Learning from past Out-of-Sample Errors: An application of the Galton method in the United Kingdom stock market

Gennaro Celardo
152222028

Dissertation written under the supervision of Pedro Monteiro E Silva Barroso

Dissertation submitted in partial fulfilment of requirements for the MSc in Economics with Major in Finance and Banking, at Universidade Católica Portuguesa and for the MSc in Economics and Finance at Università degli Studi di Napoli Federico II, 1 June 2023.

Learning from past Out-of-Sample Errors:
An application of the Galton method in the United Kingdom stock market

Gennaro Celardo

Abstract

This Thesis is focused on an out-of-sample application of the Galton strategy in the United Kingdom stock market from January 1996 to December 2022. This allocation, exploiting the lack of perfect randomness of past out-of-sample errors, is able to provide a useful alternative to the classical plug-in approach in portfolio optimization. A crucial advantage of this investment strategy is the better risk-adjusted performance with respect to the benchmark and a classic version of Mean-Variance portfolio in terms of a higher Sharpe and Sortino ratio. Galton risk estimates are not too optimistic in predicting future volatility as opposed to competitors, recording a ratio of realized volatility over its ex-ante expectation close to unity. Five hundred random horse races confirm these results. In this context, the Ledoit and Wolf portfolio is the sole competitor beating in 26% and 61% of cases the Galton GMV and MV versions in terms of annualized Sharpe ratio. From a Risk Management perspective, Galton allocations have the best VaR hit rates at 95% and 99% confidence levels.

Keywords: Portfolio Optimization; Estimation Errors; Galton Strategy; Shrinkage Estimator

Learning from past Out-of-Sample Errors:

An application of the Galton method in the United Kingdom stock market

Gennaro Celardo

Sumário

A tese está focada numa implementação out-of-sample da estratégia de Galton no mercado de ações do Reino Unido desde Janeiro 1996 até Dezembro 2022. Esta optimização de carteiras, que explora presença de previsibilidade dos erros out-of-sample do passado, é capaz de dar uma alternativa útil para a otimização de portfólios clássica. Uma vantagem crucial dessa estratégia de investimento é o melhor desempenho ajustado ao risco comparando com a benchmark e a uma versão clássica do portefólio média-variância em termos dos índices de Sharpe e Sortino mais elevados. As estimativas do risco de Galton não são excessivamente otimistas para prever a futura volatilidade em relação à competição, tendo um rácio de volatilidade realizada sobre a expectativa ex-ante abaixo da unidade. Quinhentas corridas de cavalo aleatórias confirmam estes resultados. Neste contexto, o portefólio de Ledoit e Wolf é o único concorrente batendo em 26% e 61% dos casos as versões Galton GMV e MV em termos de Sharpe ratio anualizado. De uma perspectiva de gestão de risco, as alocações de Galton têm o melhor VaR a níveis de confiança de 95% e 99%.

Palavras-chave: Otimização de Portefólio; Erros de Estimativa; Estratégia de Galton; Estimador de Contração

Contents

Introduction

Markowitz (1952) lays the groundwork for the modern portfolio theory, sparking, after the publication of his work, a vast amount of research in the asset allocation field that continues today. Based on a set of assumptions, in Markowitz's framework, investors can decide how to allocate their wealth using two fundamental parameters: the mean and the covariance of the distribution of future returns. The former variable can be interpreted as a proxy of the reward that an investor expects to receive for bearing the risk, measured in terms of contribution to portfolio volatility, of a certain financial asset. The latter factor is a useful statistical measure able to capture the degree of linear similarity between two assets and it is a necessary input to compute the standard deviation of the portfolio. These two variables are the only required moments of distribution needed to determine the entire set of attainable portfolios given all the risky assets that are available in the financial universe. Within these possible allocations, there exist some which are efficient accordingly to the maximization problem that enables the investors to attain the highest expected return by fixing a certain level of risk. Alternatively, the same outcome can be obtained by employing a minimization problem (it is also known as the Dual Problem) which aims to find the portfolios with the lowest risk maintaining a certain level of expected return.

It is pivotal to take into consideration a significant limitation when implementing this model in the real world: the replacement of the real distribution parameters (i.e., mean and covariance matrix) with their sample estimates extracted from the historical sample. This substitution would lead consequently to the presence of estimation errors in the optimization techniques used in portfolio theory. As Stein (1955) points out in his research, the choice of the sample mean is suboptimal, namely, it is not "admissible", because, under specific regularity conditions, there exist other better estimators that can achieve lower levels of the loss function.

Michaud (1989) demonstrates that the mean-variance optimization technique amplifies the estimation errors computed by investors leading them to choose less efficient allocations. This phenomenon arises because investors assign greater (lesser) weights to assets that have recorded positive (negative) performance in the past, exhibit lower (bigger) sample variance or display negative (positive) sample correlation. Jobson and Korkie (1980) appraise the magnitude of estimation errors by using Monte Carlo simulations and a known distribution of monthly returns. They find that the Sharpe ratio, computed as the average of this metric for all the simulations, significantly differs from the one calculated based on the known distribution.

Barroso and Saxena (2021) are the first two researchers to apply the Galton strategy to solve the estimation error issue in the United States stock market. In their paper "Lest we forget: learn from out-of-sample forecast errors when optimizing portfolios", they adjust the inputs of mean-variance portfolios drawn from the historical sample by the past out-of-sample errors. Being these not completely random, the scholars prove the possibility of constructing more stable and efficient portfolios. Their methodology yields portfolios with the highest Sharpe Ratio and the lowest root mean squared forecast error (RMSFE) among a wide spectrum of alternative investment strategies.

The Galton optimization draws inspiration from an extensive body of literature on portfolio optimization and estimation error and the first empirical application of this model in the original paper is limited to United States. Goyal and Wahal (2015) prove that findings obtained within the US context do not always replicate when tested using other international market. Furthermore, Jacobs and Muller (2020) show that the effect of some anomalies disappear after their discovery. Therefore, to assess the effectiveness of this strategy in a different major market, I employ the Galton method in the United Kingdom stock market. The time span starts in January 1996 and ends in December 2022. I compute Galton Mean-Variance, Galton Global Minimum Variance and Galton Empirical Bayes using the ten, thirty and fifty most capitalized companies of the United Kingdom stock market.

I find that the Galton allocations outperform the benchmarks (the naïve 1/N and the value-weighted portfolios) in most of the cases in terms of Sharpe Ratio, Sortino Ratio and maximum drawdown. Furthermore, Galton portfolios are the best portfolios in terms of higher Morningstar Risk-Adjusted Return measure (MRAR) which returns the equivalent risk-free rate that a constant relative risk aversion (CRRA) investor would accept for the performance of its portfolio. The returns of these allocations rarely exceed the Value at Risk hit rates at 95% and 99% confidence levels. After the Galton correction, the vector of weights displays greater stability and requires fewer negative positions compared to the simple versions. Examining the alphas of these investment strategies, they are all positive but not statistically different from 0 using both the Capital Asset Pricing Model and 3 Factor Models as benchmarks to construct expected return. The outcome slightly changes when I use the value-weighted portfolio instead of the FTSE All share index, with Galton allocations recording statistically significant alphas for some portfolio sizes.

Literature Review

In this section, I examine the main problems related to the classical plug-in approach and subsequently the major benefits of the Galton portfolios and other strategies developed by scholars to try to create optimal portfolios in an out-of-sample context. I investigate a wide range of alternatives, varying from the introduction of new constraints to the shrinkage of the mean vector or of the variance-covariance matrix.

2.1 The mathematical Framework of Markowitz optimization

Under the feasibility condition that the sum of all weights is one and assuming a prior knowledge of the true population mean and variance-covariance matrix, the original formula utilized for the computation of the vector of weights of the Mean-Variance portfolio (MV) is the following:

$$\omega_t = \frac{\Sigma_t^{-1}\mu}{|\mathbf{1}\Sigma_t^{-1}\mu|} \tag{1}$$

where ω_t is the Nx1 vector of weights, μ is the Nx1 vector of the true mean, $\mathbf{1}$ is a 1xN vector of all ones and Σ_t^{-1} is the NxN inverse of the true variance-covariance matrix. However, in practice, the two unknown inputs of Equation 1 are typically replaced with their historical estimates $\bar{\mu}$ and $\bar{\Sigma}_t$, thus resorting to the classical "plug-in" approach. Furthermore, during the horse race, to avoid the change in the sign of the weights, I compute the absolute value of the denominator as in Barroso and Saxena (2021).

