Markowitz optimization: correcting past out-of-sample errors\textsuperscript{1}

Pedro Barroso\textsuperscript{2}

This version: October 2015

\textsuperscript{1}The present work has benefited from comments and suggestions of several of my colleagues at UNSW such as Li Yang, Chang Mo, Konark Saxena and Paul Krehnke. This is a very preliminary work and comments are welcome.

\textsuperscript{2}University of New South Wales. E-mail: p.barroso@unsw.edu.au.
Abstract
Portfolio optimization inputs differ widely from their subsequent out-of-sample (OOS) values. As a result, optimized portfolios have 2 to 28 times more risk OOS than their ex-ante estimates suggest. I propose a simple solution to this problem: let the data speak for itself and pick, in real time, the correction that most reduces past OOS errors. The resulting optimized portfolios consistently outperform OOS the 1/N benchmark. More significantly, the corrected covariance matrix captures very well the OOS risk of those portfolios, including the hit rates at extreme quantiles.

JEL classification: G11; G12; G17.

Keywords: Portfolio optimization; estimation error; covariance matrix.
1. Introduction

A method that consists on a plain implementation of the Markowitz optimization, with no constraints on portfolio weights or information on stock characteristics, consistently outperforms OOS the 1/N rule in terms of Sharpe ratio. More significantly, it succeeds in estimating the OOS risk of optimal stock portfolios, an achievement other methods I compare with fail to accomplish. This strategy relies on a simple correction of the optimization inputs using their mean OOS errors in the past.

It is well known that the inputs of Markowitz (1952) optimization are usually estimated with large errors (e.g. Jobson and Korkie (1980), Michaud (1989)). Hence the out-of-sample (OOS) performance of optimized portfolios often lags that of a naive 1/N benchmark (DeMiguel et al. (2009b)). I study portfolios of individual stocks and find the OOS errors of optimization inputs are indeed large, but they are also quite consistent and predictable to some extent.

Mean returns, variances, pairwise correlations and covariances inferred from historical samples, all regress substantially to the mean OOS. This suggests a rather simple approach to estimation error: correct an input’s estimate by how far its past estimates were from subsequent OOS values.

Correcting for these past OOS errors in the inputs, provides a simple and effective method to estimate the covariance matrix of a set of stocks and also filters out successfully most of the noise in mean returns.

An accurate estimate of the covariance matrix is of paramount importance for risk management. As pointed out by Ledoit and Wolf (2014), there is not a lack of methods available to estimate successfully the mean returns of a set of stocks. Green et al. (2013) show there are about 300 factors or characteristics that
seem to forecast expected returns in the cross section (for a recent comprehensive example of a study of the cross section of stock returns see Lewellen (2014)). In spite of the challenges faced by portfolio optimization, there are available portfolio optimization methods that convincingly outperform OOS naive benchmarks (e.g. DeMiguel et al. (2009a), Brandt et al. (2009)). But to achieve that, these methods circumvent the issue of improving the covariance matrix estimation and that is important by itself for other purposes, such as risk management.\footnote{Brandt et al. (2009) use asset characteristics to model weights directly, avoiding the issue of estimating the covariance matrix altogether. DeMiguel et al. (2009a) focus on imposing constraints on the final output of the optimization process, the vector of portfolio weights, and show this improves performance OOS.}

It is an empirical fact that institutional investors often have relatively concentrated stock portfolios. Agarwal et al. (2013) show that the Herfindhal index of a typical mutual fund stock portfolio is 0.018 and that of a hedge fund is 0.047. This implies that the equivalent number of holdings, defined as the reciprocal of the Herfindhal index, is respectively 56 and 21 stocks. This concentration of the bulk of a portfolio in a relatively small set of securities can seem inefficient from a diversification perspective, but Kacperczyk et al. (2005) show it is associated with superior performance once controlling for risk. This shows relatively small stock portfolios are important for institutional investors. To manage the risk of those portfolios, whatever the information set used to estimate returns, institutional investors need a reasonable estimate of the covariance matrix. This should be particularly relevant to estimate the value-at-risk (VaR) for extreme quantiles of the distribution, particularly in the case of hedge funds pursuing long-short strategies.

As Basak (2005) show, historical estimates of risk grossly understate true OOS
risk. In my OOS exercise the mean variance portfolio’s standard deviation OOS is on average more than 28 times higher than its ex ante estimate would suggest. Losses exceeding the 1% VaR for that portfolio occur approximately 20 times more than they should assuming the ex ante distribution.

I find the problem is substantially reduced if instead of the historical covariance matrix one uses the constant correlation approach of Elton and Gruber (1973). In fact, the OOS performance of this approach for the global minimum portfolio is quite robust. Still losses exceeding the 1% VaR occur 5 times more than they should for that portfolio and 8 times more in the mean variance portfolio. So both the historical sample and the constant correlation matrix unquestionably underestimate OOS risk, especially in extreme quantiles.

In sharp contrast, the covariance matrix corrected for past errors the I propose produces risk estimates that are very close to actual OOS risk. For the global minimum portfolio, for example, losses exceeding the 1% VaR only happen in 1.4% of the OOS observations. This illustrates the potential of correcting past OOS errors for managing the risk of small stock portfolios.

