identical for all investors. However, as will be discussed in the next two subsections, the two-fund separation theorem may not hold for alternative portfolio strategies.

Our asset selection approach utilises the connection of the expected utility $M\left(p^{\mathrm{mv}}\right)$ with S_{mv} and N as expressed in (14). For fixed T, a reduction in N through the application of the sample mean-variance strategy to a subset of A, actually increases the coefficients κ and λ, whereas it reduces the value of S_{mv}. The smaller set should be preferred to the original asset space when the improvement in the expected utility caused by the increase in κ and λ exceeds the deterioration due to the decrease in S_{mv}. As discussed in the introduction, this is a trade-off between portfolio specification error and estimation error. In order to provide some further interpretation of this balance, we represent the specification error introduced by reducing the asset space to B with the difference between the squared Sharpe ratios of the *ex ante* tangency portfolios in the whole asset space and in B, respectively:

$$SE\left(B\right) = S_{\mathrm{mv}} - S_{\mathrm{mv}}(B) \tag{16}$$

The specification error in (16) corresponds to the efficiency losses that the investor experiences under parameter certainty if he selects his portfolio using just B instead of all N available risky assets. Amongst the subsets of A with the same number of assets, the highest out-of-sample performance is obtained for the subset that minimises $SE\left(B\right)$. This is due to the equivalence of the asset selection problem (15) to the maximax optimisation problem:

$$\max_{0 \leq n \leq N} \left(\kappa\left(n, T\right) \max_{B \in C_n} S_{\mathrm{mv}}\left(B\right) - \lambda\left(n, T\right) \right), \tag{17}$$

where C_n is the set which contains all sets of n assets. The next proposition connects the specification error with the *ex ante* mean-variance weights on the eliminated assets.

Proposition 1 *The specification error generated by the elimination of a subset B of the asset space can be expressed in terms of the ex ante mean-variance weights and the covariance matrix as*

$$SE\left(A - B\right) = \gamma^2 \left(w_B^{mv}\right)' \left(\Omega_B\right)^{-1} w_B^{mv} \tag{18}$$

where $w_B^{mv} = \left(w_1^{mv}, ..., w_{|B|}^{mv}\right)'$, $\Sigma^{-1} = \left(\omega_{ij}\right)_{i,j=1,\ldots,N}$ *and* $\Omega_B = \left(\omega_{ij}\right)_{i,j=1,\ldots,|B|}$. *In the special case that only the i asset is eliminated:*

$$SE\left(A - \{i\}\right) = \frac{\left(\gamma w_i^{mv}\right)^2}{\omega_{ii}} \tag{19}$$

On the basis of the last equation we can infer that the asset that contributes the least to the efficiency of the *ex ante* mean-variance portfolio is associated with the smallest ratio $\frac{(w_i^{\text{mv}})^2}{\omega_{ii}}$. In the following proposition, we also provide the conditions satisfied for a subset of the asset space that leads to an improvement in the out-of-sample performance of the sample tangency portfolio.

Proposition 2 *Define* $\phi(N) = 2(N+2) + \sqrt{2(N+4)(N-1)}$

1. *If* $T \leq \phi(N)$*, then the contraction of the asset space to any subset of* A *will improve the out-of-sample performance of the sample-based mean-variance strategy.*

2. *If* $T > \phi(N)$*, then the contraction of the asset space to a subset* B *will improve the out-of-sample performance of the sample-based mean-variance strategy if*

$$SE(B) < \delta_1(|B|, T)S_{mv} + \delta_2(|B|, T) \tag{20}$$

where $\delta_1(|B|, T) = \left(1 - \frac{\kappa(N,T)}{\kappa(|B|,T)}\right)$, $\delta_2(|B|, T) = \left(\frac{\lambda(N,T) - \lambda(|B|,T)}{\kappa(|B|,T)}\right)$.

The first part of Proposition 2 indicates that when $T \leq \phi(N)$, the out-of sample performance of the mean-variance strategy is enhanced by removing any assets irrespectively of the specification error that this removal induces. For instance, if there are 50 assets, then for less than 15 years of monthly observations ($T = 180$), the contraction of the asset space to any subset will increase the performance of the sample tangency portfolio. The second part of Proposition 2 is more interesting since it reflects the trade-off between specification and estimation error exploited by our asset selection approach. Note that the right-hand side of (20) increases with the number of assets that are eliminated from the asset space. In this manner, it can be considered as a measure of the gains obtained from estimating fewer weights. Hence, (20) illustrates that the out-of-sample performance of the sample mean-variance strategy is improved in a subset of the asset space if the *ex ante* efficiency losses are less than the gains due to the reduction of estimation errors. In this context, we express the optimal asset selection problem as

$$\max_{B \subseteq A} \left(\delta_1(|B|, T)S_{\text{mv}} + \delta_2(|B|, T) - SE(B)\right) \tag{21}$$

Hence, the optimal set of assets maximises the out-of-sample performance by optimally exploiting the trade-off between specification and estimation error. Finally, the combination of Propositions 1 and 2 yields the following

Corollary 1 *The elimination of the* i *asset leads to improved out-of-sample performance if*

$$(w_i^{mv})^2 < \omega_{ii}\left(\delta_1(N-1, T)S_{mv} + \delta_2(N-1, T)\right) \tag{22}$$

Once the true moments of the excess returns on the risky assets are known, an asset can be eliminated from the portfolio without impact on its efficiency only if the *ex ante* weight on the asset is zero. The mean-variance spanning literature introduced by Huberman and Kandel (1987) has developed various statistical tests for this hypothesis. As Corollary 1 indicates, this hypothesis is relaxed under parameter uncertainty and the elimination of an asset results in increased performance if its squared weight is less than the upper bound provided in the right-hand side of (22).

B Asset Selection for the Sample-Based Global Minimum Variance Strategy

The asset allocation literature attributes most of the underperformance in sample mean-variance portfolios to estimation errors in the means rather than the variances and covariances (e.g., Chopra and Ziembra, 1993). Motivated by these findings, several studies suggest the use of the sample global minimum variance portfolio. The global minimum variance strategy does not involve the estimation of expected returns, since the portfolio weights are given by

$$p^{\min}\left(I_T\right) = \hat{w}^{\min} := \frac{\hat{\Sigma}^{-1} 1_N}{1_N' \hat{\Sigma}^{-1} 1_N}, \tag{23}$$

Jagannathan and Ma (2003) and DeMiguel et al. (2009a) compared the performance of the above portfolio and the sample mean-variance portfolio in terms of Sharpe ratios and certainty-equivalent returns in both real and simulated datasets and found that the first generally outperforms the second. Recently, Kourtis et al. (2009) also provide analytical support to these studies by first deriving the expected utility of the above strategy:

$$M\left(p^{\min}\right) = \mu_{\min} - \frac{\gamma}{2} v\left(N, T\right) \sigma_{\min}^2 \tag{24}$$

where $\mu_{\min} = \dfrac{\mu' \Sigma^{-1} 1_N}{1_N' \Sigma^{-1} 1_N}$, $\sigma_{\min}^2 = \dfrac{1}{1_N' \Sigma^{-1} 1_N}$ are respectively the mean and variance of excess returns on the minimum variance portfolio and $v\left(N, T\right) = \dfrac{(T-2)}{(T-N-1)}$ for $T > N+1$. They subsequently perform a comparative assessment of the performances of the sample tangency and global minimum variance portfolios using (14) and (24), respectively. The authors find that under most scenarios the sample tangency portfolio is preferable only when an unrealistically lengthy estimation window is available (for instance, 1800 months of returns for 50 risky assets).

In light of these studies, the global minimum variance portfolio is an attractive alternative to the mean-variance portfolio in the presence of estimation risk. It is then natural to apply the asset selection framework developed in Section I in order to examine if its performance can be further enhanced. This application actually adapts the global minimum variance portfolio to the mean-variance objective since this portfolio is constructed solely to minimise the portfolio variance. The corresponding optimal asset selection problem is then $\max_{B \subseteq A} M\left(p^{\min}\left(B\right)\right)$, or equivalently

$$\max_{B \subseteq A} \left(\mu_{\min}\left(B\right) - \frac{\gamma}{2} v\left(\left|B\right|, T\right) \sigma_{\min}^2\left(B\right)\right), \tag{25}$$

where $\mu_{\min}(B) = \dfrac{\mu(B)'(\Sigma(B))^{-1} 1_{|B|}}{1'_{|B|}(\Sigma(B))^{-1} 1_{|B|}}$ and $\sigma^2_{\min}(B) = \dfrac{1}{1'_{|B|}(\Sigma(B))^{-1} 1_{|B|}}$.

Similarly to the case of the mean-variance portfolio, the optimal subset of A, defined as $B^{\min}_* = \arg\max_{B\subseteq A} M\left(p^{\min}(B)\right)$, can be obtained by computing the expected utility for all subsets when N is small. When N is large the stepwise algorithm discussed previously can be adopted. However, the structure of the optimal subset now depends on γ, even though γ does not affect the original global minimum variance weights (23). This interesting result implies that investors with different levels of risk aversion may not choose the same assets to build a minimum variance portfolio. For example, consider an infinitely risk averse investor ($\gamma \to \infty$) and an infinitely risk tolerant investor ($\gamma = 0$). The optimal assets for the risk averse investor minimise the portfolio's variance, whereas the optimal assets for the risk tolerant investor maximise the portfolio's expected excess return. Consequently, the two-fund separation theorem does not apply for the mean-variance investor who uses the global minimum variance portfolio. In Section IV, we utilise this finding to shed some light on the asset allocation puzzle.

C Asset Selection for the 1/N Strategy

The equal weighted $1/N$ strategy allocates an even proportion of the wealth across each risky asset

$$p^{1/N}(I_T) = w^{1/N} = \frac{1_N}{N} \tag{26}$$

A large proportion of investors tend to follow this simple heuristic rule (Benartzi and Thaler, 2001; Huberman and Jiang, 2006), even though it has no theoretical basis. However, recent studies indicate that this naive diversification technique is not entirely irrational in the presence of parameter uncertainty. In particular, Duchin and Levy (2009) perform a comparison of the sample tangency portfolio and the $1/N$ rule and find that the second results in higher out-of-sample performance than the first in most scenarios examined. DeMiguel et al. (2009a) reach a similar conclusion for the sample-based Markowitz rule and many of its improvements that have been developed over the years. DeMiguel et al. also show that an extraordinarily large sample is required for the sample tangency portfolio to outperform the $1/N$ allocation (namely 3,000 months for a portfolio with 25 assets and 6,000 months for a portfolio with 50 assets). These findings question the practical value of the optimal portfolio choice and emphasise the need for a more effective treatment of estimation risk.[10]

In analogy to the previous section, we apply our optimal asset selection approach to the $1/N$ rule in order to impose the mean-variance objective and obtain performance improvements. The corresponding optimal asset selection problem is $\max_{B\subseteq A} M\left(p^{1/N}(B)\right)$, or, equivalently

$$\max_{B\subseteq A}\left(\mu_{1/N}(B) - \frac{\gamma}{2}\sigma^2_{1/N}(B)\right), \tag{27}$$

[10] In this direction, Tu and Zhou (2009) suggest to combine theoretical portfolio rules with the 1/N strategy.

where $\mu_{1/N}(B) = \frac{\mu(B)' 1_{|B|}}{|B|}$ and $\sigma^2_{1/N}(B) = \frac{1'_{|B|}\Sigma(B) 1_{|B|}}{|B|^2}$ correspond to the expectation and variance of the excess returns on the equal weighted portfolio of the assets in B. Note that the optimal subset $B^{1/N}_* = \arg\max_{B\subseteq A} M\left(p^{1/N}(B)\right)$ also depends on the risk aversion coefficient γ.

III Evaluation of Optimal Asset Selection

This section explores the potential of optimal asset selection for the three portfolio strategies discussed. To this end, we employ 4 datasets of monthly returns which are listed in Table I. These consist of a set of returns on 5 national stock indices obtained from MSCI, two sets acquired from the website of Kenneth French that contain returns on 10 industry portfolios and on 25 size and book-to-market portfolios, respectively and a set of returns on 50 randomly selected S&P 500 stocks obtained from Reuters. The 5 national indices have been used in a variety of studies examining international diversification (eg., see French and Poterba, 1991). The datasets comprising of 10 and 25 assets constitute common framework for testing portfolio choice models in the literature. Our selection of 50 individual stocks also allows us to test the performance of our approach on plain stocks rather than indices or portfolios. We derive excess returns on the assets in each dataset using the 30-day nominal US T-bill as a riskless asset that is also taken from Kenneth French's website. We also assume that the mean and the covariance matrix of the excess returns are the true values of these parameters. These values are applied to the formulas from Section II for the analytical derivation of the average expected utilities of the mean-variance investor for each strategy in the asset space and its subsets. We further solve the asset selection problem for each portfolio strategy using the all subsets method when $N = 5$ or 10. For $N = 25$ and 50, we employ the stepwise approach proposed to derive a "near optimal" set of assets. Unless otherwise stated, our analysis is undertaken assuming four sample sizes ($T = 120$, 240, 360 and 480) and two values for the risk aversion parameter ($\gamma = 1$ and 3).