2.2 The Normality Assumption of Financial Returns

One potential solution to address this fundamental issue of estimating the population parameters consists of the assumption that the returns of financial assets are independently and identically drawn from a Normal Distribution. In this case, the sample mean and covariance (hence, also the standard deviation) will be the Maximum Likelihood Estimators (MLE) of their respective population parameters.

The sample mean and the covariance of a Normal Distribution have important and desirable statistical properties: i) the variance of these estimators is equal to the Cramér-Rao Lower Bound (CRLB) which corresponds to the lower bound of the variance of any unbiased estimators of a certain parameter, ii) they are consistent, this means that as the size of the sample raises, the distribution of the estimator becomes more condensed near the true value of the parameter. Unfortunately, many studies highlight some stylized aspects that confirm this assumption to be too unrealistic. Mandelbrot (1963) and Barnea and Downes (1973) show that the distribution of financial assets often exhibits a fat tail shape implying that extreme events can occur with a higher probability than the one incorporated in the Normal Distribution. The presence of negative skewness which involves that the likelihood of generating negative return is higher than what is expected using a symmetric distribution has led Rubinstein (1973) to create a model in which the expected return incorporates the skewness. Fang and Lai (1997) elaborate a pricing model that incorporates also the third and fourth moment of the distribution of returns showing that expected returns are also a function of the systematic skewness and kurtosis. Furthermore, in distressed periods with high volatility, some empirical research (Engle (1993), Bollerslev and Zhou (2006)) prove the existence of serial correlation among the returns of financial assets leading to a disruption of the assumption of independence. The abovementioned factors contribute to a decrease in the degree of similarity between the sample distribution of returns and the Gaussian one, making it very hard to rely on the MLE estimators of the true population parameters.

2.3 The Out-of-Sample Performance

Another direct consequence of using the sample mean and covariance instead of the true population parameters is the poor performance registered by Markowitz's portfolio out-of-sample (OOS). Out-of-sample optimization is a technique that relies solely on the usage of the available information in real-time to construct the vector of weights. Utilizing an expanding or rolling window, it is possible to appreciate how much the estimates of these two parameters are unstable and lead to estimation errors. Consequently, it is possible to obtain that a global minimum variance portfolio attains a higher Sharpe Ratio than a Markowitz portfolio which initially aimed to maximize that metric by selecting the same financial assets in the identical time horizon. Kan and Zhou (2007) study the poor OOS performance reporting the losses in average out-of-sample utility (OOSU) of Mean-Variance investors. DeMiguel et al (2009B) demonstrate in their work that in an out-of-sample context, the benefits of employing complex portfolios optimization techniques are offset by estimation errors, resulting in an intriguing finding: no complex allocation strategy consistently outperforms the naïve 1/N allocation in terms of Sharpe Ratio and other critical metrics for investors, such as the turnover of the portfolio.

2.4 The history of regression to the mean

Sir Francis Galton (1822-1911) is a pioneering figure in the fields of genetics and statistics. His studies focus on eugenics and the inheritance of traits. He develops many concepts that are still relevant today in statistical analysis, such as correlation and percentiles. In 1886, Sir Galton publishes a compelling anthropological study titled “Regression Towards Mediocrity in Hereditary Stature” in which he finds a connection between parental and offspring height. These two variables do not vary randomly, but rather they follow some patterns. Specifically, he shows that offspring height is a constant fraction of the deviation of the height of parents from the mean. On average, the offspring of tall parents, whose heights are higher than the average, have a lower height convergent to a mediocre point that corresponds to the mean. In statistical terms, given the occurrence of an extreme event, the likelihood of observing an even more extreme event next time is very low and it is more probable an outcome close to the mean. Nobel Prize winner in psychology, Daniel Kahneman analyzes in “Thinking Fast and Slow” some biases that can impact the human mind when it comes to the regression toward the mean. The scholar points out that the existence of the statistical concept of regression toward the mean is based on the absence of a perfect correlation (positive or negative) between two variables. Furthermore, this regression can occur over long horizons, implying that it may be very detrimental to extrapolate relationships from small samples that might not be fairly representative of the true distribution.

In the portfolio optimization universe, the Galton strategy capitalizes on the favorable property of learning from past OOS errors, which starts from the beginning of the learning period and continues during the length of the horse race. It is possible to correct the inputs by looking at the past as OOS errors do not follow a total random scheme. The usage of these errors allows investors the construction of robust portfolios that deliver coherent risk-adjusted performance even in a demanding out-of-sample context where they must rely only upon available information to allocate their wealth across all the possible alternatives. All the other investment strategies involved in the horse race fail to account for the errors that they commit over time.

By minimizing the Mean Squared Forecast Error (MSFE), the Galton allocation reaches also a better forecast of risk, avoiding the tendency of being too optimistic. In contrast, the MV and GMV strategies, based on sample estimates, exhibit huge inconsistencies between the ex-ante volatility and its ex-post realization. However, Galton’s ability to forecast risks seems to be decreasing in the number of companies involved in the optimization process. This issue is probably related to a higher T/N ratio where T represents the order of the variance-covariance matrix and N is the sample size. If the size of the matrix is large enough, one viable option consists in finding alternative estimators of the variance-covariance matrix, such as the one proposed by Ledoit and Wolf (2004). Nevertheless, in this work, the Galton portfolios register better performance in terms of lower realized volatility over the ex-ante expectation than the other compared portfolios across all the different sizes of the investment universe.

2.5 The Global Minimum Variance Portfolio as substitute of Mean-Variance Portfolio

One possible approach to partially getting rid of a part of estimations errors involves the adoption of the Global Minimum Variance (GMV) portfolio. This allocation is simply based on the minimization of the variance of portfolio under the linear constraint that the sum of weights is equal to one. The problem is formulated as follows:

$$\min_{\omega} \sigma^2(\omega)$$

$$Subject\ to\ \omega'\mathbf{1} = 1$$

The solution to this optimization problem implies the following vector of weights:

$$\omega_t = \frac{\Sigma_t^{-1}\mathbf{1}}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}} \tag{2}$$

Following the classical plug-in approach, the true and unknown population covariance matrix is replaced by its most common estimator: the sample covariance matrix. However, this portfolio allows to reduce the sampling errors since the weights depend only on the sample variance-covariance matrix and not on the sample mean. Additionally, Merton (1980) underlines an important advantage pertaining to the usage of this portfolio demonstrating that the variance and covariance matrix exhibit a greater stability over time compared to the sample mean.

2.6 The introduction of Constraints in Portfolio Optimization

Jagannathan and MA (2003) state in their research that the introduction of non-negative constraints or upper-bound limits to weights is an indirect alternative of shrinking the variance-covariance matrix. By incorporating such constraints, the new variance-covariance matrix would contribute to the reduction of sampling errors. Moreover, short sale constraints would ensure only the computation of positive weights while upper bound limits make the optimization process more similar to reality, since most of the institutional investors like pension and mutual funds work with these constraints. The abovementioned constraints, reducing the presence of extreme events, are fundamental to obtain a higher risk reduction for large portfolios.

2.7 The Shrinkage techniques

An important branch of Finance is studying how to adopt shrinkage techniques in the portfolio optimization problem. Stein (1955) introduces the concept of shrinkage as a solution to the problem of inadmissibility of the sample mean estimator when the number of stocks is greater than 3. The scholar proves that by shrinking the sample mean estimator towards a target value, it is possible to get a new estimator which has better performance in terms of mean squared error, particularly for large portfolios. Subsequently, Jorion (1986) shows the benefits of the James-Stein estimator for the mean vector adopting a quadratic loss function.

Ledoit and Wolf (2003) extend this field proposing the usage of shrinkage for the construction of a variance-covariance matrix. The most extreme values of the matrix are pulled toward more central ones trying to reduce both negative and positive sampling errors. They show that the portfolio constructed according to this rule has a higher appraisal ratio, an important objective of active investors. They find a compromise (which corresponds to the shrinkage intensity) using a linear combination between the sample variance-covariance matrix and a structured estimator F. The equation of the shrunk matrix (Σ_t^*) is the following:

$$\Sigma_t^* = \gamma * \Sigma_t + (1 - \gamma) * \mathrm{F} \tag{3}$$

Where γ is a number between 0 and 1. An example of the shrinkage target can be an estimator that assumes constant correlation for each entry out of the diagonal of the variance-covariance matrix. This means that all the pairwise correlations between the N stocks are the same and corresponding to the mean of all the sample correlations. The shrinkage intensity, instead, is obtained by the minimization of the expected value of a loss function based on the Frobenius norm.