Correcting the inputs for past OOS errors can be seen as a form of shrinkage. In that sense, my paper is related to a vast literature on portfolio methods that rely on Bayesian approaches to estimation error (Barry (1974), Bawa et al. (1979), Jobson and Korkie (1980), Jobson and Korkie (1981), Jobson et al. (1979), Jorion (1985), Jorion (1986)). It is particularly close to literature on shrinking the estimation error of the covariance matrix (Best and Grauer (1992), Ledoit and Wolf (2004a), Ledoit and Wolf (2004b), Disatnik and Benninga (2007)). In that literature it is shown that one effective way to reduce the estimation error in the covariance matrix is to shrink the estimates to some target. The exact target and the extent of shrinkage
to apply is a matter of ongoing research (Benninga (2014), Ledoit and Wolf (2014)). In comparison, there is no arbitrarily chosen prior in this paper, I just let the data speak for itself and chose the adjustment in inputs that would reduce past OOS errors the most.

It is out of the scope of the present work to compare the performance of the portfolios using corrected inputs with all possible alternatives. DeMiguel et al. (2009b) show that the 1/N is a remarkable and simple benchmark, so I focus on that for most comparisons. Compared to the existing literature the contribution of this paper relies mostly on: i) it proposes a new solution to the problem of how to pick the shrinkage coefficient for the covariance matrix; ii) the method proposed is a simple and intuitive adjustment to the inputs of Markowitz optimization, both the covariance matrix and the vector of mean returns; iii) unlike related work that focuses more on the OOS Sharpe ratio of optimized portfolios, my work gives more emphasis to the equally deserving topic of accurately estimating the risk of portfolios of individual stocks, especially at extreme quantiles.

The paper is organized as follows. Section 2 presents the evidence on regression to the mean in the inputs of Markowitz optimization. Section 3 proposes the method to correct for past OOS errors in the inputs and explains the construction of the alternative portfolio strategies. Section 4 shows the OOS performance of the strategies compared for one sequence of (approx.) 43 years of randomly drawn stocks out of the 500 largest firms. Section 5 shows the results of the simulations of 1000 similarly defined sequences of stock universes. Section 6 concludes.
2. Regression to the mean in optimization inputs

The Markowitz (1952) approach shows how to solve for the optimal weights of a portfolio given the information on the expected returns of the assets available and the respective covariance matrix. The vector of relative weights of the optimal risky portfolio is:

\[ w_t = \frac{\Sigma_t^{-1} \mu}{1_N \Sigma_t^{-1} \mu} \]  

(1)

where \( \mu \) is a N-by-1 vector of mean returns, \( 1_N \) is a N-by-1 vectors of ones, \( N \) is the number assets, and \( \Sigma \) is the covariance matrix. The inputs to solve the optimization are unknown in practice and have to be estimated. As DeMiguel et al. (2009b) explain, the classic “plug-in” approach solves this problem replacing the true mean and variances by their sample counterparts in some rolling window. In one out-of-sample testing framework, where the weights must be determined using only information available to each point in time, this amounts to estimating the inputs \( \hat{\mu} \) and \( \hat{\Sigma} \) in the historical sample. Implicitly, the approach relies on the strong assumption that historical sample moments offer the best estimate of their true unobservable counterparts. Throughout this paper, I call this the historical (or plug-in) method and denote the respective estimates as \( \mu_H \) and \( \Sigma_H \).

Goyal and Welch (2008) show that OOS the historical mean performs quite well predicting the equity premium when compared to most alternative methods. So it is not the general case that using the historical sample moments results in poor OOS estimates. But in the case of portfolio optimization, it is well known that historical estimates are plagued with large sampling errors (e.g. Michaud (1989), Kan and Smith (2008)) and hence result in poor OOS performance (DeMiguel et al.
This motivates a comparison of the inputs of Markowitz optimization in historical samples with their ex post, out-of-sample, counterparts.

Figure 1 shows the relation between the historical sample and the ex post periods for the entire universe of US stocks. For each period and variable, the observations are sorted into deciles according to their values over the previous 60 months. The y-axis shows the average value for each decile in the subsequent 12 months.

It is apparent that past covariances and correlations are positively related to their future counterparts, but the slope is clearly below one. This shows that assuming the past value is the correct estimate, as in the historical approach, is on average excessive. But it also shows it is sub-optimal to assume past correlations are best forecast by their cross section mean, as is the case in the constant correlation matrix of Elton and Gruber (1973). The slope of the OOS values is clearly not zero either.

Panel C shows the variances are relatively well approximated by their past values. Panel D shows there is actually mean reversion in mean returns (a result known at least since Bondt and Thaler (1985). So $\mu_H$ seems to be negatively correlated with true expected returns. Given this result, it is not surprising that the global minimum variance portfolio tends to outperform the mean variance portfolio OOS. Besides pervasive measurement error issues, the mean variance portfolio (estimated using the plug-in approach) in effect tends to overweight stocks with low expected returns and short those with high expected returns.

Table 1 presents the results of Fama and MacBeth (1973) predictive regressions for covariances, correlations, variances, and mean returns on their historical estimates. For each month, I run a regression of ex post OOS values on their respective
historical estimates. This regression draws power from a very large cross sample. For instance, a set of N assets has N(N-1) covariances. This implies, as the last row of the table shows, that the cross section of covariances is quite large, its maximum number of observations exceeds six million. Usually, the large number of covariances to estimate is pointed out as a limitation of optimization methods. But in this regression exercise, it is quite the opposite. The large cross section should lead to more accurate estimation.