A Solving the Optimal Asset Selection Problem for the Mean-Variance Strategy

As a first step in evaluating the potential of asset selection for the sample tangency portfolio, we study the out-of-sample performance changes due to the removal of one asset from the portfolio. Table II reports the annualised percentage expected utility in the case of 10 and 25 risky assets for the whole asset space $M(p^{\text{mv}}(A))$ along with the minimum, maximum and average utility for its subsets with N-1 assets $M(p^{\text{mv}}(A - \{i\}))$. When 120 observations are available, the magnitude of estimation risk is so large that eliminating any asset affects performance positively. For higher values of T, the out-of-sample performance of the mean-variance strategy is also improved in most cases. For example, for 40 years of monthly returns ($T = 480$), the performance is enhanced in 8 and 17 subsets of $N - 1$ assets when $N = 10$ and 25, respectively.

In Figures 1 ($N = 10$) and 2 ($N = 25$), we demonstrate that the theoretical gains from increasing diversification boundlessly turn into out-of-sample performance losses in the presence of parameter uncertainty. These figures depict the utility of the *ex ante* mean-variance portfolio $U\left(w^{\text{mv}}\left(B_n^{\text{mv}}\right)\right)$ and the average utility of its sample counterpart $M\left(p^{\text{mv}}\left(B_n^{\text{mv}}\right)\right)$ as a function of the number of assets n, where B_n^{mv} is the optimal subset of A of n assets ($B_n^{\text{mv}} = \arg\max_{B \in C_n} S_{\text{mv}}(B)$). As expected, the efficiency of the *ex ante* optimal portfolio increases with n in agreement with the dictates of mean-variance theory. In contrast, the performance of the sample tangency portfolio attains its maximum for a relatively small number of assets and then declines.

The second and the third rows of all panels in Table III report the average expected utilities of the investor in annualised and percentage terms for the sample mean-variance strategy in the original asset space (Sample mv (A)) and its optimal subset (Sample mv (B_*^{mv})), respectively. In this manner, we assess the extent to which the sample Markowitz portfolio can be enhanced through the contraction of the asset space. As is evident from our results, the improvement in the performance is significant for all datasets and sample sizes. In line with past studies, the expected utility for the sample mean-variance strategy in the entire asset space (A) is low and in many cases negative. In contrast, when this strategy is applied to the optimal set of assets (B_*^{mv}), the corresponding utilities are positive and considerably higher. For example, when $N = 10$ and $\gamma = 1$ the average utility is negative for up to 30 years of historical monthly returns and equal to 4.81 when 40 years of data are available. However, when the asset space is reduced to its optimal subset, the utility becomes positive in all cases and raises to 13.55 for $T = 480$. Naturally, the improvement in the performance is more prominent for the larger values of N and for the smaller sample lengths T. Indicatively, for $N = 50$ and $T = 120$, the expected utility when the investor employs all N assets is -1,539, whereas the utility increases significantly to 21.55 when he invests in the optimal subset.

The fourth row of the panels of Table III reports the number of assets $(|B_*^{\text{mv}}|)$ that lead to optimal out-of-sample performance for the sample tangency portfolio. A first conclusion is that this number is substantially smaller than N due to the severe effects of estimation risk. In particular, the number of optimal assets varies from 1 to 3 for the smaller datasets and from 3 to 12 for the larger ones. To obtain a clearer picture with respect to the connection of the number of optimal assets with the sample length, Table IV also presents $|B_*^{\text{mv}}|$ as a function of T. As expected, the number of optimal assets increases with the length of the estimation window. Interestingly, the investor includes all N assets in his portfolio only if an unrealistically lengthy sample is available (53,368 to 1.3 million monthly observations). For realistic values of T, the number of optimal assets is in agreement with the actual numbers of assets held by investors as is reported in the existing literature. We return to this finding and discuss it further in Section IV.

As mentioned, our asset selection approach is equivalent to an expansion of the available sample, since it increases the degrees of freedom available for estimating portfolio weights. Table V illustrates this equivalence by presenting the number of observations needed for the sample tangency portfolio to perform in the whole asset space as well as it performs in its optimal subset when the sample size is T. In fact, the performance gains from optimal asset selection are equivalent to the gains from a sample size increase by approximately a factor which ranges between 2 and 5.

B Solving the Optimal Asset Selection Problem for the Global Minimum Variance Strategy

We revisit Table III to discuss the results concerning the impact of our approach on the performance of the global minimum variance strategy. As expected, the performance of the *N-assets* sample minimum variance portfolio (Sample min (A)) is not as sensitive to estimation errors as the sample mean-variance portfolio. Specifically, the reported expected utilities $M\left(p^{\min}(A)\right)$ for the sample-based portfolio are close to the utilities $U\left(w^{\min}(A)\right)$ of the population counterpart (*ex ante* min (A)). However, one must keep in mind that asset selection aims also aligning the performance of this portfolio with the mean-variance preferences of the investor. From this perspective, we find that asset selection improves considerably the sample minimum variance portfolio (Sample min $\left(B_*^{\min}\right)$), especially for the larger values of N considered. For example, in the 25 size and book-to-market portfolios dataset, the expected utility for the asset-space portfolio is 10 when $T = 120$ and $\gamma = 1$, whereas it raises to 17.28 in the optimal subset $B_*^{\min}$. Surprisingly, the improvement from asset selection is so large that the *optimal-subset* portfolio manages to greatly outperform the *ex ante N-assets* portfolio. As previously, the number of optimal assets ($\left|B_*^{\min}\right|$) reported in the 8th row is much smaller than the asset space. We also observe that the size of the optimal subset is not affected by changes in the length of the sample due to the small sensitivity of this portfolio to sampling errors. However, for $N = 50$ the number of optimal assets appears to vary with γ. This finding will be further discussed in the next section.

C Solving the Optimal Asset Selection Problem for the $1/N$ Strategy

The last three rows in the panels of Table III respectively report the expected out-of-sample performance of the $1/N$ strategy in the asset space A and its optimal subset $B_*^{1/N}$ as well as the size of this subset ($\left|B_*^{1/N}\right|$). The conclusions from these results are similar to those drawn for the minimum variance strategy. In particular, the performance of the $1/N$ strategy is also enhanced in the optimal subset and the size of this subset is again much smaller than N. This improvement is more significant for larger values of N. For example, when $N = 25$ and $\gamma = 3$, the expected utilities raise from 4.32 in A to 8.02 in $B_*^{1/N}$. Furthermore, we find that for $N = 50$ the optimal subset also varies with γ.

D Comparison between the Mean-Variance Strategy and its Two Alternatives

As discussed in Section II, several studies show that the sample Markowitz portfolio is underperformed by the global minimum variance and the $1/N$ strategies. It is therefore instructive to ask whether optimal asset selection can render mean-variance portfolios superior to these two competitors. As Figure 3 illustrates, the answer is positive. In panel A, we depict the expected utilities of the mean-variance ($M\left(p^{\mathrm{mv}}\left(B_n^{\mathrm{mv}}\right)\right)$ and the minimum variance strategy $M\left(p^{\min}\left(B_n^{\min}\right)\right)$ versus the number of assets n (assuming $T = 240$ and $\gamma = 1$). In panel B, we replace the utility of the minimum variance portfolio with the utility of the $1/N$ strategy $M\left(p^{1/N}\left(B_n^{1/N}\right)\right)$. The utilities for each strategy are computed for the respective optimal set of n assets ($B_n^p = \underset{B \in C_n}{\arg\max}\, M\left(p\left(B\right)\right)$). Interestingly, we find that while the mean-variance rule is underperformed when $n > 4$, it outperforms the minimum variance and the $1/N$ portfolios for $n \leq 4$.

Kourtis et. al. (2009) and DeMiguel et al. (2009a) showed that a very lengthy dataset is needed for the mean-variance rule to beat the minimum variance and the $1/N$ portfolios, respectively. In tables VI and VII, we compute again these lengths for the four datasets we considered with the modification that the Markowitz rule is applied to the optimal subset of A, B_n^{mv}, of size $n = 1,...,9$ and N. In line with the aforementioned studies, a large number of observations is found to be necessary when there is no asset selection. For example, when $N = 50$ and $\gamma = 3$, the required T for the Markowitz rule to outperform the minimum variance and the $1/N$ rules is 635 and 638, respectively. In contrast, when optimal asset selection is employed, the length assumes a more realistic value varying between 32 and 254. This result re-establishes the value of mean-variance optimisation in practical applications.

Our final observation regarding the comparative performance of the three portfolios concerns the parameter of risk aversion. From the results in Table III, we observe that it is more difficult for the mean-variance strategy to beat the minimum variance and $1/N$ strategies, respectively for larger values of γ in the smaller datasets. For $\gamma = 3$, the *optimal-subset* mean-variance strategy is outperformed by the *optimal-subset* minimum variance and $1/N$ strategies for $N = 5$ and 10. However, for $\gamma = 1$ and for the remaining cases with $\gamma = 3$, the *optimal-subset* mean-variance strategy is superior to its alternatives in almost all scenarios.

E Further Performance Results

E.1 Sharpe Ratios

Our formulation of the asset selection problem adopted the average expected utility as a measure of out-of-sample performance. It is also interesting to study the potential of our methodology using alternative performance metrics, such as the Sharpe ratio. For this purpose, we generate 10,000 sets of size T (120, 240, 360 and 480) of returns for each of the 4 aforementioned sets of actual returns. These returns are normally *i.i.d.* distributed with the mean μ and covariance matrix Σ calibrated to the respective sample moments of the excess returns in the original sets. In each simulated set, we derive the portfolio weights $\hat{w}^p\left(B\right)$ for each strategy in the whole asset space ($B = A$) and its optimal subset ($B = B_*^p$) and compute the corresponding out-of-sample Sharpe ratios by

$$\widehat{\text{SR}}_p\left(B\right) = \frac{\hat{w}^p\left(B\right)' \mu}{\sqrt{\hat{w}^p\left(B\right)' \Sigma \hat{w}^p\left(B\right)}} \tag{28}$$

The averages of these Sharpe ratios are reported in percentage terms in Table VIII for $\gamma = 3$. Even though the optimal subsets are constructed on the basis of maximising expected utilities, the improvement in the out-of-sample Sharpe ratio is analogous to the improvement in the utilities in Table III. The gains from asset selection are large for all strategies and become more significant for the larger asset spaces. For example, for $N = 50$, the Sharpe ratios of the sample tangency portfolio, the global minimum variance portfolio and the naive diversification increase from 14.67, 14.42 and 16.04 in the entire asset space to 22.51, 24.27 and 24.76, respectively, in its optimal subset.

We also note that for $N = 5$ or 10, the highest Sharpe ratios are obtained for the *optimal-subset* minimum variance portfolio. These are even higher than the Sharpe ratios obtained for the *N-assets ex ante* global minimum variance portfolio. Yet, the *optimal-subset* mean-variance portfolio is superior to its competitors when $N = 25$ and 50. In Figure 4, we also plot the evolution of the Sharpe ratio of the sample tangency portfolio against the number of optimal assets (in analogy to Figures 1 and 2). The graph is also inverse U-shaped with the maximum obtained for a small number of assets.

E.2 Standard Errors

It is also necessary to analyze the uncertainty in performance obtained via the optimal asset selection approach. Table IX reports the standard errors of the expected utilities whose averages are given in Table III. These errors are computed using the simulated sets we employed for the computation of the Sharpe ratio. The standard errors for the *ex ante* strategies and the $1/N$ rule are omitted from the Table, since these are zero. We find that asset selection significantly reduces the standard errors of the expected utilities for both the mean-variance and minimum variance strategies. For instance, for $T = 120$ and $N = 50$, the standard error for the tangency and minimum variance portfolios declines from 250 to 5.7 and from 1.55 to 0.36, respectively.