In order to have a fairer horse race, I adopt the approach of Barroso and Saxena (2021) to compute the weights of a Mean-Variance investor which uses the Ledoit and Wolf matrix. Rather than relying on the historical mean vector, I employ the positive part of the James-Stein shrinkage estimator (James and Stein (1992)). This is a biased estimator of the mean that for N greater or equal than 4 achieves a lower mean squared error than the unbiased maximum likelihood estimator. Equations 4 and 5 outline the calculation of this estimator:

$$\mu^{JS} = \mu_0 \mathbf{1} + \beta_{1,JS}(\mu_H - \mu_0 \mathbf{1}) \tag{4}$$

$$\beta_{1,JS} = Max\left(1 - \frac{(N-3)\sigma_0^2/T}{\|\mu_H - \mu_0\mathbf{1}\|^2}, 0\right) \; for\; N \geq 4 \qquad (5)$$

Where μ_H corresponds to the historical mean vector of dimension 1xN, μ_0 is a scalar which represents the average cross-section of all the historical returns considering all the stocks and it is multiplied by a 1xN unity vector. σ_0^2 is another scalar which corresponds to the average cross-section of the historical variance of all returns in the panel data.

DeMiguel et all (2013) propose an alternative shrinkage technique applied to portfolio's weights. They adjust the Mean-Variance weights (ω_{sp}) accordingly to a certain shrinkage intensity ($\propto$) and a scale parameter (τ) which is used to minimize the bias of the target (ω_{tg}) and consequently, the loss function. Also in this case, the parameter $\propto$ can be obtained both by the minimization of the expected quadratic loss function and by the maximization of the Sharpe Ratio. Equation 6 represents the final formula applied by these scholars to derive the vector of weights:

$$\omega_{sh} = (1-\propto) * w_{sp} + \propto \tau\omega_{tg} \qquad (6)$$

Methodology

3.1 Data

In this Thesis, the analysis is based on the out-of-sample performance of real market data of stocks listed in the United Kingdom. I download the Total Return Index (RI) from Refinitiv Eikon DataStream, which also incorporates the reinvestments of dividends for all the stocks that have been constituents of the FTSE All Share index for at least one month. The analysis starts in January 1996 (the first available data for the constituents offered by Refinitiv Eikon although the index was founded in November 1962) and it ends in December 2022, thus including 324 monthly observations. The FTSE all Share index is composed of three main subsets: the FTSE 100 which comprises the 100 companies with highest market capitalization, the FTSE 250 which includes the next 250 firms according to size and the FTSE SmallCap that groups the smallest companies belonging to the index. Moreover, the FTSE All Share[1] covers more than 98% of the UK total market capitalization making it a reliable proxy to measure the

[1] https://www.ftserussell.com/products/indices/uk

performance of this country. After the cleaning of the dataset in which I do not consider all the investment trusts, the total number of companies is 1648.

For the computation of the Galton coefficients, in each month, I select only the companies that recorded at least 72 observations without missing values, thus also including those that were delisted or subject to extraordinary finance operations. On average, 529 companies meet this requirement each month.

I use the United Kingdom Sterling 1 Month deposit (Symbol: ECUKP1M) as the risk-free rate and the FTSE All Share (Symbol: FTALLSH) as a proxy of the market portfolio. I download the value and size factors for the United Kingdom from Global Factor Data[2]. Since this website offers the excess returns of factors for more than 93 countries in USD, I convert them according to the exchange rate between USD and GBP. These factors will be used for the computation of the alphas using the Fama and French (1992) model.

Following Barroso and Saxena (2021), a rolling window is used to appraise the financial performance of various portfolios in an out-of-sample context. At any given time t, I utilize the historical window composed by the previous 60 months to estimate all the inputs that are required to form the vector of weights. In the following period t+1, I shift the rolling window one month forward and again the same process is repeated.

I compute the arithmetic returns of the Total Return Index for all companies using the following equation:

$$r_{i,t} = \frac{RI_{i,t} - RI_{i,t-1}}{RI_{i,t-1}} \tag{7}$$

For the calculation of the risk-free rate in percentage points, I divide the time series of the UK Sterling 1 Month Deposit Middle Rate by 1200. Then, I subtract the risk-free rate to simple returns to obtain excess returns. The final dataset of excess returns comprises 323 observations since the first one is lost while computing the returns.

3.2 The estimation of Galton Coefficients

For the construction of the Galton coefficients, I use a rolling historical window (H) of 60 months to calculate the past estimates of the sample mean, standard deviation, variance, correlation and covariance. Subsequently, I determine the ex-post realizations of these variables using the following 12 months (E) which correspond to the length of the ex-post window.

[2] https://jkpfactors.com/?country=gbr

In doing so, I only consider the companies with 72 consecutive observations, thus excluding the firms with missing values because they are perhaps too illiquid or because they had been delisted in the meantime. As a result, the dimensions of the variables involved in the regressions may vary over time. Moreover, this technique requires a learning period (L) composed by 108 months. It is a necessary time window to stabilize the values estimated by the regressions from the presence of possible outliers. Consequently, the horse race begins after the 181st (which corresponds to the sum of H+L+E+1) observation, although the construction of the other portfolios such as the naïve 1/N or the value-weighted portfolio do not require that time. Therefore, the horse race lasts 143 months.

In order to obtain accurate appraisal of the Galton coefficients, I use the two steps Fama-MacBeth estimation process elaborated in their paper "Risk, Return and Equilibrium: Empirical Tests" which is also utilized by Professors Barroso and Saxena in their main research. This methodology seeks to minimize the mean squared forecast error (MSFE) using the ordinary least square (OLS) regressions. The equation of MSFE is described below:

$$Mean\ Squared\ Forecast\ Error = E[(X_{H,t} - X_{E,t+E})^2]$$

Where X is a generic variable that can represent the mean, standard deviation (or the variance), the correlation or the covariance matrix.

In the first step, to compute the values required for the Galton correction, I run the following OLS regressions for the mean, standard deviation and variance.

$$\mu_{E,t+E} = \gamma_\mu + g_\mu \mu_{H,t} + \varepsilon_{i,t} \tag{8}$$

$$\sigma_{E,t+E} = \gamma_\sigma + g_\sigma \sigma_{H,t} + u_{i,t} \tag{9}$$

$$\sigma^2_{E,t+E} = \gamma_{\sigma^2} + g_{\sigma^2} \sigma^2_{H,t} + \eta_{i,t} \tag{10}$$

Where μ, σ and σ^2 are three N_tx1 vectors and N_t represents the total number of liquid companies available at time t. It is important to point out that during the computation of Galton coefficients, the investors will use only the available information at time t and thus the variables involved in the regressions are lagged twelve months before to ensure that OLS regressions meet this requirement. Furthermore, the adjustment imposed by the Galton strategy has an insightful Bayes interpretation. In this context, the prior is related to the idea that all the variables that can be used in the optimization process (mean, standard deviation and so on) are equal for all the assets. This would inevitably lead to a naïve 1/N allocation as the cross-section average of a particular variable would be the same for each company.

In addition, if the value of the variable measured in the ex-post window is close to the one assumed in the historical window, then the slope coefficient approaches to 1 and the intercept becomes null. For this reason, in Table 1, I compute the t-statistic for the slope coefficient under the null hypothesis that it is equal to 1. Additionally, I calculate the number of times the estimate is statistically lower than 1 at 95% confidence level. Conversely, if the value of the variable in the historical window is not a good prior, the slope coefficient tends towards 0 and the ex-post observation can converge to the mean of the cross-section. To examine the validity of the historical prior, I also construct the t-statistic for the slope coefficient under the null hypothesis that it is equal to 0.

The second part of the procedure begins after running these regressions for the entire length of my dataset and consists of synthesizing all the information contained in the OLS regressions into two unique values which correspond to the expanding mean of the intercept and of the slope. The computation of these averages starts at time s which I use to indicate the first available period after the end of learning period.

$$\bar{\gamma}^{G}_{X,t} = \frac{\sum_{s=1}^{T} \gamma_{X,s}}{t} \tag{11}$$

$$\bar{g}^{G}_{X,t} = \frac{\sum_{s=1}^{T} g_{X,s}}{t} \tag{12}$$

At this point, I incorporate the expanding mean of the intercept and slope into the historical estimates to generate the Galton mean and standard deviation that are plugged into the MV and GMV formulas.

$$\mu^{G}_{t} = \bar{\gamma}^{G}_{\mu,t} + \bar{g}^{G}_{\mu,t}\mu_{H,t} \tag{13}$$

$$\sigma^{G}_{t} = \bar{\gamma}^{G}_{\sigma,t} + \bar{g}^{G}_{\sigma,t}\sigma_{H,t} \tag{14}$$

I reiterate the two steps Fama-MacBeth procedure also for the correlation and covariance matrix, although I never use the latter variable in the optimization process for the construction of weights. This choice stems from the uncertainty of obtaining a positive definite covariance matrix following the Galton adjustments. A covariance matrix which is not positive definite could result in the computation of negative variances.