[Insert table 1 near here]

The reported slopes and t-statistics are inferred from the time series averages of the Fama-MacBeth regression coefficients. For covariances, correlations, and variances the null hypothesis that the slope is zero (no predictability in the variable) is clearly rejected with t-statistics of 10.03 and 21.44 respectively. The hypothesis of a slope of zero is rejected at the 5% level in every cross section regression for the covariance and correlation, and in 98.18% of the regressions for the variance. This strongly suggests optimization could do better than just using the constant correlation matrix of Elton and Gruber (1973).\footnote{For correlations, the R-square of the regression is quite low (only 1.59%). This shows that making inference with a relatively small number of assets, as in Elton and Gruber (1973), the constant correlation matrix should provide a very good approximation. In fact, the high statistical significance of the slope coefficient in my results benefits from the very large cross section that is only tractable using modern computing power.}

On the other hand, the null hypothesis that slope coefficients of these variables are one (implying the historical approach is correct on average) is clearly rejected too. The t-statistics for covariances, correlations, and variances are, respectively, -17.12, -71.16, and -8.01. This is illustrative of the problems of the plug in approach that implicitly assumes a value of one for the slope. In the case of mean returns,
slope coefficient is also significantly negative with a t-statistic of -4.83.

All in all, this section shows there is a clear regression to the mean in the covariance matrix. The best estimate for the future correlation between a pair of stocks is somewhere between the past correlation of the same pair of stocks and the mean correlation between all pairs of stocks. In the case of mean returns, there is even mean reversion. Those stocks the historical approach estimates to have high-expected returns assets, are in fact, on average, those with lower expected returns.

In unreported results, I find this regression to the mean phenomenon to be pervasive among other possible optimization inputs. Besides variances, correlations, mean returns, I also examine CAPM betas, the betas with respect to the 3 Fama and French (1993) factors, the alpha of those regressions, the variance of the residuals of those regressions, and the pairwise correlations of the respective residuals between individual stocks. Regression to the mean is the norm in all of those results.\(^3\)

3. **Correcting past out of sample errors**

The previous section shows there are large differences between historical estimates of the optimization inputs and the values they assume on average OOS. But those differences are also consistent and predictable to some extent. For instance, above average historical correlations tend to become smaller. This leads to the possibility that correcting past OOS errors in the inputs can lead to more robust portfolio optimization.

\[^3\]In the case of CAPM beta estimation, the pattern of regression to the mean is reminiscent of Vasicek (1973)
For any individual variable of interest $X$ (variance, pairwise correlation, or mean return) let $X_{H,t}$ denote its historical estimate at time $t$ computed from a rolling window of $H$ observations. The value it assumes in a subsequent ex post window of $E$ months is denoted by $X_{E,t}$. The results shown throughout this paper are for $H = 60$ and $E = 12$ (note that $X_{E,t}$ only becomes known at $t + E$).

For each period $t$ in the sample, I run the cross-section regression:

$$X_{E,t,j} = g_{0,t} + g_{1,t}X_{H,t,j} + \epsilon_{t,j}$$  \hspace{1cm} (2)

for $j = 1, ..., N_t$, where $N_t$ is the number of stocks available in the sample at time $t$. If the historical approach is correct, the best estimate of $X_{E,t}$ is $X_{H,t}$ and so $g_{0,t} = 0$ and $g_{1,t} = 1$. If regression to the mean is total, the best estimate of $X_{E,t}$ is the cross section average, so $E(X_{E,t,j}) = g_{0,t}$ and $g_{1,t} = 0$.

The history of cross sectional estimates $\hat{g}_{0,t}$ and $\hat{g}_{1,t}$ can be used to form a corrected expectation of $X_{E,t,j}$, which I denote as $X_{G,t,j}$:

$$X_{G,t,j} = \bar{G}_{0,t-E} - E + \bar{G}_{1,t-E}X_{H,t,j}$$  \hspace{1cm} (3)

where $\bar{G}_{0,t-E} = \frac{\sum_{s=1}^{t-E} \hat{g}_{0,s}}{(t - E)}$ and $\bar{G}_{1,t-E} = \frac{\sum_{s=1}^{t-E} \hat{g}_{1,s}}{(t - E)}$. This estimate, $X_{G,t,j}$, is simply the historical estimate, $X_{H,t,j}$, corrected by how close (or how far) all known past historical estimates were of their subsequent OOS values ($X_{E,t,j}$). It is essentially a correction of past OOS errors. As $X_{G,t,j}$ only uses information available until time $t$, it can be used for OOS tests. The correction should become more accurate as more past OOS errors are available. As such, besides the usual

\footnote{The $G$ stands for the surname of Sir Francis Galton (1822-1911) who first proposed the concept of regression to the mean.}
initial in-sample period needed in OOS tests, the results in this paper also require one additional learning period ($L$) to correct for past OOS errors. Given this, the first truly OOS return for a strategy using $X_{G,t,j}$ will occur at time $H + L + E + 1$. The results in this paper use an initial period ($L$) of 120 months. For the chosen values of $H$, $L$, and $E$ that amounts to 193 months.\footnote{Please note that after the initial learning period, for a given set of assets, the only requirement is to have $H$ past observations. The method ‘learns’ from past OOS errors of similar assets.}

Using the correction above, I compute corrected inputs for the Markowitz optimization. The corrected correlation matrix, $\rho_{G,t}$, has in each entry the corrected pairwise correlation and ones in the diagonal. Similarly, I obtain the $N$-by-$1$ vector of corrected estimates of the variances, $\sigma_{G,t}$, and mean returns, $\mu_{G,t}$. The $\sigma_{G,t}$ is restricted to be strictly positive to machine precision. The Galton corrected covariance matrix is:

$$\Sigma_{G,t} = diag(\sigma_{G,t})^\frac{1}{2} \rho_{G,t} diag(\sigma_{G,t})^\frac{1}{2}$$

From this the weights of the Galton mean-variance (MV) portfolio are:

$$w^{MV}_{G,t} = \frac{\Sigma_{G,t}^{-1}\mu_{G,t}}{1_N\Sigma_{G,t}^{-1}\mu_{G,t}}$$

Similarly, the Galton global minimum variance (GMV) portfolio is:

$$w^{GMV}_{G,t} = \frac{\Sigma_{G,t}^{-1}1_N}{1_N\Sigma_{G,t}^{-1}1_N}$$

The weights of these portfolios do not use any information about the characteristics of the stocks and are not adjusted ex post to respect any constraints.
are simply the result of the Markowitz optimization applied to the corrected inputs. I compare the results of these two portfolios with the MV and GMV obtainable using the historical approach and the also the Elton and Gruber (1973) constant-correlation approach.