E.3 Results from Real Data

Finally, it is also useful to investigate the potential of our approach in a distribution-free setting. To this end, we follow a common practice in the portfolio choice literature and assess the asset selection methodology in actual datasets. Interesting results can be drawn in this context since, as shown by DeMiguel et al. (2009a), the sample mean-variance rule and many of its extensions can produce losses on a risk-adjusted basis when real data are used. Adopting the rolling window method followed by DeMiguel et al. (2009a), we first derive the sample mean $\hat{\mu}_t$ and covariance matrix $\hat{\Sigma}_t$ using the returns for the months t-T,..., t-1, $t > T$. We then use these estimates to calculate the weights w_t^p for each strategy p in the entire asset space ($B = A$) and in its optimal subset ($B = B_*^p$). For both cases, we calculate the respective portfolio return $R_t^p(B) = (w_t^p)' R_t(B)$ in each month leading to a series of returns, with mean and variance denoted by $\hat{\mu}_p(B)$ and $\hat{\sigma}_p^2(B)$. We measure out-of-sample performance using the certainty equivalent return on each portfolio strategy:

$$CER_p(B) = \hat{\mu}_p(B) - \frac{\gamma}{2}\hat{\sigma}_p^2(B). \tag{29}$$

We report the results in Table X for T =120 and 240 and $\gamma = 1$ and 3. The improvement in the performance is similar to the normality case for all strategies. Contrary to the *N-assets* mean-variance portfolio, the optimised portfolio leads to positive certainty equivalent returns in all but one case and outperforms the $1/N$ and the minimum variance portfolios in most scenarios.

IV Relation to Empirical Evidence

A Underdiversification

Contrary to what Modern Portfolio Theory dictates, a large body of empirical evidence indicates that investors typically hold a small number of assets. Early studies include Lease et al., (1974) and Blume and Friend, (1975). In particular, Blume and Friend report that the majority of investors hold up to 2 stocks. More recently, Kelly (1995) and Polkovnichenko (2005) employ household data from the Survey of Consumer Finances and evince a similar investment pattern. Specifically, Kelly finds that the median portfolio consists of one stock while Polkovnichenko reports that the median number of directly held stocks has increased from 2 in 1983 to 3 in 2001, with most households holding one to five stocks. Barber and Odean (2000), Ivković et al. (2004) and Goetzman and Kumar (2008) use data from a large U.S. discount brokerage and find that the average investor holds just four stocks. At an international level, several studies included in Guiso et al. (2002) confirm the small number of stocks in household portfolios for different countries. Finally, Calvet et al. (2007) find portfolio underdiversification using a dataset that reflects the entire population of Sweden.

A number of potential explanations for this underdiversification puzzle have been suggested in the literature. These include transaction costs (Mao, 1970; Jacob, 1974; Brennan, 1975), higher moment preferences (Kraus and Litzenberger, 1974; Conine and Tamarkin, 1981; Harvey and Siddique, 2000; Mitton and Vorkink, 2007), information only about a fraction of all available assets (Merton, 1987), preference for portfolio insurance (Liu, 2009), financial constraints such as margin requirements and restrictions on borrowing (Roche et al., 2009) and information acquisition (Van Nieuwerburgh and Veldkamp, 2010). Explanations have also been based on behavioral biases such as preference towards a two layer-portfolio for poverty protection and hope for riches (Shefrin and Statman, 2000), rank-dependent preferences (Polkovnichenko, 2005) and ambiguity aversion (Uppal and Wang, 2003; Boyle et. al. 2009).

Our work provides an alternative rational justification to underdiversification. Specifically, our results imply that investors should not diversify across many assets because estimation errors will offset diversification benefits and this may result in heavily suboptimal portfolio performance. Given that portfolio weights are typically computed using less than 240 months, we found in Section III that the number of optimal assets ranges from 1 to 5. This finding is strikingly consistent with the empirical evidence on existing portfolio practices. Moreover, our explanation of underdiversification offers three significant advantages. First, we do not deviate from the classical mean-variance framework. Second, our approach encapsulates estimation risk, the latter being a crucial obstacle to the practical implementation of portfolio theory. Third, even though it is beyond the scope of the present work, factors such as transaction costs and financial constraints can also be incorporated into our model to further investigate empirical patterns such as the connection of the level of diversification with the investor's wealth.

B International Diversification

French and Poterba (1991) report that the proportion of wealth allocated to domestic assets for the U.S. investor is 92%. More recently, Thomas, et al. (2006) find that international exposure of equity portfolios of US investors has not considerably increased since then and investors allocate just 14% of their wealth to foreign assets. These findings are in contrast to the theoretical benefits of international diversification which imply that the mean-variance US investor should just hold around 60% in US stocks (Lewis, 1999). As a result a large body of research tries to rationalise this so-called international diversification puzzle (see, for example, Adler and Dumas, 1983; Black, 1990; Brennan and Cao, 1997; Serrat, 2001; Ang and Bekaert, 2002; Dahlquist et al. 2004; Guidolin, 2005; Guidolin and Timmermann, 2008)

In this context, we argue that our approach can provide a justification to the behaviour of investors with respect to international diversification. In particular, we solve the optimal asset selection problem for the mean-variance US investor who selects his portfolio in a market of 5 national stock indices corresponding to large international stock markets (namely France, Germany, Japan, UK and USA). The parameters that define the asset selection problem are calibrated using a dataset of monthly excess returns on the 5 indices for the period 1/1970-12/2007. This dataset is obtained from MSCI and converted to US dollar values using the MSCI exchange rate series.[11] The outcome of our analysis is quite astonishing. Specifically, we find that for an estimation window that contains less than 45 years of monthly returns, the US investor who selects his portfolio using the sample mean-variance strategy should hold just the domestic index. The same result applies for the equal-weighted strategy irrespective of the sample size. The optimal set for the global minimum variance portfolio also includes a foreign index. However, the weight assigned to the domestic asset is still high at around 74%. These findings justify the preference of US investors towards their home market.

C The Asset Allocation Puzzle

Canner et al. (1997) document that financial advisors recommend a high ratio of bonds to stocks for the conservative investors. Conversely, a low ratio is recommended for aggressive investors. The authors coin this empirical pattern as an "asset allocation puzzle" since it contradicts the two-fund separation theorem. Chakroun et al (2008) use data on the portfolio composition of a brokerage firm and confirm this empirical irregularity in actual allocations. A commonly suggested solution to this puzzle is intertemporal hedging activity (see Brennan and Xia, 2000, 2002; Bajeux-Besnainou

[11] This is actually the dataset used in the performance evaluation of Section III.

et al., 2003). Yet, Lioui (2007) shows that this solution is not convincing under general assumptions and finds that a highly risk averse investor with intertemporal hedging demands could invest more in stocks than in bonds.

We briefly discuss here how optimal asset selection can provide a potential solution to this puzzle by studying a mean-variance investor who allocates using the global minimum variance portfolio or the naive diversification. As we discussed in Section II, the first portfolio is proposed in many studies as a superior alternative to the mean-variance portfolio under parameter uncertainty. The second portfolio rule is a popular allocation method for many investors as various empirical studies report. To illustrate our point, we use a simple numerical example. Consider two assets to replicate a typical bond-stock choice situation. The excess return on the first asset ("bond") is assumed to have an annual mean of 4% and volatility of 10% while the second ("stock") is assumed to have an annual excess return of 8% and volatility of 20%. The covariance between the two assets is assumed to be zero. We solve the asset selection problem for two values of γ (1 and 3) and two values of T (120 and 240) for each one of the portfolio strategies studied in Section II. The optimal sets of assets, portfolio weights and expected utilities obtained are reported in Table XI. We can note that the less risk-averse investor holds only the stock for both the minimum variance and the $1/N$ strategies. Interestingly though, when risk aversion parameter increases to 3, the bond enters into the portfolio for both strategies with weight of 0.8 and 0.5, respectively. This result indicates that asset selection under parameter uncertainty may result in a ratio of bonds to stocks which increases with risk aversion. To the best of our knowledge, this is the first rational explanation to the asset allocation puzzle in a static setting.

V Implementation of Optimal Asset Selection in Practice

We have shown that optimal asset selection can significantly improve the out-of-sample performance of portfolio strategies and we have provided the conditions that should hold for the optimal set of assets. The implementation of asset selection in practical applications requires the knowledge of the average expected utilities in each subset of assets and this section provides suggestions as to their estimation.

For the sample-based mean-variance strategy, the average expected utility of the investor given in (14) is a function of the squared Sharpe ratio $S_{\mathrm{mv}}(B)$. Hence, in practice, one has to estimate $S_{\mathrm{mv}}(B)$. The ML estimate of $S_{\mathrm{mv}}(B)$ is $\hat{S}_{\mathrm{mv}}(B) = \hat{\mu}(B)' \left(\hat{\Sigma}(B)\right)^{-1} \hat{\mu}(B)$. However, this estimator is known to be heavily biased (e.g., see Kubokawa et al. 1993). In the Appendix, we derive an accurate estimate of $S_{\mathrm{mv}}(B)$ given by

$$\tilde{S}(B) = \zeta(|B|, T)\, \hat{S}_{\min}(B) \tag{30}$$

where

$$\hat{S}_{\min}(B) = \frac{\left(\hat{\mu}(B)'\left(\hat{\Sigma}\left(\mathrm{B}\right)\right)^{-1} 1_{|B|}\right)^2}{1'_{|B|}\left(\hat{\Sigma}\left(\mathrm{B}\right)\right)^{-1} 1_{|B|}} \tag{31}$$

is the ML estimator of the squared Sharpe ratio of the *ex ante* global minimum variance portfolio of the assets in B and

$$\zeta\left(|B|, T\right) = \frac{(T - |B| - 1)^2 (T - |B| - 2)^2}{T(T-1)(T-2)^2}. \tag{32}$$

Then, the estimated optimal asset selection problem is $\max_{B \subseteq A}\left(\kappa\left(|B|, T\right) \tilde{S}(B) - \lambda\left(|B|, T\right)\right)$ and $\hat{B}_*^{\mathrm{mv}}$ denotes its solution. To further improve the accuracy of the estimated optimal subset, we suggest to force an upper bound constraint $\bar{n}$ to its cardinality $\left|\hat{B}_*^{\mathrm{mv}}\right|$. To provide some interpretation for imposing this constraint, we note that the larger the number of assets in $\hat{B}_*^{\mathrm{mv}}$, the higher the probability that suboptimal assets have been included in the portfolio. Therefore, we induce an upper bound to the number of optimal assets to account for underperformance issues caused by the choice of wrong assets. Overall, we propose the estimation of the optimal assets by solving the following problem

$$\begin{gathered} \max_{B \subseteq A}\left(\kappa\left(|B|, T\right) \tilde{S}(B) - \lambda\left(|B|, T\right)\right) \\ |B| \leq \bar{n} \end{gathered} \tag{33}$$

We also use the standard ML estimates of the moments of the excess returns on the global minimum variance portfolio and the $1/N$ strategy to estimate the corresponding asset selection problems

$$\begin{gathered} \max_{B \subseteq A}\left(\hat{\mu}_{\min}\left(B\right) - \frac{\gamma}{2} v\left(|B|, T\right) \hat{\sigma}^2_{\min}\left(B\right)\right) \\ |B| \leq \bar{n} \end{gathered} \tag{34}$$

$$\begin{gathered} \max_{B \subseteq A}\left(\hat{\mu}_{1/N}\left(B\right) - \frac{\gamma}{2} \hat{\sigma}^2_{1/N}\left(B\right)\right) \\ |B| \leq \bar{n} \end{gathered} \tag{35}$$

The respective estimates of the optimal subsets will be denoted by $\hat{B}_*^{\min}$ and $\hat{B}_*^{1/N}$.