Firstly, I vectorize the upper triangle of historical and ex-post realizations of the correlation matrix (I implement the same procedure for the covariance matrix). This process would allow to pick up only the unique pair-wise correlations without including the ones on the main diagonal and the values in the lower triangle of the N_txN_t matrix. The resulting vector has dimensions [N_t*(N_t-1)/2]x1. At this time, having the dependent and independent variable in vector notation, I run the following regressions:

$$\rho_{E,t+E} = \gamma_\rho + g_\rho \rho_{H,t} + \varepsilon_{i,t} \tag{15}$$

$$cov_{E,t+E} = \gamma_{cov} + g_{cov} cov_{H,t} + \varepsilon_{i,t} \tag{16}$$

Once I store all the intercepts and the slopes, I compute the expanding average of these variables using the equations 11 and 12. The pair-wise Galton correlation will be computed as:

$$\rho_t^G = \bar{\gamma}_{\rho,t}^G + \bar{g}_{\rho,t}^G \rho_{H,t} \tag{17}$$

To ensure a positive definitive N_txN_t Galton Variance-Covariance matrix, I apply the following linear algebra expression:

$$\Sigma_t^G = diag(\sigma_t^G)\, \rho_t^G\, diag(\sigma_t^G) \tag{18}$$

Where the diag(·) function is an operator that creates a square matrix N_txN_t where all the elements out of the main diagonal are zeros and the diagonal elements correspond to the elements of the vector of the standard deviation.

An alternative to this formula is represented by the usage of the Galton Variance instead of the standard deviation in the following equation:

$$\Sigma_t^G = \sqrt{diag(\sigma_t^{2G})}\, \rho_t^G\, \sqrt{diag(\sigma_t^{2G})} \tag{19}$$

Both formulas are theoretically correct, but I choose to use the equation 18 for the computation of the variance-covariance matrix because it enables to create portfolios with a slightly higher Sharpe Ratio.

Table 1 summarizes the most important average statistics derived by running 251 OLS regressions of ex-post values on their historical estimates for each variable.

Table 1: Summary statistics of OLS regressions of ex-post realizations on the historical values.
Intercept is the average value of all the intercepts estimated through the OLS regressions. T-stat (=0) reports the t-statistic under the null hypothesis that the intercept is statistically different from 0. Sign Intercept reports in decimal terms how many times the intercept is statistically different from 0 at 5% significance level. Slope is the average value of all the slopes estimated through the OLS regressions. T-stat (=0) indicates the t-statistic under the null hypothesis that the slope coefficient is equal to 0. Greater than 0 reports in decimal terms how many times the slope is statistically positive at 5% significance level. T-stat (=1) indicates the t-statistic under the null hypothesis that the slope coefficient is equal to 1. Smaller than 1 reports in decimal terms how many times the slope is statistically lower than 1 at 5% significance level. R-square indicates the average R-square expressed in decimal points. The last three rows indicate respectively the average, the minimum and the maximum number of observations involved in the OLS regressions for each variable.

	Mean	Variance	Standard Deviation	Correlation	Covariance
Intercept	0.01	0.05	0.05	0.12	0.00
T-stat (=0)	6.27	2.21	29.64	23.01	11.57
Sign Intercept	0.75	0.89	0.96	1.00	1.00
Slope	-0.06	0.67	0.53	0.28	0.32
T-stat (=0)	-2.36	4.48	34.85	51.82	10.49
Greater than 0	0.28	0.76	0.99	1.00	0.98
T-stat (=1)	-39.72	-2.17	-30.32	-132.85	-22.51
Lower than 1	0.99	0.78	0.88	1.00	0.88
R-square	0.02	0.09	0.21	0.03	0.02
Average Obs	529	529	529	285,670	285,670
Minimum Obs	409	409	409	167,281	167,281
Maximum Obs	673	673	673	452,929	452,929

An interesting aspect that can be noticed by looking at the Table 1 is the presence of mean reversion in the United Kingdom stock market since the average slope coefficient for the mean is equal to -0.06. The profitability of the value strategy corroborates this result. In the same time span of my thesis, the value factor, which consists in being long in undervalued stocks (high Book-to-Market ratio) and short in overvalued firms (low Book-to-Market ratio), returns a positive annual Sharpe Ratio of 0.33. This outcome is perfectly coherent with the results of Table 1 since some scholars (Asness et all (2013), Fama and French (1992)) find that the mean reverting attitude of the stocks can be a possible cause of the good performance of a such factor.

By analyzing the other variables, it can be observed that the variance, standard deviation and covariance exhibit the greatest stability, as evidenced by Merton (1980), with higher slope coefficients among the five variables. All the slope coefficients are on average statistically different from 0 meaning that the historical prior can be enhanced with some adjustments. By learning from past OOS errors, the Galton Strategy is able to capture this improvement.

Figures 1 to 5 illustrate the values assumed by the slope, the intercept, the expanding average of the slope and of the intercept across the different regressions.

Notably, the estimated slope coefficients for the covariance exceed one in two distinct time clusters: the first jump occurred around 89th and 100th regressions and the second one took place from the 215th to 231st regressions. These periods correspond to major crises such as the Global Crisis in 2008-2009, the Brexit and the Covid-19 in 2019-2020, which induced outliers in the ex-post window. Furthermore, the slope coefficients of variance and standard deviation take extreme values around the 40th regression. These events demonstrate the importance of the learning period in the Fama-MacBeth methodology, a necessary time to create robust coefficients that can accommodate the outbreak of some turmoil. Indeed, after the learning period, the expanding average of the slope coefficient for covariance (as well as for other variables except for the mean) is always between 0 and 1, indicating that an investor who is applying the Galton strategy, also after the occurrence of some extreme periods, is benefiting from a shrinking technique plus the reduction of risk by the minimization of the MSFE.

Figure 1: Galton Coefficients for the Mean

This figure shows the estimated intercept, the estimated slope, the expanding average for intercept and slope for all OLS regression of ex-post mean on the historical estimate.

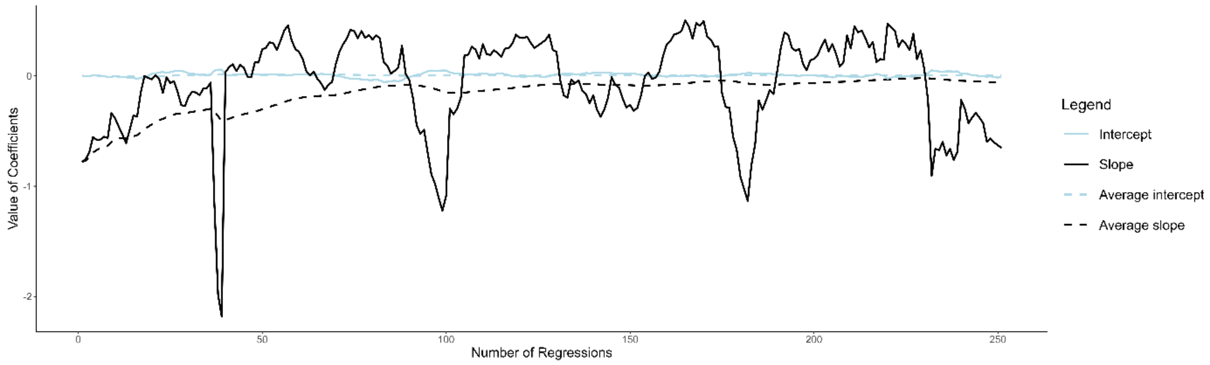


Figure 2: Galton Coefficients for the Variance

This figure shows the estimated intercept, the estimated slope, the expanding average for intercept and slope for all OLS regression of ex-post variance on the historical estimate.

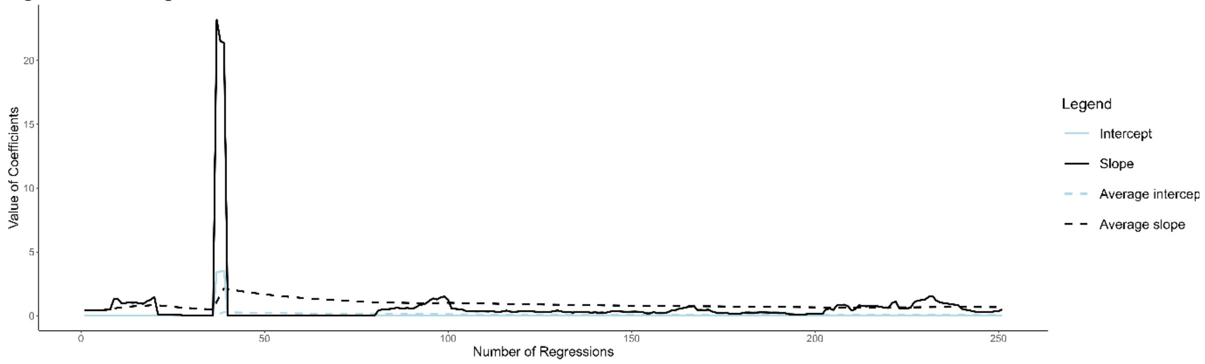


Figure 3: Galton Coefficients for the Standard Deviation

This figure shows the estimated intercept, the estimated slope, the expanding average for intercept and slopes for all OLS regression of ex-post standard deviation on the historical estimate.