In the Elton and Gruber (1973) constant correlation approach, $\rho_{EG,t}$ consists of a matrix where the non-diagonal elements are the average of pairwise correlations in the rolling historical sample and the diagonal elements are all ones. Then the covariance matrix is:

$$\Sigma_{EG,t} = \text{diag}(\sigma_{H,t}) \rho_{EG,t} \text{diag}(\sigma_{H,t})$$

and $\sigma_{H,t}$ is the N-by-1 vector of estimated variances from the historical sample. The GMV portfolio of the Elton-Gruber approach is determined as in equation 6 but with $\Sigma_{EG,t}$ instead of $\Sigma_{G,t}$. Assuming that $\Sigma_{EG,t}$ is a reasonable estimate of the covariance matrix, I combine it too with $\mu_{H,t}$ to obtain the weights of a Elton-Gruber MV portfolio.

Besides these four portfolios, I also compare the two Galton-corrected portfolios with the 1/N benchmark that DeMiguel et al. (2009b) show compares favourably with most optimization methods.

4. The OOS performance

The dataset consists of monthly returns of the entire universe of US listed stocks on the Center for Research in Security Prices (CRSP). The monthly returns data start in 1950:03 and end in 2010:12. I follow Ledoit and Wolf (2014) and, at the
start of the sample, select randomly a universe of 50 stocks picked out of the 500 largest firms by market capitalization with a complete history of returns in the previous 60 months and the subsequent 12 months. The sample is kept fixed for 12 months and then a new sample is drawn each 12 months until the end of the time series. For the OOS exercise, one initial period of 193 months is needed, this implies that the first OOS return is in 1966:04. Also, the requirement of a subsequent history of 12 months means that no OOS returns can be computed for the last 11 months, so the OOS period is of 526 monthly returns from 1966:04 to 2010:01. So in total 43 universes of 50 stocks are sequentially drawn in the OOS period, one for each year \((526/12 = 43.83)\).

Table 2 shows the OOS performance of the 6 portfolios. The plain historical approach performs extremely badly. The minimum variance portfolio, not only achieves a negative Sharpe ratio as it also has extremely high standard deviation and crash risk (an excess kurtosis of 335.39 combined with a left skew of -17.70). This is illustrative of the well known problems of Markowitz optimization. By comparison with the historical approach, the simple 1/N portfolio, with a Sharpe ratio of 0.31, performs quite well.\(^7\) This confirms with portfolios of individual stocks the result DeMiguel et al. (2009b) obtain with industry and size / book-to-market sorted portfolios.

The Elton-Gruber GMV portfolio has an interesting performance, with a Sharpe ratio OOS almost identical to the one of the 1/N. But the Elton-Gruber MV has a very poor performance, with a Sharpe ratio of only 0.09 and, more importantly,

---

\(^6\)The selections are made starting from the default random number generator of Matlab2014.

\(^7\)The performance of the Historical GMV portfolio in this sample of 43 stock universes is not typical of the average results of the approach. The section on simulations shows more representative results for this portfolio.
with a extremely high standard deviation of 1725.55. Both the Galton methods, the GMV and the MV, outperform the the 1/N portfolio in terms of Sharpe ratio and with reasonable levels of risk (standard deviation of 13.59 and 20.70, respectively).

The search of robust mean variance portfolios in OOS tests as received relatively more attention than an equally deserving question: the OOS predictability of the risk of those same portfolios. According with the two-fund separation theorem, in a conventional Markowitz setting, an hypothetical investor solves the allocation in problem in two steps: i) identify the tangency risky portfolio; ii) determine the allocation of wealth between the tangent portfolio and the risk free rate depending on the portfolio’s risk and the investor preferences. The first step has received a lot of attention, and the difficulties of finding the mean-variance efficient in a realistic OOS setting are well documented. Still, there are successful methods that achieve robust OOS performance in terms of Sharpe ratio (Brandt et al. (2009), DeMiguel et al. (2009a), Kirby and Ostdiek (2012)). But even if an investor is able to solve the first step using one of those methods, he is still left with the problem of how to estimate, ex ante, the risk of his chosen portfolio. For that second step, it does not matter if the standard deviation of the optimal portfolio is high or low, but it does matter if it is predictable.

Basak (2005) show that using the historical method to estimate the risk of the GMV results in a dramatical understatement its true risk OOS. Also, assuming a multivariate normal distribution, Kan and Smith (2008) show analytically that historical estimates of the risk and mean return of the GMV portfolio are overly optimistic. I examine this problem in an OOS setting, with real stock data, and add to the historical sample estimation two other methods: the constant correlation matrix of Elton and Gruber (1973) and my proposed correction using past OOS
errors.

Table 3 shows the ex ante risk of each of the portfolios and the respective ex post OOS risk. An investor wary of the fact that mean returns are difficult to estimate might decide to follow the advice of Jobson et al. (1979) and pursue a GMV strategy. This investor would be quite surprised to see that the strategy, in real time, has more than 7 times the risk he anticipates. The ex ante standard deviation of the historical GMV is of 3.85 percentage points (annualized) but the ex post standard deviation of the strategy is 26.90.