We apply the estimated optimal subsets to the corresponding portfolio strategies and result in three new strategies, namely $p^{\mathrm{mv}}\left(\hat{B}_*^{\mathrm{mv}}\right)$, $p^{\min}\left(\hat{B}_*^{\min}\right)$ and $p^{1/N}\left(\hat{B}_*^{1/N}\right)$. To assess the performance of the new strategies, we perform a comparison between each one of them and its asset space counterpart in a simulation setting. For each of the 4 datasets of realised returns in Table I, we generate 10,000 simulated datasets of size T (120, 240, 360 and 480) of normally distributed i.i.d. returns with means and covariances calibrated using the original datasets. In each set, we solve the asset selection problems (33), (34) and (35) by setting $\bar{n} = 2$ for $T = 120$ and 240 and $\bar{n} = 3$ for T = 360 and 480. We subsequently derive the portfolio weights $\hat{w}^p\left(B\right)$ for each strategy in the whole asset space ($B = A$) and its estimated optimal subset ($B = \hat{B}_*^p$) and measure the corresponding out-of-sample performance by the expected utility

$$U\left(\hat{w}^{p}\left(B\right)\right)=\hat{w}^{p}\left(B\right)'\mu-\hat{w}^{p}\left(B\right)'\Sigma\hat{w}^{p}\left(B\right) \tag{36}$$

Table XII reports the annualised percentage averages of the above measure for the 3 dyads of strategies considered. In agreement with Table III, the *N-assets* sample tangency portfolio $p^{\text{mv}}\left(A\right)$ generally leads to highly negative utilities and is greatly outperformed by the estimated subset rule $p^{\text{mv}}\left(\hat{B}_{*}^{\text{mv}}\right)$ which in most cases yields positive utility. For example, in the case of 25 assets, 120 historical observations and a risk aversion parameter of 1, the utility is -204.60 in the initial asset space whereas the corresponding utility for the subset-based portfolio $p^{\text{mv}}\left(\hat{B}_{*}^{\text{mv}}\right)$ is 41.86. The utilities for the minimum variance portfolio and the $1/N$ strategy also increase in most cases with the increase being more significant for the larger values of N. Overall, most conclusions drawn in Section III apply in this setting as well.

VI Conclusions

Wide diversification benefits are dissipated when finite samples are available for parameter estimation. We show that in practice the best possible out-of-sample investment performance is accomplished by using a relatively small subset of the total assets available. We develop a generic framework for determining the optimal subset and apply it to three popular investment strategies. The empirical analysis clearly demonstrates that our asset selection approach has substantial merit for investors. It also lends support to rational explanations of three empirical puzzles in the literature.

Our paper offers some reconciliation in the longstanding argument between academics and practitioners on the issue of optimal portfolio selection. Our results allow a validation and reinterpretation of the practical value of Markowitz's mean-variance theory. The diversification principle is alive and kicking but it should be applied with caution in order to avoid the perils of estimation risk. The competing view held by many investors is epitomised by the legendary Warren Buffet: "wide diversification is only required when investors do not understand what they are doing". Strictly speaking, our paper indicates the opposite since we find that unlimited diversification is optimal only when you know exactly what you are doing, i.e., when estimation risk is absent. However, in the presence of estimation risk, the advice of Buffet and the practice of investors are justified after all, since the optimal implementation of portfolio theory calls for investments in a small number of assets.

Appendix

A Stepwise Asset Selection

1. Forward Selection

   (a) Define $B_0^{forw} = \emptyset$ and $B_1^{forw} = \arg\max_{i \in A} M\left(p\left(\{i\}\right)\right)$

   (b) Compute $B_2^{temp} = \arg\max_{i \in A - B_1^{forw}} M\left(p\left(B_1^{forw} \cup \{i\}\right)\right)$

   (c) Compute $B_3^{temp} = \arg\max_{i \in A - B_2^{temp}} M\left(p\left(B_2^{temp} \cup \{i\}\right)\right)$

   (d) Compute $B_2^{forw} = \arg\max_{i \in B_3^{temp}} M\left(p\left(B_3^{temp} - \{i\}\right)\right)$

   (e) Proceed likewise until the derivation of B_N^{forw}.

2. Backward Elimination

   (a) Define $B_N^{back} = A$ and $B_0^{back} = \emptyset$

   (b) Compute $B_{N-1}^{back} = \arg\max_{i \in A} M\left(p\left(A - \{i\}\right)\right)$

   (c) Compute $B_{N-2}^{temp} = \arg\max_{i \in B_{N-1}^{back}} M\left(p\left(B_{N-1}^{back} - \{i\}\right)\right)$

   (d) Compute $B_{N-3}^{temp} = \arg\max_{i \in B_{N-2}^{temp}} M\left(p\left(B_{N-2}^{temp} - \{i\}\right)\right)$

   (e) Compute $B_{N-2}^{back} = \arg\max_{i \in A - B_{N-3}^{temp}} M\left(p\left(B_{N-3}^{temp} \cup \{i\}\right)\right)$

   (f) Proceed likewise until the derivation of B_1^{back}.

A *near optimal* set of assets is then

$$B_*^p = \arg\max_{\substack{\left\{B_i^{forw}, B_i^{back}\right\} \\ i \in \{0,1,\ldots,N\}}} \left\{M\left(p\left(B_i^{forw}\right)\right), M\left(p\left(B_i^{back}\right)\right)\right\}$$

B Proofs of Theoretical Results

Proof of Proposition 1: We first express the matrix $(\Sigma(A-B))^{-1}$ as a function of the elements of Σ^{-1}. We partition the covariance matrix and its inverse as:

$$\Sigma = \begin{pmatrix} \Sigma(B) & \Sigma_1' \\ \Sigma_1 & \Sigma(A-B) \end{pmatrix}, \Sigma^{-1} = \begin{pmatrix} \Omega_B & \Omega_1' \\ \Omega_1 & \Omega_{A-B} \end{pmatrix}, \tag{B.1}$$

where $\Sigma = (\sigma_{i,j})_{i,j=1,\dots N}$, $\Sigma(B) = (\sigma_{ij})_{i,j=1,\dots,|B|}$, $\Sigma(A-B) = (\sigma_{ij})_{i,j=|B|+1,\dots,N}$, $\Sigma_1 = (\sigma_{ij})_{\substack{i=|B|+1,\dots,N \\ j=1,\dots,|B|}}$, $\Omega_{A-B} = (\omega_{ij})_{i,j=|B|+1,\dots,N}$ and $\Omega_1 = (\omega_{ij})_{\substack{i=|B|+1,\dots,N \\ j=1,\dots,|B|}}$. Multiplying the covariance matrix to its inverse results in

$$\begin{aligned} &\Omega_1\Sigma_1' + \Omega_{A-B}\Sigma(A-B) = I_{|A-b|,|A-b|} \\ &\Omega_B\Sigma_1' + \Omega_1'\Sigma(A-B) = 0_{|A-b|} \end{aligned} \tag{B.2}$$

where $I_{|A-b|,|A-b|}$ is the $|A-B| \times |A-B|$ identity matrix and $0_{|A-b|}$ is vector of $|A-B|$ zeros.

Solving for $(\Sigma(A-B))^{-1}$ leads to

$$(\Sigma(A-B))^{-1} = \Omega_{A-B} - \Omega_1(\Omega_B)^{-1}\Omega_1' \tag{B.3}$$

The specification error losses due to investing in $A-B$ instead of A are given by

$$SE(A-B) = \mu'\Sigma^{-1}\mu - \mu(A-B)'(\Sigma(A-B))^{-1}\mu(A-B) \tag{B.4}$$

Applying (B.3) to the last equation results after some algebra to (18). ■

Proof of Proposition 2: We first note that the condition $T \leq \phi(N)$ is equivalent to $\kappa(N,T) \leq 0$. If $\kappa(N,T) \leq 0$, then for any proper subset B of A, we have that $\kappa(N,T) < \kappa(|B|,T)$, $\lambda(N,T) > \lambda(|B|,T)$ and $S_{\text{mv}}(B) \leq S_{\text{mv}}$. As a result $M(p^{\text{mv}}(A)) \leq M(p^{\text{mv}}(B))$. Otherwise, if $T > \phi(N)$, then $\kappa(N,T) > 0$. Reordering the terms in the inequality

$$\kappa(N,T)S_{\text{mv}} - \lambda(N,T) < \kappa(|B|,T)S_{\text{mv}}(B) - \lambda(|B|,T) \tag{B.5}$$

gives (20). ■

Derivation of (30): We provide an accurate estimate of the squared Sharpe ratio of the *ex ante* tangency portfolio $S_{\text{mv}} = \mu'\Sigma^{-1}\mu$. We first note that S_{mv} can be computed as

$$S_{\text{mv}} = \gamma^2 (w^{\text{mv}})'\Sigma w^{\text{mv}} \tag{B.6}$$

Naturally, the main source of error in the estimation of S_{mv} is the estimation of w^{mv}. To tackle this estimation risk we suggest the estimation of the *ex ante* optimal weights with a vector proportional to the global minimum variance portfolio, since this portfolio is much less sensitive to sampling errors. Therefore, we seek for the estimator of the form $c\hat{w}^{\min}$ that maximises the standard statistical loss function[12]

$$D\left(w^{\mathrm{mv}}, c\hat{w}^{\min}\right) = E\left(\left(w^{\mathrm{mv}} - c\hat{w}^{\min}\right)' \Sigma \left(w^{\mathrm{mv}} - c\hat{w}^{\min}\right)\right) \tag{B.7}$$

Expanding (B.7) results in

$$D\left(w^{\mathrm{mv}}, c\hat{w}^{\min}\right) = \left(w^{\mathrm{mv}}\right)' \Sigma w^{\mathrm{mv}} - 2c\mu_{\min} + c^2 v\left(N,T\right)\sigma^2_{\min} \tag{B.8}$$

since $E\left(\hat{w}^{\min}\right) = w^{\min}$ and $E\left(\left(\hat{w}^{\min}\right)' \Sigma \hat{w}^{\min}\right) = v\left(N,T\right)\sigma^2_{\min}$ (e.g., see Kourtis et al., 2009). Applying first-order conditions to (B.8), gives the optimal coefficient

$$c^* = \frac{\mu_{\min}}{\gamma v\left(N,T\right)\sigma^2_{\min}} \tag{B.9}$$

The unknown ratio $\dfrac{\mu_{\min}}{\sigma^2_{\min}} = 1'_N \Sigma^{-1}\mu$ in the above equation can be computed using $\hat{\mu}$ and the unbiased estimate of the inverse covariance $\dfrac{T-N-2}{T}\hat{\Sigma}^{-1}$. This computation leads to the following estimate of the optimal coefficient:

$$\hat{c}^* = \frac{(T-N-1)(T-N-2)}{\gamma T(T-2)} \frac{\hat{\mu}_{\min}}{\hat{\sigma}^2_{\min}} \tag{B.10}$$

and the corresponding estimate of the mean-variance weights $\hat{c}^*\hat{w}^{\min}$. We substitute the latter weights in (B.6) and estimate the covariance matrix using its unbiased estimator $\dfrac{T}{T-1}\hat{\Sigma}$ to obtain (30). ■

[12]Examples of studies using this loss function are Jorion (1986), Frost and Savarino (1986), Kan and Zhou (2007) and DeMiguel et al. (2009a) among others.

References

Adler, M., and B. Dumas, 1983, "International Portfolio Choice and Corporation Finance: A Synthesis," *Journal of Finance* 38, 925–984.

Ang, A., and G. Bekaert., 2002, "International Asset Allocation with Regime Shifts," *Review of Financial Studies* 15, 1137–1187.

Bajeux-Besnainou, I., J. Jordan, and R. Portait, 2003, "Dynamic asset allocation for stocks, bonds and cash," *Journal of Business*, 76, 263–287.

Barber, B. M., and T. Odean, 2000, "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55, 773-806.

Benartzi, S., and R. Thaler, 2001, "Naive diversification strategies in defined contribution saving plans," *American Economic Review*, 91, 79-98.

Best, M. J., and R. R. Grauer, 1991, "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies*, 4, 315–342.

Black, F., 1990, "Equilibrium Exchange Rate Hedging", *Journal of Finance* 45, 899-908.

Blume, M., and I. Friend, 1975, "The Asset Structure of Individual Portfolios and Some Implications for Utility Functions," *Journal of Finance*, 30, 585-603.

Boyle, P., L. Garlappi, R. Uppal, and T. Wang, 2009, "Keynes meets Markowitz: the tradeoff between familiarity and diversification," Working paper, London Business School.

Brandt, M., 2009, "Portfolio Choice Problems," *Handbook of Financial Econometrics*, Y. Ait-Sahalia and L P. Hansen, eds, pp 269-336.

Brennan, M., and H. Cao, 1997, "International Portfolio Flows", *Journal of Finance* 52, 1851-1880.

Brennan, M., and J. Xia, 2000, "Stochastic interest rates and the bond-stock mix", *European Finance Review* 4, 197-210.

Brennan, M., and J. Xia, 2002, "Dynamic asset allocation under inflation", *Journal of Finance* 57, 1201-1238.