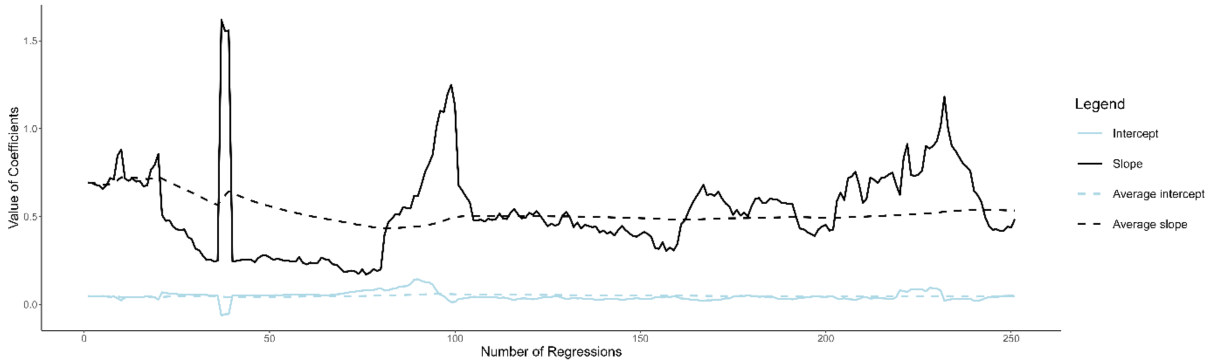


Figure 4: Galton Coefficients for the Correlation

This figure shows the estimated intercept, the estimated slope, the expanding average for intercept and slopes for all OLS regression of ex-post correlation on the historical estimate.

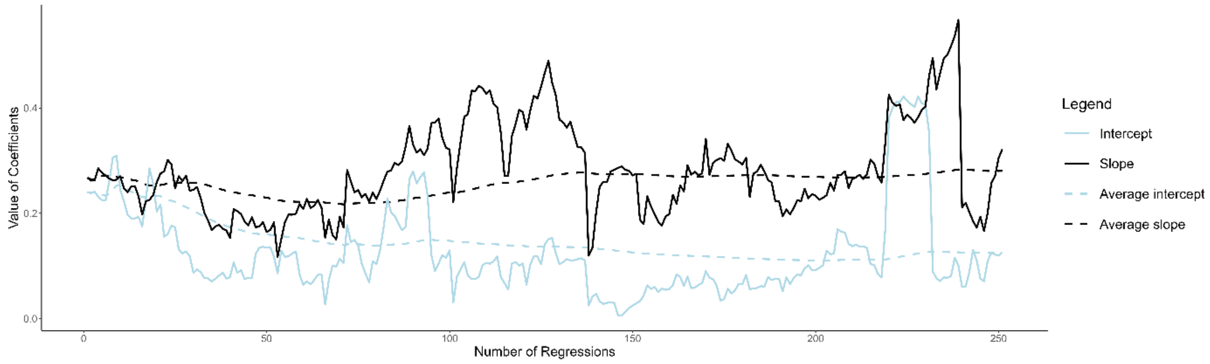
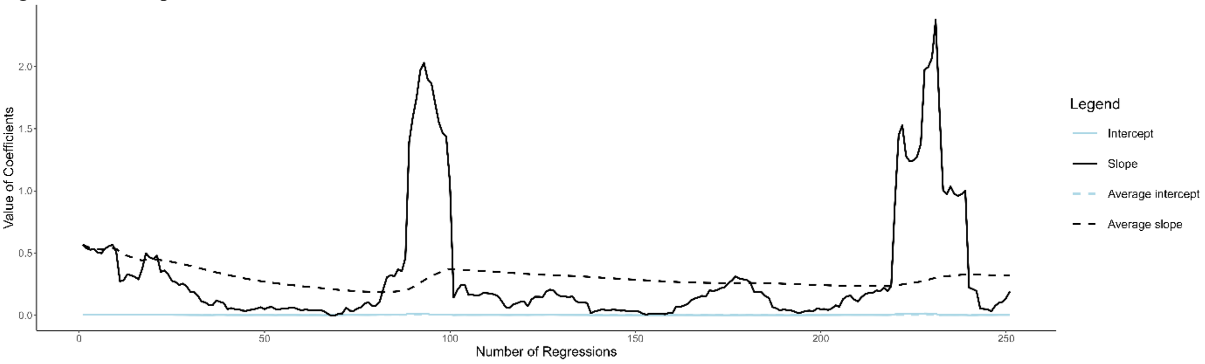


Figure 5: Galton Coefficients for the Covariance

This figure shows the estimated intercept, the estimated slope, the expanding average for intercept and slopes for all OLS regression of ex-post covariance on the historical estimate.



3.3 The portfolios involved in the horse race

In this work, I analyze the out-of-sample performance of eight different portfolios for the 10, 30 and 50 companies with the largest market capitalization at the beginning of each year. In addition to the simple Mean Variance and Global Minimum Variance portfolios, I construct the naïve 1/N and a value-weighted portfolio as benchmark. Furthermore, I create the Galton Mean Variance (Equation 20) and Galton Global Minimum Variance portfolio (Equation 21) according to the following portfolios allocation. The G as superscript suggests that these variables are estimated using the Galton strategy.

$$\omega_t^{Galton\ MV} = \frac{\Sigma_t^{G-1}\mu^G}{|\mathbf{1}\Sigma_t^{G-1}\mu^G|} \tag{20}$$

$$\omega_t^{Galton\ GMV} = \frac{\Sigma_t^{G-1}\mathbf{1}}{|\mathbf{1}'\Sigma_t^{G-1}\mathbf{1}|} \tag{21}$$

I apply also the "Empirical Bayes" version for the Galton strategy which involves constraining the expanding average of slope coefficients estimated through the Fama-MacBeth procedure. Specifically, the new slope coefficient ($\bar{g}^{G^*}_{X,t}$) utilized in the optimization process is obtained using the following equation:

$$\bar{g}^{G^*}_{X,t} = \max\left(\min\left(\bar{g}^{G}_{X,t}, 1\right), 0\right) \tag{22}$$

Through equation 22, any slope coefficient above 1 is reduced to the unity and each negative variable is adjusted to 0. I find that the expanding slope coefficient of mean is always below 0 during the horse race while the standard deviation and correlation vary between 0 and 1 as it is possible to notice in Figures 1,3 and 4. As a result, the slope coefficient for the mean is shrunk toward 0 and the Empirical Bayes allocation perfectly coincides with the Galton GMV portfolio. This result is possible because looking at the data, there is mean reversion in the United Kingdom stock market. Regarding the LW allocation, it is based on the shrinkage of the variance-covariance matrix towards a constant correlation target (Σ_t^*) according to Ledoit and Wolf (2003) and on the James-Stein shrinkage for the mean vector. The resulting vector of weights is given by:

$$\omega_t^{LW} = \frac{\Sigma_t^{*-1}\mu^{JS}}{|\mathbf{1}\Sigma_t^{*-1}\mu^{JS}|} \tag{23}$$

Performance

4.1 Risk-adjusted Performance

When comparing the performance of alternative portfolios, it is not sufficient relying solely on their realized returns, as it does not take into account the amount of risk taken to achieve these returns. To avoid this issue, it is essential to use risk-adjusted measure to compare portfolios with different level of risk. The analysis of the performance of the portfolios starts with the computation of the Sharpe Ratio (SR) which is an extensively adopted metric to evaluate risk-adjusted returns.

This indicator is expressed by the ensuing formula:

$$SR_i = \frac{E(R_i)}{\sigma_{R_i}} \tag{24}$$

Where $E(R_i)$ denotes the average of the excess return of the portfolio i and σ_{R_i} depict the standard deviation of its excess return. I compute the Sharpe ratio in this way because the risk-free rate changes each month. Moreover, I calculate the relative t-statistic of the Sharpe ratio to test whether it is statistically different from 0, as shown in Table 2. To derive the standard errors for SR (Lo (2002)), I employ the following expression:

$$Standard\ Errors\ (SR_i) = \frac{\sqrt{1 + SR_i^2/2}}{\sqrt{T}} \tag{25}$$

Where T denotes the number of months of the out-of-sample horse race and SR_i the out-of-sample Sharpe ratio of strategy i.