The problem is even worse for the historical approach MV portfolio with an OOS risk 69 times higher than the ex ante estimate. So a hypothetical investor following the historical approach to estimate risk should soon conclude there is something wrong with his estimates. A clear illustration of this is that 44.22 percent of the OOS returns are losses exceeding the investor’s ex ante estimate of the 1% level value-at-risk (VaR).

The Elton-Gruber constant correlation approach performs much better predicting the risk of stock portfolios. The standard deviation of the GMV portfolio is 15.35 percentage points, versus 8.34 expected ex ante. So in the case of stock portfolios, the Elton-Gruber approach substantially reduces the dramatic problems shown in Basak (2005) with historical sample estimates. Still, an investor in the robust Elton-Gruber GMV portfolio, would find on average 84% more volatility OOS than anticipated. The number of occurrences in the extreme quantiles of the distribution is much higher than anticipated too. In the left tail, losses exceeding his estimate of 1% level VaR would occur 7.45 times more often than anticipated, the null hypothesis that the hit rate is 1% is clearly rejected. For the MV portfolio, the constant correlation matrix does not capture accurately the OOS risk either.
The standard deviation OOS is 14 times higher than the ex ante estimate. So while the constant correlation matrix achieves a somehow acceptable performance describing risk of the GMV, an investor tempted to use it to estimate the risk of a MV portfolio would be dramatically surprised.

The third column shows the performance of the portfolios that use the covariance matrix (and mean returns) corrected for past OOS errors. The most striking result is that there is no significant difference between ex ante and out of sample risk. For the GMV portfolio, the expected standard deviation is 15.19 percentage points while the standard deviation in the OOS period is 13.59. So risk is actually lower OOS than expected for this portfolio. None of the hit rates is significantly higher than the respective target rates, two are even significantly lower in a statistical sense. Even for the case of the MV portfolio, OOS standard deviation (20.70 percentage points) is close in magnitude to the estimated ex ante (18.54 percentage points). Half the hit rates are not significantly different than their respective targets. So the covariance matrix corrected for past OOS errors captures well the risk of these optimized stock portfolios.

5. The result of simulations

The results in the previous section are based upon only one sequence of 43 stocks universes (each universe comprising 50 stocks) covering an OOS period of 526 months. In spite of the long period, there is still significant sampling error resulting from the fact that the individual stocks picked each year will be different. This is particularly important for the methods that rely on the plug in covariance matrix, as their estimates will be more noisy. To handle this I simulate 1000
such sequences of 526 months, resulting in 43,000 stock universes with a total of 526,000 OOS returns. All returns are OOS, so within each sequence the investor following a strategy only uses the available data up to the month in question. The simulations use the actual OOS returns of the stocks in each portfolio, so they do not assume a multivariate normal distribution data generating process as in, for example, Jobson and Korkie (1980) or Kan and Smith (2008).

Table 4 shows the summary of the performance of these strategies in the OOS simulations. On average the Sharpe ratio of the historical GMV is 0.24, much higher than the one in table 2. This difference illustrates how a single sequence of portfolios of 50 stocks, even with a long sample of 43 years, can be affected by randomness. Still the strategy has a Sharpe ratio smaller than the 1/N strategy and it only outperforms the naive strategy in 22% of the simulations. The investor expects, ex ante, a standard deviation of 7.48 percentage points, but OOS the actual standard deviation is 17.08 percentage points, more than double his expectation. In 100% of the simulations, OOS risk (as measured by the standard deviation) is higher than the ex ante expectation. So even if the strategy delivers a reasonable performance, it is consistently inferior to the naive portfolio in OOS simulations and it also surprises investors with risks substantially higher than they anticipate.

The Elton-Gruber constant correlation matrix GMV has a Sharpe ratio very close to the 1/N on average. In fact, in 54% of the simulations it outperforms in this metric the naive benchmark. It is remarkable as this approach to portfolio

---

8In unreported results I found that on average the 1/N strategy loads positively on the size and value Fama-French factors. This possibly contributes to its consistent performance.
9All values in the second column are statistically different from 50% in two-tailed tests at the 1% level.
management, proposed in 1973, has performed so well out of sample in the context of stock portfolios. Still the ex-post risk of the strategy is 44% higher on average than the ex ante estimates. So the constant correlation matrix systematically underestimates the risk of the GMV.

Both the historical and Elton-Gruber mean variance portfolios perform consistently badly OOS. The average Sharpe ratio is close to 0 and the risk OOS is more than 10 times higher the ex ante estimates from the respective covariance matrices. This shows that even the constant-correlation matrix has non negligible difficulties estimating the risk of concentrated stock portfolios.

The methods that correct past OOS errors have, on average, Sharpe ratios of 0.40 (GMV) and 0.36 (MV). Both are higher than the Sharpe ratio of the 1/N strategy (0.31). The differences in OOS performance are statistically significant at the 1% level and the second column shows the outperformance occurs in most simulations (85% for the GMV and 63% for the MV). The performance of MV portfolio is not has impressive as the GMV. This suggests the most relevant information for the optimization is in the covariance matrix, not the vector of expected returns.\textsuperscript{10} As there is not a lack of successful approaches to estimate expected returns, this is not a serious limitation of the correction proposed though. The relevant challenge lies in estimating the covariance matrix, which is exactly where the Galton correction proposed here seems to achieve an interesting performance.

For risk management purposes, the most relevant issue is the OOS predictability of risk for each stock portfolio. The results show ex ante estimated risk is, on average, 16.21 percentage points for the GMV and 18.99 percentage points for

\textsuperscript{10}Still the correction of past OOS errors seems to effectively filter out most of the noisy information in the estimation of past returns.
the MV. This compares to ex post OOS risk of 14.21 percentage points and 19.72 percentage points, respectively. So the most striking result is that OOS risk is close to the ex ante estimate when using the corrected covariance matrix. This contrasts sharply with the historical and constant correlation approaches. Even the MV portfolio, that shows dismal performance OOS and unpredictable risk with the other methods, achieves to outperform the 1/N on average. This is particularly noteworthy as it achieves this without using any portfolio constraints or stock characteristics.