Brennan, M., 1975, "The Optimal Number of Securities in a Risky Asset Portfolio when there are Fixed Costs of Transacting: Theory and some Empirical Results," *Journal of Financial and Quantitative Analysis*, 10, 483-496.

Britten-Jones, M., 1999, "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights," *Journal of Finance*, 54, 655-671.

Calvet, L. E., J. Y. Campbell, and P. Sodini, 2007, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115, 707-747.

Canner, N., G. Mankiw, and D. Weil, 1997, "An asset allocation puzzle," American Economic Review 87, 181–191.

Chakroun, O., G. Dionne, and A., Dugas-Sampara, "Empirical evaluation of the asset-allocation puzzle", *Economics Letters*, 100, 304-307.

Chopra, V. and W. Ziemba, 1993, "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice", *Journal of Portfolio Management* 19, 6-12.

Conine, Jr., T. E., and M. J. Tamarkin, 1981, "On Diversification Given Asymmetry in Returns," *Journal of Finance*, 36, 1143-1155.

Constantinides, G. M. and A. G. Malliaris, 1995, "Portfolio theory," *Handbooks in Operations Research and Management Science: Finance*, R. A. Jarrow, V. Maksimovic and W. T. Ziemba eds, pp. 1-30.

Dahlquist, M., L. Pinkowitz, R. Stulz, and R. Williamson, 2003, "Corporate Governance and the Home Bias," *Journal of Financial and Quantitative Analysis*, 38, 87-110.

DeMiguel, V., L. Garlappi and R. Uppal, 2009a, "Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" *Review of Financial Studies*, 22, 1915-1953.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal, 2009b, "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science*, 55, 798-812.

Duchin, R., and H. Levy, 2009, "Markowitz versus the Talmudic portfolio diversification strategies," *Journal of Portfolio Management*, 35, 71-74.

Efroymson, M. A., 1960, "Multiple regression analysis," In *Mathematical Methods for Digital Computers,* Wiley, Ralston, A. and H. S, Wilf, eds.

Fama, E., 1976, *Foundations of Finance*, Oxford: Blackwell.

French, K., and J. Poterba, 1991, "Investor Diversification and International Equity Markets", *American Economic Review*, 81, 222-226.

Frost, P. A., and J. E. ,Savarino, 1988, "For Better Performance: Constrain Portfolio Weights", *Journal of Portfolio Management* 14, 29-34.

Goetzman, W. N., and A. Kumar, 2008, "Equity Portfolio Diversification," *Review of Finance*, 12, 433-463.

Guidolin, M., 2005, “Home Bias and High Turnover in an Overlapping Generations Model with Learning,” *Review of International Economics* 13, 725-756.

Guidolin, M., and A. Timmermann, 2008, “International asset allocation under regime switching, skew and kurtosis preferences,” *Review of Financial Studies*, 21, 889–935.

Guiso, L., M. Haliassos and T., Jappelli, 2002, Household Portfolios, MIT Press, Cambridge, MA.

Harvey, C. R., and A. Siddique, 2000, “Conditional Skewness in Asset Pricing Tests,” *Journal of Finance*, 55, 1263-1295.

Huberman, G., and S., Kandel, 1987, “Mean-variance spanning,” *Journal of Finance* 42, 873-888.

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

Ivković, Z., and S. Weisbenner, 2005, “Local Does as Local Is: Information Content of the Geography of Individual Investors’ Common Stock Investments,” *Journal of Finance*, 60, 267-306.

Jacob, N. L., 1974, “A Limited-Diversification Portfolio Selection Model for the Small Investor,” *Journal of Finance*, 29, 847-856.

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

Jobson, J. D., and B. Korkie, 1980, “Estimation for Markowitz Efficient Portfolios,” *Journal of the American Statistical Association*, 75, 544-554.

Jorion, P., 1986, “Bayes-Stein Estimation for Portfolio Analysis,” *Journal of Financial and Quantitative Analysis*, 21, 279-292.

Kan, R., and G. Zhou, 2007, “Optimal portfolio choice with parameter uncertainty,” Journal of *Financial and Quantitative Analysis*, 42, 621-656.

Kelly, M., 1995, “All Their Eggs in One Basket: Portfolio Diversification of U.S. Households,” *Journal of Economic Behavior and Organization*, 27, 87-96.

Keynes, J. M., 1983, The Collected Writings of John Maynard Keynes, MacMillan, Cambridge University Press, New York, Moggridge, Donald Edward, ed.

Kourtis A., G. Dotsis, and R. Markellos, 2009, “Parameter Uncertainty in Portfolio Selection: Shrinking the Inverse Covariance Matrix”, Working Paper, AUEB.

Kraus, A., and R. H. Litzenberger, 1976, “Skewness Preference and the Valuation of Risk Assets,” *Journal of Finance*, 31, 1085-1100.

Kubokawa T, C. P., Robert, and A. K. Md. E Saleh, 1993, “Estimation of the noncentrality parameters,” *Canadian Journal of Statistic*s, 21, 45-57.

Lease, R. C., W. G., Lewellen, and G. G. Schlarbaum, 1974, “The Individual Investor: Attributes and Attitudes”, *Journal of Business* 29, 413-433.

Lewis, K., 1999, “Trying to Explain Home Bias in Equities and Consumption,” *Journal of Economic Literature*, 37, 571-608.

Lioui, Abraham, 2007, ”The asset allocation puzzle is still a puzzle,” *Journal of Economic Dynamics and Control*, 31, 1185-1216.

Litterman, B., 2003, *Modern investment management: an equilibrium approach.* New York: Wiley.

Liu, H., 2009, “Portfolio Insurance and Underdiversification,” Working Paper, Olin Business School, Washington University.

Mao, J. C. T., 1970, “Essentials of Portfolio Diversification Strategy,” *Journal of Finance*, 27, 127-129.

Markowitz, H. M., 1959, *Portfolio selection: efficient diversification of investments.* New Jersey: Wiley.

Markowitz, H. M., 1952, “Portfolio Selection,” *Journal of Finance*, 7, 77-91.

Merton, R. C., 1987, “A Simple Model of Capital Market Equilibrium with Incomplete Information,” *Journal of Finance*, 42, 35-55.

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

Mitton, T., and K. Vorkink, 2007, “Equilibrium Underdiversification and the Preference for Skewness,” *Review of Financial Studies*, 20, 1255-1288.

Polkovnichenko, V. 2005, “Household Portfolio Diversification: A Case for Rank-Dependent Preferences,” *Review of Financial Studies*, 18, 1467-1502.

Roche, H., S. Tompaidis, and C. Yang, 2009, “Asset Selection and Under-Diversification with Financial Constraints and Income: Implications for Household Portfolio Studies,” Working paper, University of Texas at Austin.

Serrat, A., 2001, “A Dynamic Equilibrium Model of International Portfolio Holdings,” *Econometrica* 69, 1467-89.

Shefrin, H., and M. Statman, 2000, “Behavioral Portfolio Theory,” *Journal of Financial and Quantitative Analysis*, 35, 127-151.

Tobin, J., 1958, "Liquidity Preference as Behavior towards Risk," *Review of Economic Studies*, 25, 65-85.

Thomas, C., F. Warnock, and J. Wongswan, 2006, "The Performance of International Portfolios," NBER Working Paper No. 12346.

Tu, J. and G. Zhou, 2004,"Data-generating Process Uncertainty: What Difference Does it Make in Portfolio Decisions?" *Journal of Financial Economics*, 72, 385-421.

Tu, J. and G. Zhou, 2009, "Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies," Working paper, Olin Business School, Washington University.

Uppal, R., and T. Wang, 2003, "Model misspecification and underdiversification," *Journal of Finance*, 58, 2465–2486.

Van Nieuwerburgh, S., and L. Veldkamp, 2010, "Information Acquisition and Portfolio Under Diversification," forthcoming in *Review of Economic Studies.*

Table I: List of datasets

This table contains the datasets of excess returns employed in this paper for the assessment of the optimal asset selection approach. It also reports the number of assets in each dataset, the time period spanned as well as their source. Excess returns are computed using the 30-day nominal US T-Bill obtained from Kenneth French's Website.

Description	Number of Assets	Time Period	Source
Country Stock Indices (France, Germany, Japan, UK and USA)	5	01/1970-12/2007	MSCI
Industry Portfolios	10	07/1963-08/2007	Kenneth French's Website
Size and Book-to-Market Portfolios	25	07/1963-08/2007	Kenneth French's Website
Random S&P 500 Stocks	50	01/1981-12/2007	Reuters

Table II: Effect of the elimination of one asset to the out-of-sample performance of the sample mean-variance strategy

This table compares the average expected utility for the mean-variance strategy ($M\left(p^{\text{mv}}\left(A\right)\right)$) when this is applied to all available N assets to the corresponding expected utilities $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$ in the subsets with N-1 assets. These utilities are computed using the formulas provided in subsection A of Section II. The mean and covariance matrix of excess returns are respectively calibrated using the 10 Industry Portfolios (Panel A) and the 25 Size and Book-to-Market Portfolios (Panel B) datasets and a coefficient of relative risk aversion equal to 3. The sample sizes considered are contained in column 1. The expected utility in the whole asset space A is reported in column 2 while the columns 3-5 report the minimum, the maximum and the average utility for the subset $A-\{i\}$ with N - 1 assets, respectively. Column 6 finally reports the number of assets whose elimination improves the average out-of-sample performance.

Panel A: 10 Industry Portfolios					
T	$M\left(p^{\text{mv}}\left(A\right)\right)$	Minimum $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Maximum $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Average $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Cases of Improved Performance
120	-17.32	-15.49	-14.39	-14.63	10
240	-3.79	-3.86	-2.69	-2.93	9
360	-0.11	-0.64	0.56	0.31	8
480	1.6	0.86	2.09	1.82	8

Panel B: 25 Size and Book-to-Market Portfolios					
T	$M\left(p^{\text{mv}}\left(A\right)\right)$	Minimum $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Maximum $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Average $M\left(p^{\text{mv}}\left(A-\{i\}\right)\right)$	Cases of Improved Performance
120	-69.16	-67.87	-61.09	-61.71	25
240	7.26	-2.44	9.26	8.2	21
360	22.74	11.13	23.8	22.86	19
480	29.21	16.86	29.92	28.74	17

Table III: Average expected utilities in the asset space and its optimal subset

For each one of the 4 datasets listed in Table I, this table reports the average (annualized and percentage) expected utility of a mean-variance investor for the sample mean-variance strategy (Sample mv), the sample global minimum variance strategy (Sample min) and the $1/N$ strategy when they are applied to all available assets (A) and to the set of assets which maximises out-of-sample performance ($B_*^{\text{mv}}, B_*^{\text{min}}$ and $B_*^{1/N}$, respectively). The optimal set for each strategy is computed by solving the respective optimal asset selection problem. The parameters that define this problem and the expected utilities are calibrated using the corresponding dataset. The utilities are analytically computed using the formulas provided in Section II for four sample sizes T (120, 240, 360 and 480) and two values for the coefficient of relative risk aversion (1 and 3). Finally, we report in parenthesis the number of optimal assets $|B_*^p|$ for each strategy as well as the expected utilities for the parameter certainty mean-variance (*Ex ante* mv) and global minimum variance (*Ex ante* min) strategies, respectively.