Table 2: Annualized Sharpe ratio

This table illustrates the annualized Sharpe ratio in decimal points for each portfolio across various investment universe. In parenthesis, the t-statistic is used to assess if the Sharpe ratio is statistically different from 0. A * and ** indicate that the t-statistic is statistically significant at 90% and 95% confidence interval. The dataset spans from February 1996 to December 2022, with the first decade which is employed for the construction of robust coefficients. The horse race commences from February 2011, considering also the initial historical window of 60 months. At the beginning of each year, the top 10, 30 and 50 companies in terms of market capitalization are selected to form the portfolios, which are rebalanced monthly.

	N=10	N=30	N=50
Galton MV	0.62	0.56	0.52
	(2.12)**	(1.94)*	(1.80)*
MV	0.28	0.43	0.20
	(0.97)	(1.48)	(0.70)
Galton GMV	0.60	0.60	0.62
	(2.05)**	(2.05)**	(2.13)**
GMV	0.56	0.31	0.42
	(1.90)*	(1.06)	(1.44)
Empirical Bayes	0.60	0.60	0.62
	(2.05)**	(2.05)**	(2.13)**
Value Weighted	0.43	0.41	0.44
	(1.46)	(1.42)	(1.50)
1/N	0.53	0.45	0.48
	(1.81)*	(1.55)	(1.64)*
Ledoit Wolf	0.41	0.37	0.53
	(1.42)	(1.28)	(1.82)*

The Galton GMV allocation, with an average SR of 0.61, delivers the highest annualized Sharpe ratio in two out of the three different sizes of the investment universe. It offers an enhanced performance of 42% in terms of SR as opposed to the simple Global Minimum Variance portfolio. It is noteworthy that the Galton MV, with an average SR of 0.57, strictly outperforms in all the cases the naïve 1/N (average SR of 0.49) and the value-weighted allocations (average SR of 0.42). Additionally, the performance of the Ledoit and Wolf allocation is quite remarkable, as it beats the benchmark when the number of stocks involved in the optimization process is greater than 50.

Although all the Sharpe ratios are positive, it should be noted that only the Galton portfolios exhibit a statistically significant Sharpe Ratio with a 90% or 95% confidence level for all the sizes of the investment universe. The lack of statistically significance of SR at 99% confidence level may be attributed to the relatively short length of the horse race, which spans only 143 months.

Furthermore, to gain additional insights, I perform the Jobson and Korkie (1981) test in Table 3. Following the correction outlined in Memmel (2003), this test aims to investigate whether the Sharpe ratios of two different portfolios, with one being the benchmark (i.e, the value-weighted allocation), are statistically distinguishable. The equation for the Jobson and Korkie t-statistic is provided below:

$$Z_{JB} = \frac{(\sigma_{ben}\mu_i - \sigma_i\mu_{ben})}{\sqrt{\eta}} \quad (26)$$

$$\eta = \frac{2\sigma_i^2\sigma_{ben}^2 - 2\sigma_i\sigma_{ben}\sigma_{i,ben} + \frac{1}{2}\mu_i^2\sigma_{ben}^2 + \frac{1}{2}\mu_{ben}^2\sigma_i^2 - \frac{\mu_i\mu_{ben}\sigma_{i,ben}^2}{\sigma_i\sigma_{ben}}}{T} \quad (27)$$

In this equation, T represents the out-of-sample time period, μ_{ben} and σ_{ben} indicates respectively the average and the volatility of the excess returns of the benchmark while $\sigma_{i,ben}$ the covariance between the excess returns of a specific portfolio and the reference portfolio.

Table 3: Jobson and Korkie test

This table showcases the annualized difference of the Sharpe ratio for each portfolio compared to the value-weighted allocation. The Jobson and Korkie (1981) t-statistic after the Memmel (2003) correction is presented in parentheses. A * and ** reveal that the t-statistic is statistically significant at 90% and 95% confidence interval.

	N=10	N=30	N=50
Galton MV	0.19	0.15	0.09
	(2.42)**	(1.13)	(0.58)
MV	-0.14	0.02	-0.23
	(-0.39)	(0.04)	(-0.55)
Galton GMV	0.17	0.19	0.18
	(1.89)*	(1.40)	(1.16)
GMV	0.13	-0.11	-0.02
	(0.60)	(-0.40)	(-0.06)
Empirical Bayes	0.17	0.19	0.18
	(1.89)*	(1.40)	(1.16)
Ledoit Wolf	-0.01	-0.04	0.10
	(-0.07)	(-0.14)	(0.34)

Out of the 9 Galton portfolios, none exhibits a negative annualized difference of the Sharpe ratio when compared to the benchmark. This interesting result indicates that the Galton's approach to portfolio construction leads to optimal performance in terms of risk-adjusted returns. On the other hand, a simple Mean Variance investor underperforms the benchmark in two out of three cases, probably due to estimation errors. Not surprisingly, only the Galton portfolios yield a statistically significant Jobson and Korkie T-statistic at the 90% and 95% confidence levels when the number of stocks is equal to 10, providing further evidence of their superiority over the other investment strategies.

In tables 4 and 5, I present the first and second moment of the distribution of excess returns for each investment strategy. It is worthy of attention that there is a considerable variability of the mean of excess returns across different universe sizes for the simple MV. I believe that this outcome may be ascribed to the usage of historical mean vector, which might contribute to the weaker performance compared to other methods. The simple GMV approach, instead, exhibits the second highest average standard deviation even if it is based only on the historical variance-covariance matrix. Historical estimators exclusively rely upon only the information available in the last historical window of 60 months, forgetting what has occurred in the more distant past. In contrast, the Galton approach takes into consideration the entire history of returns, enabling a more robust assessment of risk and return characteristics.

Table 4: First moment of distribution

This table shows the annualized mean of excess returns expressed in percentage points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	8.14	7.17	6.60
MV	8.04	45.39	130.11
Galton GMV	7.79	7.37	7.66
GMV	8.30	5.56	12.30
Empirical Bayes	7.79	7.37	7.66
Value Weighted	5.70	5.33	5.73
1/N	7.24	6.18	7.02
Ledoit Wolf	6.12	5.90	8.21

Table 5: Second moment of distribution

This table shows the annualized standard deviation of excess returns expressed in percentage points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	13.16	12.70	12.59
MV	28.62	105.61	641.16
Galton GMV	13.03	12.29	12.35
GMV	14.96	18.16	29.46
Empirical Bayes	13.03	12.29	12.35
Value Weighted	13.39	12.95	13.14
1/N	13.75	13.67	14.69
Ledoit Wolf	14.86	15.91	15.45

In finance, measuring risk has always been a critical issue since the standard deviation, the most common proxy of risk, captures the amount of dispersion of both positive and negative returns. Roy (1952) introduces the concept of downside risk, which measures the standard deviation of returns below a certain threshold, defined as the minimum acceptable return (MAR). In this work, since the risk-free rate is changing over time, I consider only the excess return that are not negative (thus the MAR in this case is set equal to 0). The downside risk, focusing on the threat of losses, can be considered as a more appropriate measure of risk. The Ang et al. (2002) study demonstrates that investors demand a higher risk premium for assets with greater downside risk because they are perceived as more uncertain and subject to greater losses. Sortino and Van der Meer (1991) propose the Sortino ratio (Table 6), which is a risk-adjusted performance measure which takes into account the downside risk instead of the standard deviation. The expression for the Sortino ratio is the following:

$$Sortino\ ratio = \frac{E(R_i)}{\sigma_{d,i}} \tag{28}$$

Let $E(R_i)$ be the average of excess return of the strategy i and $\sigma_{d,i}$ is the standard deviation of the negative excess returns.

Table 6: Annualized Sortino ratio

This table shows the annualized Sortino ratio expressed in decimal points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	1.04	0.86	0.77
MV	0.43	0.89	0.55
Galton GMV	1.00	0.92	0.92
GMV	0.89	0.45	0.65
Empirical Bayes	1.00	0.92	0.92
Value Weighted	0.66	0.61	0.64
1/N	0.85	0.68	0.70
Ledoit Wolf	0.64	0.52	0.76

Table 7: Annualized downside risk

This table shows the annualized downside risk of excess returns expressed in percentage points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	7.86	8.34	8.56
MV	18.62	50.96	235.14
Galton GMV	7.77	7.99	8.37
GMV	9.28	12.41	19.05
Empirical Bayes	7.77	7.99	8.37
Value Weighted	8.57	8.71	8.93
1/N	8.51	9.12	9.98
Ledoit Wolf	9.60	11.37	10.76

Considering the Sortino ratio as the main indicator to evaluate risk-adjusted performance, it is remarkable that the Galton GMV and Empirical Bayes strategies boast the highest average, reaching a score of 0.95. Even after the substantial reduction in risk achieved using the downside risk instead of the standard deviation, the simple MV outperforms the Galton MV allocation only once. Nevertheless, the MV strategies still carries the highest level of risk which is typically undesirable for a risk-averse investor.