Table 5 shows the hit rates in the OOS period for the 1000 simulations. For the historical and the constant correlation approaches, all hit rates for the extreme quantiles are statistically different than the target. In fact, extreme observations happen consistently more frequently than the ex ante risk estimate would suggest. Generally, the problem is more pronounced in the left tail than in the right tail, reflecting the left skewness of the strategies. Losses that should only happen 1% of the time, according with ex ante estimation of risk, occur with a frequency between 5.30% and 21.76% on average. One interpretation of these results is that both historical and constant correlation covariance matrices leave considerable scope for investors to be mislead (or mislead!) about the true OOS risk of their stock portfolios.

For the portfolios using past OOS errors to correct the inputs, the hit rates are, on average, close to the target levels. In fact, they are significantly below the target in 4 cases (the G-GMV at 10%, 90%, 95% and the G-MV at the 90%) and insignificantly different from the target in 6 cases. They only seem to capture insufficiently the risk in the extreme left tail. Losses that should happen with 1% probability occur in the OOS period with 1.40% and 1.80% frequency for the
GMV and MV portfolios, respectively. This is consistent with the interpretation that the distribution of returns is not multivariate normal. Still this is a relatively minor deviation in the VaR estimate when compared to the other methods. This shows that correcting the covariance matrix for past OOS errors has the potential to improve the estimation of risk of concentrated stock portfolios, in particular in extreme quantiles of the distribution.

6. Conclusion

The historical and the constant correlation approaches to the estimation of the covariance matrix have a major problem. They systematically underestimate risk. The actual OOS standard deviation of the strategies is on average from 2 to almost 29 times higher the ex ante estimate would suggest. The problem is more severe for the MV portfolios but is also pertinent in the GMV. This leads to a gross mis-estimation of risk, particularly at the extreme quantiles of the distribution. In simulations with real stock returns, losses exceeding the 1% level VaR occur in as much as 22% of the OOS periods (for the historical MV portfolio).

Correcting the covariance matrix for past OOS errors dramatically reduces this estimation problem. The ex post risk of the optimal portfolios is very close in magnitude to its ex ante risk. For most of the extreme quantiles, the hit rates are either not statistically different from the respective targets or even below them.

This suggests that correcting past OOS errors provides a simple method to estimate the covariance matrix of stock portfolios, with particularly useful applications in risk management.
Table 1
Regression to the mean

Regression of ex post values on historical estimates. For each month in the sample from 1955:03 to 2009:12, I regress the ex post value of some variable (computed using the subsequent 12 months of data) on its estimate from the historical sample in the previous 60 months. The variables in the columns are: i) the pairwise covariance of stock returns; ii) the pairwise correlation of stock returns; iii) the variance of individual stocks; iv) the mean return of the individual stocks. The rows show the output of Fama-MacBeth (1973) regressions of the ex post values on the historical estimates. The outputs are: i) the average slope coefficient in the monthly regressions; ii) the Newey-West (1987) t-statistic of the slope coefficient (computed with 12 lags); iii) the percentage of cross sectional regressions where the slope coefficient is significantly positive in a one-tailed test at a significance level of 5%; iv) the percentage of regressions where the slope coefficient is significantly smaller than one in a one-tailed test at a significance level of 5%; v) the average R-square of the regressions. Rows 6 to 8 show, respectively, the minimum, average, and maximum number of observations in the regressions.

<table>
<thead>
<tr>
<th></th>
<th>Covariance</th>
<th>Correlation</th>
<th>Variance</th>
<th>Mean return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>0.37</td>
<td>0.23</td>
<td>0.58</td>
<td>-0.18</td>
</tr>
<tr>
<td>t-stat(=0)</td>
<td>10.03</td>
<td>21.44</td>
<td>11.07</td>
<td>-4.83</td>
</tr>
<tr>
<td>Greater than 0 (%)</td>
<td>100.00%</td>
<td>100.00%</td>
<td>98.18%</td>
<td>18.54%</td>
</tr>
<tr>
<td>t-stat(=1)</td>
<td>-17.12</td>
<td>-71.16</td>
<td>-8.01</td>
<td>-32.06</td>
</tr>
<tr>
<td>Smaller than 1 (%)</td>
<td>93.01%</td>
<td>100.00%</td>
<td>86.93%</td>
<td>100.00%</td>
</tr>
<tr>
<td>R-square</td>
<td>4.16%</td>
<td>1.59%</td>
<td>13.24%</td>
<td>2.14%</td>
</tr>
<tr>
<td>Min</td>
<td>371091</td>
<td>371091</td>
<td>862</td>
<td>862</td>
</tr>
<tr>
<td>Average</td>
<td>3297770</td>
<td>3297770</td>
<td>2353</td>
<td>2353</td>
</tr>
<tr>
<td>Max</td>
<td>6579378</td>
<td>6579378</td>
<td>3628</td>
<td>3628</td>
</tr>
</tbody>
</table>
Table 2
OOS performance of the portfolio