Panel A: 5 Country Stock Indices

	$\gamma=1$				$\gamma=3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	9.39	9.39	9.39	9.39	3.13	3.13	3.13	3.13
Sample mv (A)	-20.92	-4.48	0.41	2.75	-6.97	-1.49	0.14	0.92
Sample mv (B_*^{mv})	2.40	**5.25**	**6.15**	**6.60**	0.80	1.75	2.05	2.20
$\lvert B_*^{\text{mv}}\rvert$	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)
Ex ante min	4.89	4.89	4.89	4.89	2.88	2.88	2.88	2.88
Sample min (A)	4.86	4.88	4.88	4.88	2.78	2.83	2.85	2.86
Sample min (B_*^{min})	**5.10**	5.10	5.10	5.10	**2.99**	**3**	**3**	**3.01**
$\lvert B_*^{\text{min}}\rvert$	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
$1/N$ (A)	3.99	3.99	3.99	3.99	1.53	1.53	1.53	1.53
$1/N$ ($B_*^{1/N}$)	4.99	4.99	4.99	4.99	2.55	2.55	2.55	2.55
$\lvert B_*^{1/N}\rvert$	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)

Panel B: 10 Industry Portfolios

	$\gamma=1$				$\gamma=3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	18.74	18.74	18.74	18.74	6.25	6.25	6.25	6.25
Sample mv (A)	-51.96	-11.36	-0.32	4.81	-17.32	-3.79	-0.11	1.60
Sample mv (B_*^{mv})	**7.30**	**10.33**	**12.21**	**13.55**	2.43	3.44	4.07	4.52
$\lvert B_*^{\text{mv}}\rvert$	(1)	(2)	(2)	(3)	(1)	(2)	(2)	(3)
Ex ante min	5.03	5.03	5.03	5.03	3.66	3.66	3.66	3.66
Sample min (A)	4.98	5.01	5.02	5.02	3.49	3.58	3.61	3.62
Sample min (B_*^{min})	7.07	7.09	7.09	7.09	**5.12**	**5.16**	**5.17**	**5.18**
$\lvert B_*^{\text{min}}\rvert$	(4)	(4)	(4)	(4)	(4)	(4)	(4)	(4)
$1/N$ (A)	5.46	5.46	5.46	5.46	3.45	3.45	3.45	3.45
$1/N$ ($B_*^{1/N}$)	6.93	6.93	6.93	6.93	4.88	4.88	4.88	4.88
$\lvert B_*^{1/N}\rvert$	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)

Table III (continued)

Panel C: 25 Size and Book-to-Market Portfolios								
	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	133.87	133.87	133.87	133.87	44.62	44.62	44.62	44.62
Sample mv (A)	-207.47	21.78	68.22	87.64	-69.16	7.26	22.74	29.21
Sample mv (B_*^{mv})	**81.38**	**99.69**	**106.31**	**110.40**	**27.13**	**33.23**	**35.44**	**36.80**
$\lvert B_*^{\text{mv}} \rvert$	(4)	(5)	(6)	(8)	(4)	(5)	(6)	(8)
Ex ante min (A)	10.17	10.17	10.17	10.17	8.87	8.87	8.87	8.87
Sample min (A)	10	10.09	10.12	10.13	8.37	8.65	8.73	8.76
Sample min (B_*^{min})	17.28	17.29	17.29	17.29	14.27	14.29	14.30	14.30
$\lvert B_*^{\text{min}} \rvert$	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
$1/N$ (A)	7.26	7.26	7.26	7.26	4.32	4.32	4.32	4.32
$1/N$ ($B_*^{1/N}$)	12.14	12.14	12.14	12.14	8.02	8.02	8.02	8.02
$\lvert B_*^{1/N} \rvert$	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)

Panel D: 50 S&P 500 Stocks								
	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	85.27	85.27	85.27	85.27	28.42	28.42	28.42	28.42
Sample mv (A)	-1539	-218.5	-69.28	-16.80	-512.90	-72.82	-23.09	-5.60
Sample mv (B_*^{mv})	**21.55**	**34.60**	**40.99**	**46.78**	7.18	**11.53**	**13.66**	**15.59**
$\lvert B_*^{\text{mv}} \rvert$	(3)	(5)	(8)	(12)	(3)	(5)	(8)	(12)
Ex ante min (A)	6.50	6.50	6.50	6.50	5.33	5.33	5.33	5.33
Sample min (A)	6.08	6.35	6.41	6.43	4.08	4.88	5.05	5.13
Sample min (B_*^{min})	15.76	15.76	15.76	15.76	10.51	10.57	10.60	10.61
$\lvert B_*^{\text{min}} \rvert$	(1)	(1)	(1)	(1)	(4)	(4)	(4)	(4)
$1/N$ (A)	7.61	7.61	7.61	7.61	5.05	5.05	5.05	5.05
$1/N$ ($B_*^{1/N}$)	15.76	15.76	15.76	15.76	**10.93**	10.93	10.93	10.93
$\lvert B_*^{1/N} \rvert$	(1)	(1)	(1)	(1)	(3)	(3)	(3)	(3)

Table IV: Number of optimal assets as a function of the sample size

For each dataset listed in Table I and each number n reported in the first column below, this table reports the minimum number of observations required in order the optimal asset selection for the sample mean-variance strategy to result in an optimal set of assets $B_*^{mv} = \arg\max_{B \subseteq A} M\left(p^{mv}\left(B\right)\right)$ that contains n assets ($|B_*^{mv}| = n$). The squared Sharpe ratios of the *ex ante* mean-variance portfolios that define the asset selection problem are calibrated using the corresponding dataset in each case. Nan indicates that there is not an optimal set of assets of this size for any value of T.

Number of Assets	5 Country Stock Indices	10 Industry Portfolios	25 Size & Book-to-Market Portfolios	50 S&P 500 Stocks
1	86	57	Nan	36
2	547	230	28	89
3	2111	370	Nan	103
4	12345	1189	82	136
5	473983	2298	149	225
6	-	3527	291	271
7	-	6394	395	Nan
8	-	Nan	437	340
9	-	14429	0	Nan
N	473983	53368	75299	1288921

Table V: Equivalence of optimal asset selection to an increase in the sample size

For each one of the datasets listed in Table I and four values of the sample size T, this table reports the sample size T' needed for the sample-based mean-variance strategy to result in average expected utility in the whole asset space equal to the average expected utility in the optimal subset when the sample size is T. T' is defined as the minimum number that fulfils the condition $\kappa\left(N,T'\right)S_{\text{mv}} - \lambda\left(N,T'\right) \geq \kappa\left(\left|B_*^{\text{mv}}\right|,T\right)S_{\text{mv}}(B_*^{\text{mv}}) - \lambda\left(\left|B_*^{\text{mv}}\right|,T\right)$. The squared Sharpe ratios of the *ex ante* mean-variance portfolios that determine the optimal set of assets are calibrated using the corresponding dataset in each case.

	T			
Dataset	120	240	360	480
5 Country Indices	458	758	964	1114
10 Industry Portfolios	578	773	985	1230
25 Size & Book-to-Market Portfolios	432	623	753	871
50 S&P 500 Stocks	688	829	928	1046

Table VI: Sample size needed for the sample mean-variance strategy to beat the global minimum variance portfolio

For each dataset listed in Table I and each number of assets reported in the first column below, this table reports the minimum sample size T needed for the sample-based mean-variance strategy to result in average expected utility in the optimal set B_*^{mv} higher than the average expected utility for the sample global minimum variance portfolio that is applied to the whole asset space $\left(M\left(p^{\mathrm{mv}}\left(B_*^{\mathrm{mv}}\right)\right) > M\left(p^{\min}\left(A\right)\right)\right)$. The parameters that define the optimal set of assets and the expected utilities are calibrated using the corresponding dataset in each case. Two values for the coefficient of relative risk aversion are considered (1 and 3). Nan indicates that there are not any sets of assets of this size that result in sample tangency portfolios superior to the minimum variance portfolio.

	$\gamma = 1$				$\gamma = 3$			
Number of Assets	5 Country Indices	10 Industry Portfolios	25 Size & Book-to-Market Portfolios	50 S&P 500 Stocks	5 Country Indices	10 Industry Portfolios	25 Size & Book-to-Market Portfolios	50 S&P 500 Stocks
1	212	88	534	47	Nan	298	Nan	83
2	306	127	31	62	2206	259	36	86
3	426	163	41	72	2173	289	48	92
4	559	208	46	83	2735	356	52	101
5	697	253	54	96	3409	428	61	114
6	-	299	63	109	-	501	70	128
7	-	345	71	122	-	577	79	142
8	-	392	80	133	-	655	89	154
9	-	439	88	145	-	732	98	167
N	697	487	224	582	3409	811	247	635

Table VII: Sample size needed for the sample mean-variance strategy to beat the $1/N$ rule

For each dataset listed in Table I and each number of assets reported in the first column below, this table reports the minimum sample size T needed for the sample-based mean-variance strategy to result in average expected utility in the optimal set B_*^{mv} higher than the average expected utility for the $1/N$ strategy that is applied to the whole asset space $\left(M(p^{\mathrm{mv}}(B_*^{\mathrm{mv}})) > M\left(p^{1/N}\left(A\right)\right)\right)$. The parameters that define the optimal set of assets and the expected utilities are are calibrated using the corresponding dataset in each case. Two values for the coefficient of relative risk aversion are considered (1 and 3).

	$\gamma = 1$				$\gamma = 3$			
Number of Assets	5 Country Indices	10 Industry Portfolios	25 Size & Book-to-Market Portfolios	50 S&P 500 Stocks	5 Country Indices	10 Industry Portfolios	25 Size & Book-to-Market Portfolios	50 S&P 500 Stocks
1	166	93	46	50	195	254	65	87
2	254	132	30	64	287	241	32	88
3	358	169	40	75	401	274	42	93
4	470	215	45	85	527	339	47	102
5	586	261	53	99	657	408	55	116
6	-	308	62	111	-	478	64	129
7	-	356	70	124	-	550	73	144
8	-	405	79	136	-	625	81	156
9	-	453	87	148	-	699	90	169
N	586	502	221	590	657	774	228	638

Table VIII: Out-of-sample Sharpe ratios

For each one of the 4 datasets listed in Table I, this table reports the average (percentage) out-of-sample Sharpe ratios of the sample mean-variance strategy (Sample mv), the sample global minimum variance strategy (Sample min) and the $1/N$ strategy when they are applied to all available assets (A) and to the set of assets which maximises out-of-sample performance ($B_*^{\text{mv}}, B_*^{\text{min}}$ and $B_*^{1/N}$ respectively). The optimal set for each strategy is computed as in the previous tables. For each of the 4 datasets, the Sharpe ratios are computed in 10,000 simulated datasets of normal i.i.d returns with mean and covariance matrix calibrated using the corresponding original dataset. We derive these ratios for four sample sizes T (120, 240, 360 and 480) and a coefficient of relative risk aversion equal to 3. Finally, we report in parenthesis the number of optimal assets $|B_*^p|$ for each strategy as well as the Sharpe ratios for the parameter certainty mean-variance (*Ex ante* mv) and global minimum variance (*Ex ante* min) strategies.

Panel A

	5 Country Indices				10 Industry Portfolios			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	12.51	12.51	12.51	12.51	17.67	17.67	17.67	17.67
Sample mv (A)	6.46	8.12	9.09	9.77	8.85	11.34	12.79	13.64
Sample mv (B_*^{mv})	9.11	10.60	11.09	11.35	13	14.68	15.20	15.83
$\|B_*^{\text{mv}}\|$	(1)	(1)	(1)	(1)	(1)	(2)	(2)	(3)
Ex ante min (A)	12	12	12	12	14.09	14.09	14.09	14.09
Sample min (A)	11.80	11.91	11.94	11.95	13.57	13.81	13.91	13.96
Sample min (B_*^{min})	**12.23**	**12.25**	**12.26**	**12.27**	**16.64**	**16.75**	**16.79**	**16.80**
$\|B_*^{\text{min}}\|$	(2)	(2)	(2)	(2)	(4)	(4)	(4)	(4)
$1/N$ (A)	9.60	9.60	9.60	9.60	13.16	13.16	13.16	13.16
$1/N$ ($B_*^{1/N}$)	11.48	11.48	11.48	11.48	16.04	16.04	16.04	16.04
$\|B_*^{1/N}\|$	(1)	(1)	(1)	(1)	(2)	(2)	(2)	(2)

Panel B

	25 Size & Book-to-Market Portfolios				50 S&P 500 Stocks			
	T							
Strategy	120	240	360	480	120	240	360	480
Ex ante mv (A)	47.23	47.23	47.23	47.23	37.70	37.70	37.70	37.70
Sample mv (A)	30.44	37.01	40.04	41.52	14.67	21.41	24.85	27.19
Sample mv (B_*^{mv})	**39.57**	**41.88**	**42.82**	**43.50**	22.51	**25.74**	**27.68**	**29.58**
$\|B_*^{\text{mv}}\|$	(4)	(5)	(6)	(8)	(3)	(5)	(8)	(12)
Ex ante min (A)	27.37	27.37	27.37	27.37	18.91	18.91	18.91	18.91
Sample min (A)	24.39	25.88	26.50	26.68	14.42	16.75	17.59	17.96
Sample min (B_*^{min})	31.26	31.32	31.33	31.36	24.27	24.42	24.47	24.49
$\|B_*^{\text{min}}\|$	(2)	(2)	(2)	(2)	(4)	(4)	(4)	(4)
$1/N$ (A)	14.70	14.70	14.70	14.70	16.04	16.04	16.04	16.04
$1/N$ ($B_*^{1/N}$)	20.19	20.19	20.19	20.19	**24.76**	24.76	24.76	24.76
$\|B_*^{1/N}\|$	(1)	(1)	(1)	(1)	(3)	(3)	(3)	(3)

Table IX: Standard errors of expected utilities

This table reports the (percentage) standard errors of the expected utilities whose average is reported in Table III. For each one of the 4 datasets of Table I, the standard errors are computed in 10,000 simulated datasets of normal i.i.d returns with mean and covariance calibrated using the corresponding original dataset. We derive these errors for two sample sizes T (120 and 240) and a coefficient of relative risk aversion equal to 3. Finally, we report in parenthesis the number of optimal assets $|B_*^p|$ for each strategy.