In gauging the performance of the different investment allocations, I also utilize the Morningstar Risk-Adjusted Return (MRAR) metric. This measure is derived assuming that investors have a Constant Relative Risk Aversion (CRRA) utility function, meaning that the relative risk aversion coefficient of investors is the same regardless their wealth. The utility functions, such as the power utility, present a concave behavior implying that the satisfaction of making money is smaller than the pain of losing the same amount. The MRAR can be viewed as the risk-free equivalent excess return of the portfolio based on a given risk aversion coefficient (γ) (Bodie and Kane (2020)). In this exercise, I adopt a risk aversion coefficient of 2 but it is possible to use different levels of risk aversion recalling that a higher coefficient implies a greater aversion to risk. The following equation is used in Table 8 for the computation of MRAR based on monthly observations:

$$MRAR(\gamma) = \left[\frac{1}{T}\sum_{t=1}^{T}(1 + r_t)^{-\gamma}\right]^{\frac{12}{\gamma}} - 1 \tag{29}$$

Where γ is the coefficient of Risk aversion and r_t the excess returns of a specific investment strategy at time t. In addition, I test whether the MRAR of each investment strategy differs significantly from that of the value-weighted allocation using the bootstrap method. In each simulation, I create random samples with replacement starting from the vector of returns of a specific investment strategy and the benchmark. Then, I compute the MRAR difference which I store in a separate vector. I repeat this process 1000 times and using the standard deviation of the MRAR differences as standard error, I calculate the t-statistic.

Table 8: Morningstar Risk-Adjusted Return and bootstrap

This table displays the MRAR expressed in percentage points for each different investment strategy using monthly observations. The coefficient of risk aversion of the investors is set equal to 2. In parenthesis, I show the t-statistic of the test to assess if the MRAR of a specific strategy is different from that of the value-weighted allocation. A *, ** and *** indicate respectively that the T-statistic is significant at 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	5.73	4.77	4.18
	(2.37)**	(1.03)	(0.52)
MV	-4.13	-100.00	-83.63
	(-0.84)	(-3.28)***	(-4.53)***
Galton GMV	5.42	5.15	5.38
	(1.82)*	(1.26)	(1.07)
GMV	5.08	0.47	-1.20
	(0.62)	(-0.53)	(-0.53)
Empirical Bayes	5.42	5.15	5.38
	(1.82)*	(1.26)	(1.07)
Value Weighted	3.07	2.81	3.11
1/N	4.52	3.36	3.68
Ledoit Wolf	2.84	1.88	4.48
	(-0.07)	(-0.19)	(0.31)

In a scenario where investors accept a positive risk-free rate as certainty equivalent, the Galton allocations consistently outperform the benchmark and the other portfolios, yielding a greater MRAR. Furthermore, in their simple version, MV and GMV allocations generate negative MRAR in 4 out of 6 cases. Specifically, MV allocation yields the lowest MRAR due to the punishment for its excessive variability of returns.

4.2 Risk analysis

The maximum drawdown is a relevant measure of risk which refers to the largest reduction in terms of wealth that a certain investment strategy experiences at its highest point of performance. To fully grasp the implications of the maximum drawdown, it is essential to understand that this metric only focuses on the proportional magnitude of the loss and does not consider the frequency of losses or the time taken to recover from it. As can be noticed in Table 9, another important advantage of implementing the Galton strategy consists in getting a lower maximum drawdown in absolute value respect to their simple version.

Furthermore, the Galton GMV and the Empirical Bayes have always a lower drawdown in absolute value for all the different size of portfolios with respect to all the allocations. In addition, the value-weighted and 1/N portfolios seem very stable allocations as in contrast to the simple MV allocation which turns out to be again the riskiest strategy, going bankrupt when the investment universe is composed by 30 and 50 stocks. Looking at this metric, the Ledoit and Wolf portfolio delivers a slightly poorer performance compared to the benchmark.

Table 9: Maximum drawdown

This table shows the maximum drawdown expressed in decimal points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	-0.28	-0.27	-0.28
MV	-0.53	-1.00	<-1
Galton GMV	-0.27	-0.26	-0.27
GMV	-0.29	-0.37	-0.50
Empirical Bayes	-0.27	-0.26	-0.27
Value Weighted	-0.31	-0.27	-0.27
1/N	-0.29	-0.27	-0.32
Ledoit Wolf	-0.31	-0.42	-0.37

Volatility forecasting has always been another key topic in Finance because it is the main proxy for risk and it is not constant over time. Looking at real data, the volatility shows some clusters pattern (first noted by Mandelbrot (1963)) where large changes of the second moment of distribution of returns are followed by similarly large changes, of either sign, and small changes tend to be followed by small variations. In 1982, the Nobel prize Robert Engle invents the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to better capture the time-varying structure of volatility. To test if the investment strategies are good in forecasting volatility, I use the following equation to compute in each month an expectation for the volatility:

$$E(\sigma) = \sqrt{\omega' \Sigma \omega} \tag{30}$$

I store all these expectations in a vector, subsequently I compute the mean obtaining an average ex-ante measure of volatility. Then I calculate the sample standard deviation of realized returns of each allocation and collect the ratio. In this test, I do not include the 1/N and the value-weighted portfolios since they do not use any measure of volatility to form weights. In table 10, I show the ratio of realized volatility over the ex-ante expectation.

The Galton strategy shows good properties in terms of risk forecast because the ratio of realized volatility over the ex-ante expectation is lower than 1 meaning that an investor is overestimating the future volatility. As it is possible to notice, the Ledoit and Wolf allocation shows some optimism with the ex-post volatility being always higher than the ex-ante measure. Table 10 displays an interesting pattern, the level of optimism for forecasting the volatility is increasing in the number of companies that belong to the investment universe.

Table 10: Risk forecast ability

This table shows the ratio of realized volatility over the expected volatility expressed in decimal points for each different investment strategy.

	N=10	N=30	N=50
Galton MV	0.85	0.94	0.96
MV	1.34	3.18	23.30
Galton GMV	0.85	0.93	0.98
GMV	1.31	2.51	6.83
Empirical Bayes	0.85	0.93	0.98
Ledoit Wolf	1.19	1.59	1.72

Value at Risk (VaR) is a widely accepted metric for quantifying the maximum potential loss of a portfolio over a specific time horizon with a particular confidence level. Its application as one of the main metrics for measuring market risk in trading portfolio can be traced back to J.P. Morgan's Risk Metrics paper in 1996. To construct a measure of Var at time t, I employ the following equation:

$$VaR = -\sigma * z \tag{31}$$

Where σ is the expected volatility obtained as $\sqrt{\omega'\Sigma\omega}$ and z is the inverse of a normal standard distribution associated to the selected confidence level. For this study, I choose the same confidence levels (95% and 99%) as Barroso and Saxena (2021). After computing the VaR measure for each month, I create a vector whose elements take value 1 if the return of the subsequent month is lower than the VaR computed at time t, and 0 otherwise. Then, I sum all the elements of this vector and divide by its length to determine the number of times in which a given strategy produced returns lower than VaR out of 100 times. Overall, VaR provides a valuable tool for measuring risk for financial applications, but it is important to acknowledge its limitations, such as its reliance on the normal distribution to compute the scaling factor z and its inability to account the losses in the tail of distribution. To provide a more complete picture of the potential risks and reward of a given investment strategy, I conduct a similar exercise also for the right tail of the distribution. To do this, I use the same formula ($Var = -\sigma * z$) but with the opposite sign then I follow the same steps. Furthermore, I test if the Var hit rate of each allocation is statistically different from 0.05 and 0.01. In doing so, I compute the standard error of a proportion (Newbold (2013)) using the following equation:

$$Standard\ Error\ (VaR) = \sqrt{\frac{pq}{T}} \tag{32}$$

Where p indicates the observed VaR hit rate in the sample, q is defined as $1 - p$ and T is the length of the sample.