Each twelve months I draw a random sample of 50 stocks among the 500 largest listed companies for which there is a complete return history over the previous 60 months and the subsequent 12 months. The sample is kept fixed for the subsequent twelve months. The weights are re-balanced monthly, each month I use a rolling window of 60 months to estimate the covariance matrix hence obtaining the global minimum variance (GMV) and the mean-variance (MV) portfolios for three different methods: the historical method, the Elton-Gruber method, and the Galton method. The columns show descriptive statistics of the out-of-sample performance of each portfolio. These are: i) the mean annual return of the portfolio; ii) the annualized standard deviation of the portfolio; iii) the excess kurtosis of the portfolio; iv) the skewness of the portfolio; and v) the Sharpe ratio of the portfolio. The sample returns are from 1955:03 to 2010:12.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Mean</th>
<th>STD</th>
<th>KURT</th>
<th>SKEW</th>
<th>Sharpe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/N</td>
<td>5.49</td>
<td>17.59</td>
<td>1.96</td>
<td>-0.21</td>
<td>0.31</td>
</tr>
<tr>
<td>Historical GMV</td>
<td>-2.24</td>
<td>28.38</td>
<td>1.94</td>
<td>-0.01</td>
<td>-0.08</td>
</tr>
<tr>
<td>Historical MV</td>
<td>-748.41</td>
<td>3322.89</td>
<td>335.39</td>
<td>-17.70</td>
<td>-0.23</td>
</tr>
<tr>
<td>Elton-Gruber GMV</td>
<td>4.75</td>
<td>15.35</td>
<td>1.50</td>
<td>-0.01</td>
<td>0.31</td>
</tr>
<tr>
<td>Elton-Gruber MV</td>
<td>150.47</td>
<td>1725.55</td>
<td>349.98</td>
<td>15.92</td>
<td>0.09</td>
</tr>
<tr>
<td>Galton GMV</td>
<td>4.99</td>
<td>13.59</td>
<td>2.72</td>
<td>-0.05</td>
<td>0.37</td>
</tr>
<tr>
<td>Galton MV</td>
<td>6.74</td>
<td>20.70</td>
<td>5.24</td>
<td>0.56</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Table 3
Accuracy of risk estimates

The first row shows the expected standard deviation of a strategy in the OOS period using the weights of the strategy and the respective estimate of the covariance matrix. The second row shows the realized standard deviation of the strategy computed from its OOS monthly returns. Both measures of standard deviation are annualized. Rows 3 to 5 show how often the strategy delivered returns smaller than the 1st, 5th, and 10th quantile, respectively, according to the ex-ante estimate of risk. Rows 6 to 8 show the same information for returns above the 90th, 95th, and 99th quantile, respectively. In the first column the variance is estimated from the historical sample, in the second column it is estimated with the Elton-Gruber correlation matrix, and in the third column with the Galton correction. Panel A shows the results for the global minimum variance portfolios and panel B for the mean-variance portfolios. The out-of-sample returns are from 1966:04 to 2010:01. All values are in percentage points. One star denotes significance at the 10% level, two stars at the 5% level, and three stars at the 1% level for the (two-tailed) test that the OOS hit rate is different than the expected ex ante.

<table>
<thead>
<tr>
<th></th>
<th>Historical</th>
<th>Elton-Gruber</th>
<th>Galton</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: The global minimum variance portfolio</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD(expt)</td>
<td>3.85</td>
<td>8.34</td>
<td>15.19</td>
</tr>
<tr>
<td>STD(realized)</td>
<td>28.38</td>
<td>15.35</td>
<td>13.59</td>
</tr>
<tr>
<td>r&lt;Qz(0.01)</td>
<td>40.43***</td>
<td>7.45***</td>
<td>0.95</td>
</tr>
<tr>
<td>r&lt;Qz(0.05)</td>
<td>44.98***</td>
<td>14.29***</td>
<td>4.56</td>
</tr>
<tr>
<td>r&lt;Qz(0.10)</td>
<td>46.96***</td>
<td>18.39***</td>
<td>8.75</td>
</tr>
<tr>
<td>r&gt;Qz(0.99)</td>
<td>37.84***</td>
<td>22.34***</td>
<td>6.27***</td>
</tr>
<tr>
<td>r&gt;Qz(0.95)</td>
<td>35.26***</td>
<td>16.26***</td>
<td>3.04**</td>
</tr>
<tr>
<td>r&gt;Qz(0.99)</td>
<td>32.98***</td>
<td>8.81***</td>
<td>1.14</td>
</tr>
<tr>
<td><strong>Panel B: The mean variance portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD(expt)</td>
<td>47.83</td>
<td>122.49</td>
<td>18.54</td>
</tr>
<tr>
<td>STD(realized)</td>
<td>3322.89</td>
<td>1725.55</td>
<td>20.70</td>
</tr>
<tr>
<td>r&lt;Qz(0.01)</td>
<td>44.22***</td>
<td>10.18***</td>
<td>2.66***</td>
</tr>
<tr>
<td>r&lt;Qz(0.05)</td>
<td>48.78***</td>
<td>21.12***</td>
<td>6.84*</td>
</tr>
<tr>
<td>r&lt;Qz(0.10)</td>
<td>51.37***</td>
<td>27.81***</td>
<td>11.22</td>
</tr>
<tr>
<td>r&gt;Qz(0.99)</td>
<td>32.67***</td>
<td>20.21***</td>
<td>8.17</td>
</tr>
<tr>
<td>r&gt;Qz(0.95)</td>
<td>30.85***</td>
<td>15.50***</td>
<td>4.56</td>
</tr>
<tr>
<td>r&gt;Qz(0.99)</td>
<td>28.57***</td>
<td>9.27***</td>
<td>1.71</td>
</tr>
</tbody>
</table>
Table 4
Results of simulations (OOS)