	5 Country Indices		10 Industry Portfolios		25 Size & Book-to-Market Portfolios		50 S&P 500 Stocks	
	T							
Strategy	120	240	120	240	120	240	120	240
Sample mv (A)	7.25	3.11	12.71	5.12	51.98	13.76	250.34	31.19
Sample mv (B_*^{mv})	2.76	1.32	2.82	1.88	7.15	3.85	5.66	3.39
$\lvert B_*^{\text{mv}}\rvert$	(1)	(1)	(1)	(2)	(4)	(5)	(3)	(5)
Sample min (A)	0.18	0.12	0.42	0.29	1.56	1.05	1.55	0.86
Sample min (B_*^{min})	0.04	0.02	0.17	0.11	1.13	0.82	0.36	0.24
$\lvert B_*^{\text{min}}\rvert$	(2)	(2)	(4)	(4)	(2)	(2)	(4)	(4)

Table X: Certainty equivalent returns from real data

For each one of the 4 datasets listed in Table I, this table reports the certainty equivalent returns (CERs) on the sample mean-variance strategy (Sample mv), the sample global minimum variance strategy (Sample min) and the $1/N$ strategy when they are applied to all available assets (A) and to the corresponding optimal set (B_*^{mv}, B_*^{min} and $B_*^{1/N}$, respectively). The optimal set for each strategy is computed as in the previous tables. For each dataset, CERs are computed using a rolling window approach with window length 120 or 240 and two values for the coefficient of relative risk aversion (1 and 3). Finally, we report in parenthesis the number of optimal assets $|B_*^p|$ for each strategy, as well as the CERs on the *N-assets* mean-variance and minimum variance rules estimated using the whole sample (In sample mv and In sample min, respectively).

	5 Country Indices		10 Industry Portfolios		25 Size & Book-to-Market Portfolios		50 S&P 500 Stocks	
Panel A: $\gamma = 1$								
	T							
Strategy	120	240	120	240	120	240	120	240
In sample mv (A)	7.50	7.50	28.02	28.02	227.28	227.28	46.68	46.68
Sample mv (A)	-33.08	-12.92	-63.51	-20.86	-325.20	-37.63	-927.46	-208.25
Sample mv (B_*^{mv})	-2.35	**10.01**	**16.53**	**19.76**	**95.54**	**114.77**	**8.41**	**13.59**
$\lvert B_*^{\text{mv}}\rvert$	(1)	(1)	(1)	(2)	(4)	(5)	(3)	(5)
In sample min (A)	4.24	4.24	7.95	7.95	14.29	14.29	4.71	4.71
Sample min (A)	4.28	4.21	8.10	8.70	16.91	15.58	-1.84	1.39
Sample min (B_*^{min})	5.04	4.32	10.15	10.39	21.02	21.14	7.11	7.11
$\lvert B_*^{\text{min}}\rvert$	(2)	(2)	(4)	(4)	(2)	(2)	(1)	(1)
$1/N$ (A)	3.13	3.13	7.24	7.24	7.58	7.58	8.02	8.02
$1/N$ ($B_*^{1/N}$)	**6.96**	6.96	9.22	9.22	11.13	11.13	8.33	8.33
$\lvert B_*^{1/N}\rvert$	(1)	(1)	(2)	(2)	(1)	(1)	(1)	(1)
Panel B: $\gamma = 3$								
	T							
Strategy	120	240	120	240	120	240	120	240
In sample mv (A)	2.50	2.50	9.34	9.34	75.76	75.76	15.56	15.56
Sample mv (A)	-11.03	-4.31	-21.17	-6.95	-108.40	-12.54	-309.15	-69.42
Sample mv (B_*^{mv})	-0.78	3.34	5.51	6.59	**31.85**	**38.26**	2.80	4.53
$\lvert B_*^{\text{mv}}\rvert$	(1)	(1)	(1)	(2)	(4)	(5)	(3)	(5)
In sample min (A)	2.33	2.33	6.56	6.56	13.11	6.56	4.13	4.13
Sample min (A)	2.54	2.33	6.59	7.10	14.17	7.10	-3.29	0.37
Sample min (B_*^{min})	3.13	2.37	**8.23**	**8.48**	19.01	8.48	2.41	3.14
$\lvert B_*^{\text{min}}\rvert$	(2)	(2)	(4)	(4)	(2)	(2)	(4)	(4)
$1/N$ (A)	1	1	5.36	5.36	5.09	5.09	6.20	6.20
$1/N$ ($B_*^{1/N}$)	**4.98**	**4.98**	7.41	7.41	8.04	8.04	**6.49**	**6.49**
$\lvert B_*^{1/N}\rvert$	(1)	(1)	(2)	(2)	(1)	(1)	(3)	(3)

Table XI: The asset allocation puzzle

This table illustrates a potential explanation to the asset allocation puzzle discussed in subsection C of Section IV. A market of a bond (b), a stock (s) and a riskless asset is considered. The annual mean and standard deviation of the excess return on the bond are assumed to be 0.04 and 0.1, respectively, whereas the corresponding moments for the stock are 0.08 and 0.2. The covariance between these assets is zero. The table reports the average (annualized and percentage) expected utilities of a mean-variance investor for the sample mean-variance strategy (p^{mv}), the sample global minimum variance strategy ($p^{\min}$) and the $1/N$ strategy ($p^{1/N}$) in the set of assets which maximise out-of-sample performance ($B_*^{\mathrm{mv}}, B_*^{\min}$ and $B_*^{1/N}$, respectively). These optimal sets of assets and the *ex ante* portfolio weights ($w^p\left(B_*^p\right)$) are also reported. We derive these quantities for two sample sizes T (120 and 240) and two values for the coefficient of relative risk aversion (1 and 3).

	$\gamma=1$		$\gamma=3$	
	T			
	120	240	120	240
$\left(M\left(p^{\mathrm{mv}}\left(B_*^{\mathrm{mv}}\right)\right)\right)$	2.5	5.34	0.83	1.78
B_*^{mv}	$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$w^{\mathrm{mv}}\left(B_*^{\mathrm{mv}}\right)$	(4,0)	(4,0)	(1.3,0)	(1.3,0)
$M\left(p^{\min}\left(B_*^{\min}\right)\right)$	6	6	3.59	3.6
$B_*^{\min}$	$\{s\}$	$\{s\}$	$\{b,s\}$	$\{b,s\}$
$w^{\min}\left(B_*^{\min}\right)$	(0,1)	(0,1)	(0.8,0.2)	(0.8,0.2)
$M\left(p^{1/N}\left(B_*^{1/N}\right)\right)$	6	6	4.12	4.12
$B_*^{1/N}$	$\{s\}$	$\{s\}$	$\{b,s\}$	$\{b,s\}$
$w^{1/N}\left(B_*^{1/N}\right)$	(0,1)	(0,1)	(0.5,0.5)	(0.5,0.5)

Table XII: Expected utilities in the practical implementation of optimal asset selection

For each one of the 4 datasets listed in Table I, this table reports the average (annualized and percentage) expected utility of a mean-variance investor for the sample mean-variance strategy (Sample mv), the sample global minimum variance strategy (Sample min) and the $1/N$ strategy when they are applied to all available assets (A) and to the estimated set of assets which maximises out-of-sample performance ($\hat{B}_*^{\text{mv}}, \hat{B}_*^{\text{min}}$ and $\hat{B}_*^{1/N}$ respectively). The estimates of the optimal set for each strategy and the expected utilities are computed in 10,000 simulated datasets of normal i.i.d returns with mean and covariance matrix calibrated using the corresponding original dataset. The estimations are performed by following the suggestions of Section V. We derive utilities for four sample sizes T (120, 240, 360 and 480) and two values for the coefficient of relative risk aversion (1 and 3).

Panel A: 5 Country Stock Indices

	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Sample mv (A)	-20.85	-4.59	0.43	2.73	-6.93	-1.48	0.14	0.90
Sample mv ($\hat{B}_*^{\text{mv}}$)	-0.56	3.55	**5.19**	**5.96**	-0.22	1.24	1.75	1.99
Sample min (A)	**4.86**	**4.87**	4.88	4.88	**2.77**	**2.83**	**2.85**	**2.86**
Sample min ($\hat{B}_*^{\text{min}}$)	4.26	4.39	4.54	4.83	2.11	2.32	2.60	2.70
$1/N$ (A)	3.99	3.99	3.99	3.99	1.53	1.53	1.53	1.53
$1/N$ ($\hat{B}_*^{1/N}$)	4.16	4.38	4.54	4.61	1.80	2.02	2.15	2.24

Panel B: 10 Industry Portfolios

	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Sample mv (A)	-52.15	-10.94	-0.37	4.79	-17.50	-3.79	-0.07	1.57
Sample mv ($\hat{B}_*^{\text{mv}}$)	-5.46	5.67	**8.96**	**10.72**	-1.63	1.86	3.01	3.55
Sample min (A)	4.98	5.00	5.02	5.02	3.49	3.58	3.61	3.63
Sample min ($\hat{B}_*^{\text{min}}$)	**5.82**	**6.11**	6.27	6.42	**3.59**	**3.95**	**4.22**	**4.37**
$1/N$ (A)	5.46	5.46	5.46	5.46	3.45	3.45	3.45	3.45
$1/N$ ($\hat{B}_*^{1/N}$)	5.60	5.87	6.04	6.19	3.47	3.69	3.89	4.00

Table XII (continued)

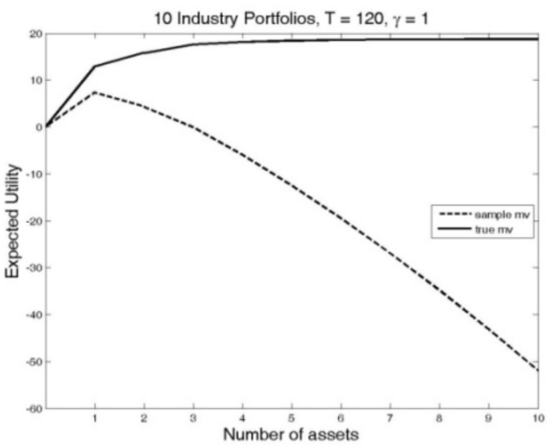
Panel C: 25 Size and Book-to-Market Portfolios								
	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Sample mv (A)	-204.60	20.29	68.41	87.62	-70.95	6.79	22.61	29.35
Sample mv ($\hat{B}_*^{\text{mv}}$)	**41.86**	**70.56**	**78.78**	**89.87**	**13.65**	**23.37**	**26.15**	**30.42**
Sample min (A)	9.99	10.06	10.11	10.15	8.42	8.68	8.71	8.78
Sample min ($\hat{B}_*^{\text{min}}$)	15.32	16.04	16.72	16.83	12.68	13.11	13.65	13.85
$1/N$ (A)	7.26	7.26	7.26	7.26	4.32	4.32	4.32	4.32
$1/N$ ($\hat{B}_*^{1/N}$)	9.28	9.85	10.28	10.59	5.53	6.09	6.52	6.81

Panel D: 50 S&P 500 Stocks								
	$\gamma = 1$				$\gamma = 3$			
	T							
Strategy	120	240	360	480	120	240	360	480
Sample mv (A)	-1527.97	-215.39	-69.31	-17.42	-505.70	-75.13	-22.80	-6.12
Sample mv ($\hat{B}_*^{\text{mv}}$)	-8.00	6.75	**15.24**	**18.98**	-2.37	2.30	5.00	6.34
Sample min (A)	6.13	6.35	6.43	6.43	4.10	4.93	5.06	5.14
Sample min ($\hat{B}_*^{\text{min}}$)	9.32	10.51	11.45	11.79	5.07	6.40	7.14	7.46
$1/N$ (A)	7.61	7.61	7.61	7.61	5.05	5.05	5.05	5.05
$1/N$ ($\hat{B}_*^{1/N}$)	**9.66**	**10.84**	11.92	12.46	**5.17**	**6.41**	**7.30**	**7.84**

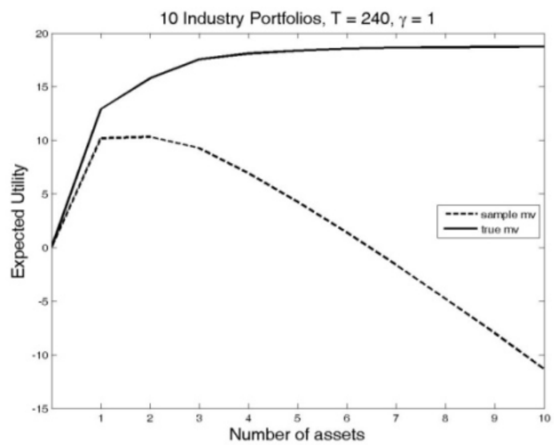
Figure 1: Expected out-of-sample performance of the sample mean-variance portfolio as a function of the number of assets ($N = 10$)

This figure shows the expected utility of a mean variance investor for the *ex ante* mean-variance strategy (solid line) and the average expected utility for the sample-based mean-variance strategy (dashed line) as a function of the number of optimally chosen assets. For each number of assets *n,* these strategies are applied to the set of assets that maximises the squared Sharpe ratio of the *ex ante* mean-variance portfolio ($B_n^{\mathrm{mv}} = \arg\max_{B \in C_n} S_{\mathrm{mv}}(B)$). The parameters that define this set and the corresponding utilities are calibrated using the excess returns on the 10 industry portfolios (Table I). Each panel in this figure corresponds to a different sample size (120, 240, 360 and 480, respectively). The value for the coefficient of relative risk aversion is 1.