Each simulation selects randomly a set of 50 stocks and uses the information available up to each point in time to form a portfolio. It rebalances the portfolio monthly, reflecting the update in the information available, and every twelve months selects another set of stocks until the end of the 526 months OOS period. So a simulation comprises 43 stock universes and a total of 526 OOS monthly returns. The results are based on 1000 simulations totalling 526,000 OOS monthly returns. The first column shows the average Sharpe ratio across the 1000 simulations for the respective strategy. The second column shows the percentage of simulations the Sharpe ratio of the strategy was superior to that of the 1/N strategy. The third column shows the average expected volatility of the strategy using the ex-ante covariance matrix. Column 4 shows the actual ex post volatility of the strategy. Column 5 shows the percentage of simulations the ex post risk was higher than the one expected ex ante. The values in columns 3 to 5 are in percentage points.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>$SR$</th>
<th>$SR &gt; SR_{1/N}$</th>
<th>$E\sigma$</th>
<th>$\sigma_{expost}$</th>
<th>$\sigma_{expost} &gt; E\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/N</td>
<td>0.33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Historical GMV</td>
<td>0.24</td>
<td>0.22</td>
<td>7.48</td>
<td>17.08</td>
<td>100.00</td>
</tr>
<tr>
<td>Historical MV</td>
<td>0.00</td>
<td>0.00</td>
<td>193.02</td>
<td>5556.97</td>
<td>99.90</td>
</tr>
<tr>
<td>Elton-Gruber GMV</td>
<td>0.34</td>
<td>0.54</td>
<td>9.90</td>
<td>14.25</td>
<td>100.00</td>
</tr>
<tr>
<td>Elton-Gruber MV</td>
<td>0.01</td>
<td>0.01</td>
<td>243.15</td>
<td>3178.35</td>
<td>100.00</td>
</tr>
<tr>
<td>MRC GMV</td>
<td>0.40</td>
<td>0.85</td>
<td>16.21</td>
<td>14.21</td>
<td>0.00</td>
</tr>
<tr>
<td>MRC MV</td>
<td>0.36</td>
<td>0.63</td>
<td>18.99</td>
<td>19.72</td>
<td>80.20</td>
</tr>
</tbody>
</table>
### Table 5
Hit rates in the OOS simulations

Each simulation selects randomly a set of 50 stocks and uses the information available up to each point in time to form a portfolio. It rebalances the portfolio monthly, reflecting the update in the information available, and every twelve months selects another set of stocks until the end of the 526 months OOS period. So a simulation comprises 43 stock universes and a total of 526 OOS monthly returns. The results are based on 1000 simulations totalling 526,000 OOS monthly returns. Columns 1 to 3 show the average hit rates of each simulation for the shown critical level (the average percentage of observations with losses greater than the ex-ante estimated VaR). Columns 4 to 6 show the same information for hit rates above the respective critical level in the right tail of the distribution (that is gains exceeding the respective quantile of the ex-ante distribution). Three stars denote significance at the 1% level, two stars at the 5% level and one star at the 10% level. All tests are two-tailed tests. The null hypothesis is that the hit rate is equal to the respective critical level. The rows show the results for the following portfolios: i-ii) the historical global minimum variance (‘H-GMV’) and mean variance (‘H-MV’); iii-iv) the constant correlation matrix global minimum variance (‘EG-GMV’) and mean variance (‘EG-MV’); v-vi) Galton corrected global minimum variance (‘G-GMV’) and mean variance (‘G-MV’).

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Hit &lt;1%</th>
<th>Hit&lt;5%</th>
<th>Hit&lt;10%</th>
<th>Hit&gt;90%</th>
<th>Hit&gt;95%</th>
<th>Hit&gt;99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-GMV</td>
<td>14.58***</td>
<td>22.44***</td>
<td>27.70***</td>
<td>25.93***</td>
<td>20.84***</td>
<td>13.37***</td>
</tr>
<tr>
<td>H-MV</td>
<td>21.76***</td>
<td>31.66***</td>
<td>37.80***</td>
<td>19.95***</td>
<td>16.10***</td>
<td>10.47***</td>
</tr>
<tr>
<td>EG-GMV</td>
<td>5.30***</td>
<td>11.15***</td>
<td>16.20***</td>
<td>17.24***</td>
<td>11.73***</td>
<td>5.39***</td>
</tr>
<tr>
<td>EG-MV</td>
<td>8.06***</td>
<td>16.86***</td>
<td>24.12***</td>
<td>14.03***</td>
<td>10.04***</td>
<td>5.27***</td>
</tr>
<tr>
<td>G-GMV</td>
<td>1.40***</td>
<td>4.36</td>
<td>7.73**</td>
<td>5.42***</td>
<td>2.62***</td>
<td>0.74</td>
</tr>
<tr>
<td>G-MV</td>
<td>1.80**</td>
<td>5.32</td>
<td>9.64</td>
<td>7.67**</td>
<td>4.39</td>
<td>1.46</td>
</tr>
</tbody>
</table>
Fig. 1. Historical versus ex post values. For each moment in time and observation (either a stock or a pair of stocks) I compute the value of a variable in the historical sample - the previous 60 months of observations - and its respective value in the future - the subsequent 12 months. Then observations are classified into deciles according to their values in the historical sample each month. The figure shows, for different variables, the time series average of the values in each decile of historical and respective ex post realizations. In panel A and B the observations are pairs of individual stocks while in C and D they are the stocks themselves. Panel A shows the historical covariances versus the ex post covariances of the returns of individual stocks. Panel B shows the same comparison for the pairwise correlations of returns. Panels C and D show the same comparison for, respectively, the variance of stocks and the mean total return. The data consists of monthly observations from 1955:03 to 2009:12.
References


Victor DeMiguel, Lorenzo Garlappi, Francisco J Nogales, and Raman Uppal.