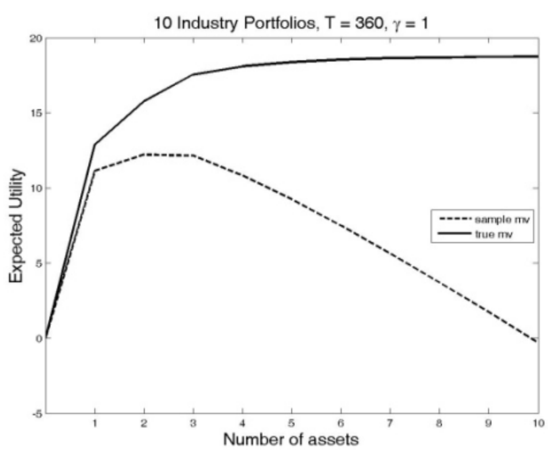
Panel A



Panel B



Panel C



Panel D

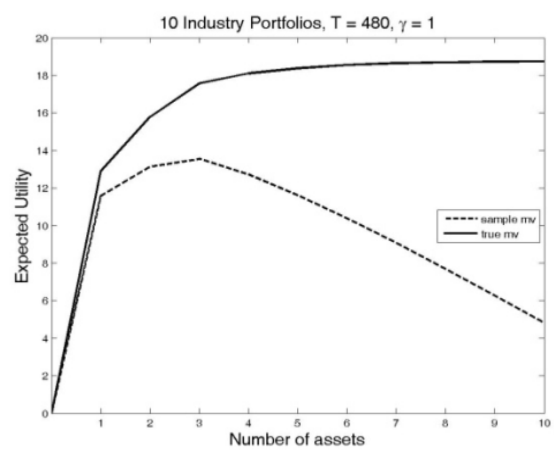


Figure 2: Expected out-of-sample performance of the sample mean-variance portfolio as a function of the number of assets ($N = 25$)

This figure shows the expected utility of a mean variance investor for the *ex ante* mean-variance strategy (solid line) and the average expected utility for the sample-based mean-variance strategy (dashed line) as a function of the number of optimally chosen assets. For each number of assets *n,* these strategies are applied to the set of assets that maximises the squared Sharpe ratio of the *ex ante* mean-variance portfolio ($B_n^{\mathrm{mv}} = \arg\max_{B \in C_n} S_{\mathrm{mv}}(B)$). The parameters that define this set and the corresponding utilities are calibrated using the excess returns on the 25 size and book-to-market portfolios (Table I). Each panel in this figure corresponds to a different sample size (120, 240, 360 and 480, respectively). The value for the coefficient of relative risk aversion is 1.

Panel A

Panel B

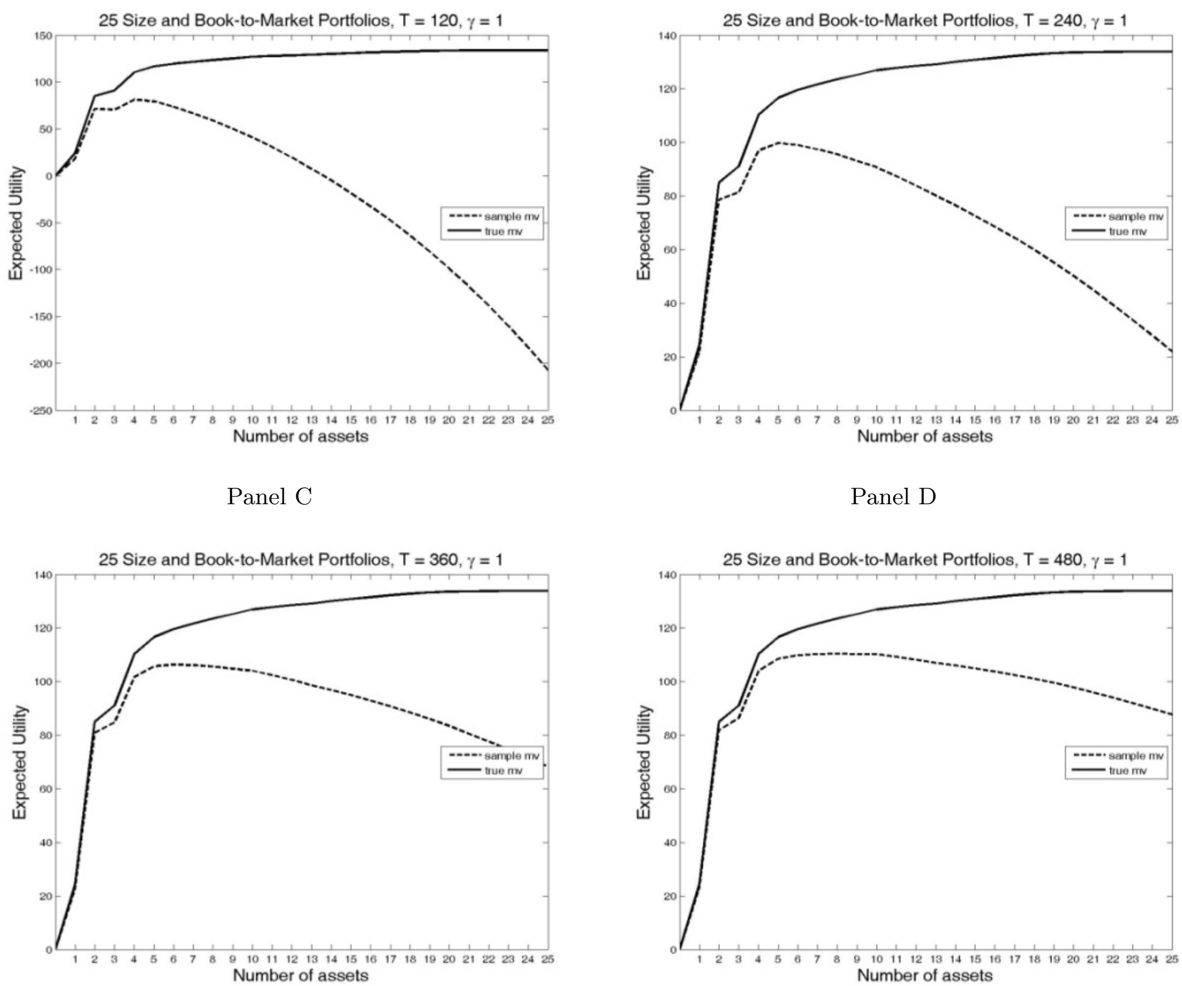
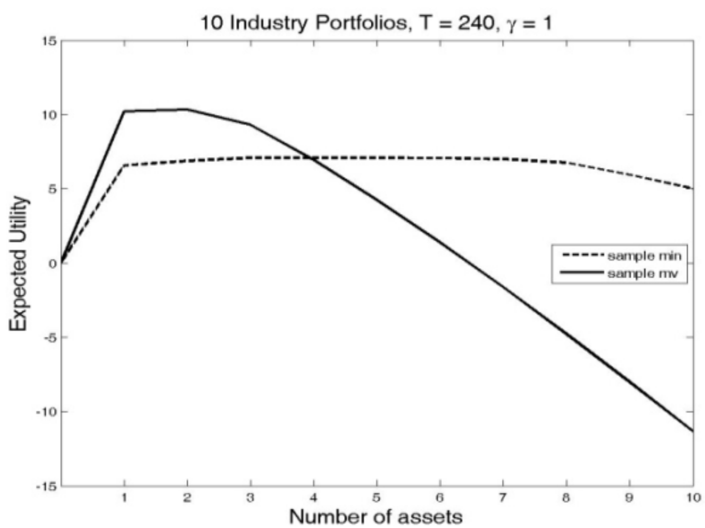


Figure 3: Comparison between the performance of the mean-variance strategy and its alternatives

This figure shows the average expected utility of a mean variance investor for the sample mean-variance strategy (solid line) and the average expected utilities for the sample global minimum variance portfolio (dashed line in Panel A) and the $1/N$ rule (dashed line in Panel B) as a function of the number of optimally chosen assets. For each number of assets n, each strategy is applied to the set of assets that maximizes its out-of-sample performance ($B_n^p = \arg\max_{B \in C_n} M\left(p\left(B\right)\right)$). The parameters that define this set and the corresponding utilities are calibrated using the excess returns on the 10 industry portfolios (Table I). The sample size is 240 and the coefficient of relative risk aversion is equal to 1.

Panel A (sample mean-variance rule & global minimum variance rule)



Panel B (sample mean-variance rule & $1/N$ rule)

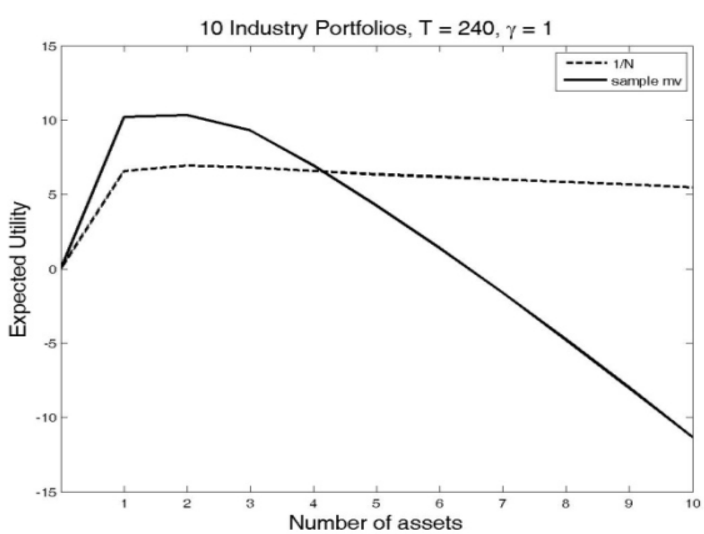


Figure 4: Out-of-sample Sharpe ratio as a function of the number assets

This figure shows the Sharpe ratio of the *ex ante* mean-variance strategy (solid line) and the average out-of-sample Sharpe ratio for the sample-based mean-variance strategy (dashed line) as a function of the number of optimally chosen assets. For each number of assets *n,* the strategies are applied to the set of assets that maximises the squared Sharpe ratio of the *ex ante* mean-variance portfolio ($B_n^{\mathrm{mv}} = \arg\max_{B \in C_n} S_{\mathrm{mv}}(B)$). The parameters that define this set and the corresponding utilities are calibrated using the excess returns on the 10 industry portfolios (Table I). The computation of the out-of-sample Sharpe ratios are from 10,000 simulated datasets of normal i.i.d returns with mean and covariance matrix calibrated from the actual data. The sample size is 240 and the coefficient of relative risk aversion is equal to 1.