Table 11: Left VaR hit rate at 95% confidence level

This table shows the VaR hit rate for the left tail of distribution at 95% confidence level in percentage points for each investment strategy. I assume that returns follow a normal distribution. I show the t-statistic in parenthesis, a *, ** and *** means respectively that the VaR hit rate is statistically different from 0.05 with a 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	2.11	2.82	2.11
	(-2.40)**	(-1.58)	(-2.40)**
MV	9.86	23.24	35.21
	(1.95)*	(5.16)***	(7.56)***
Galton GMV	2.11	2.11	3.52
	(-2.40)**	(-2.40)**	(-0.96)
GMV	7.75	17.61	34.51
	(1.23)	(3.96)***	(7.42)***
Empirical Bayes	2.11	2.11	3.52
	(-2.40)**	(-2.40)**	(-0.96)
Ledoit Wolf	6.34	8.45	11.97
	(0.66)	(1.48)	(2.57)**

Table 12: Right VaR hit rate at 95% confidence level

Table 12 shows the VaR hit rate for the right tail of distribution at 95% confidence level in percentage points for each investment strategy. I assume that returns follow a normal distribution. I show the t-statistic in parenthesis, a *, ** and *** means respectively that the VaR hit rate is statistically different from 0.05 with a 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	4.93	4.93	4.23
	(-0.04)	(-0.04)	(-0.46)
MV	11.27	23.94	40.14
	(2.37)**	(5.31)***	(8.57)***
Galton GMV	4.23	4.93	5.63
	(-0.46)	(-0.04)	(0.33)
GMV	14.08	25.35	45.07
	(3.12)***	(5.59)***	(9.63)***
Empirical Bayes	4.23	4.93	5.63
	(-0.46)	(-0.04)	(0.33)
Ledoit Wolf	9.86	17.61	19.01
	(1.95)*	(3.96)***	(4.27)***

An allocation is successful in estimating risk if the proportion of times in which it exceeds the VaR at 95% confidence level is close to theoretical value of 5%. Galton allocations seem to be too conservative indeed Galton MV is the only portfolio which delivers returns that do not exceed the VaR threshold more than 5 times out of 100 for both tails of distribution. For the Galton GMV strategy, only one out of six cases have a hit rate above 5% indicating a smaller but higher likelihood of returns above or below the VaR. Moreover, simple MV and GMV strategies show the potential to generate extreme returns on both sides of the distribution, but it seems unrealistic the idea of some investors able to manage this high variability in their investment outcome. Furthermore, the Ledoit Wolf allocation records a lower performance in this Risk Management test showing some instability in the tails of distribution, with 4 out of 6 hit rates statistically different from the theoretical proportion. As for the ratio of realized volatility over the ex-ante measure, the VaR hit rates are an increasing function of the number of companies involved in the portfolio optimization. A smaller number of companies would lead to less extreme performances.

In tables 13 and 14, I display the results of the VaR analysis at 99% confidence level for the left and right tails of the distribution.

Table 13: Left VaR hit rate at 99% confidence level

This table shows the VaR hit rate for the left tail of distribution at 99% confidence level in percentage points for each investment strategy. I assume that returns follow a normal distribution. I show the t-statistic in parenthesis, a *, ** and *** means respectively that the VaR hit rate is statistically different from 0.01 with a 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	0.70	1.41	2.11
	(-0.42)	(0.41)	(0.93)
MV	4.23	14.79	30.28
	(1.92)*	(4.64)***	(7.62)***
Galton GMV	0.70	1.41	1.41
	(-0.42)	(0.41)	(0.41)
GMV	3.52	12.68	29.58
	(1.64)*	(4.20)***	(7.49)***
Empirical Bayes	0.70	1.41	1.41
	(-0.42)	(0.41)	(0.41)
Ledoit Wolf	2.82	4.23	4.23
	(1.31)	(1.92)*	(1.92)*

Table 14: Right VaR hit rate at 99% confidence level

This table shows the VaR hit rate for the right tail of distribution at 99% confidence level in percentage points for each investment strategy. I assume that returns follow a normal distribution. I show the t-statistic in parenthesis, a *, ** and *** means respectively that the VaR hit rate is statistically different from 0.01 with a 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	1.41	2.11	1.41
	(0.41)	(0.93)	(0.41)
MV	5.63	17.61	36.62
	(2.40)**	(5.21)***	(8.84)***
Galton GMV	2.11	1.41	0.70
	(0.93)	(0.41)	(-0.42)
GMV	4.93	15.49	41.55
	(2.17)**	(4.79)***	(9.84)***
Empirical Bayes	2.11	1.41	0.70
	(0.93)	(0.41)	(-0.42)
Ledoit Wolf	3.52	6.34	9.15
	(1.64)*	(2.62)***	(3.38)***

While the results of the Value at Risk analysis at the 99% confidence level show some changes in the outcome, Galton allocations still demonstrate the lowest number of hit rates meaning that they are the safest allocations under a Risk Management view.

Mean Variance optimizers only care about the first two moment of distributions neglecting higher orders moments such as skewness and kurtosis, however these characteristics play an important role in defining the performance of an investment strategy. In table 15, I present the skewness of each allocation. Most of the portfolios exhibit a negative skewness meaning that the probability of observing negative returns is potentially higher than what is estimated by using a symmetrical distribution. This feature is more apparent when analyzing the VaR hit rates at 99% confidence level. Another important aspect to consider is the ratio of VaR hit rate in the left tail of 99% confidence level over the 95% one. Interestingly, the MV strategy is the riskiest strategy with more than 64% of occurrences of losses exceeding both the VaR at 95% and 99% confidence levels. The corresponding ratios for the Galton MV and GMV allocations are 61.11% and 46.67% respectively. In 43% of the cases, a loss exceeding the VaR at 95% is higher than the VaR at 99% confidence level for the Ledoit and Wolf strategy. The theoretical proportion of losses exceeding both the VaR at 95% and 99% confidence level should be 20% meaning that the tails of distribution of these portfolios are quite fat.

Table 15: Third moment of distribution

This table shows the skewness of excess returns in decimal points for each investment strategy.

	N=10	N=30	N=50
Galton MV	0.18	-0.63	-0.96
MV	0.08	2.82	8.20
Galton GMV	0.19	-0.67	-1.08
GMV	-0.04	-0.36	-0.06
Empirical Bayes	0.19	-0.67	-1.08
Value Weighted	0.02	-0.38	-0.52
1/N	0.13	-0.32	-0.57
Ledoit Wolf	-0.11	-0.93	-1.00

Additionally, in assessing the risk of extreme values, I compute the fourth moment of distribution of returns (Table 16) for each investment scheme. The outcomes of this exercise confirm that most of the unconditional distribution of returns is leptokurtic (Cont 2001), indicating a higher probability of observing extreme values compared to what would be estimated with a normal distribution. The positive skewness displayed by the simple MV allocation is completely offset by its huge kurtosis. Furthermore, the investment strategies with higher VaR hit rates exhibit higher kurtosis. Finally, most of the allocations show a greater negative skewness and higher kurtosis increasing the number of stocks involved in the portfolio optimization. I believe that a possible cause of such result can be the decreasing benefits of diversification when the dimension of the portfolio rises. However, it is also crucial to consider that regardless of the level of diversification, the market exhibits negative skewness and excess kurtosis. Consequently, investors are inevitably exposed to the potential impact of tail risks driven by market dynamics when holding a portfolio of equities.

Table 16: The fourth moment of distribution

This table shows the kurtosis of excess returns in decimal points of each investment strategy.

	N=10	N=30	N=50
Galton MV	4.28	6.91	7.81
MV	3.01	20.77	92.84
Galton GMV	4.07	6.60	7.96
GMV	3.04	4.68	5.00
Empirical Bayes	4.07	6.60	7.96
Value Weighted	4.08	5.60	6.76
1/N	4.32	7.67	10.45
Ledoit Wolf	3.59	6.22	6.51

4.3 500 Random Horse races

Considering the scarce size of out-of-sample observations, I run 500 random horse races to test if my results are robust or they are dependent on the particular sample of stocks. This type of test is present in Barroso and Saxena (2021). At the beginning of each year, I randomly select 50 companies among the 100 with the highest market capitalization and I perform the optimization process for all the investment strategies (Table 17). This exercise gives the opportunity of computing 71,500 (143*500) OOS returns for each investment strategy with a dynamic sample of stocks that may belong to the portfolio. Furthermore, I test if each investment strategy delivers at least the same Sharpe ratio of Galton allocations in at least half of the races.

The t-statistic is computed using the following equation:

$$t - statistic = \frac{Sharpe \geq Galton - 0.5}{\sqrt{0.5^2/500}} \tag{33}$$

Where Sharpe ≥ Galton represent the proportion of times in which each investment strategy produces an annualized Sharpe ratio greater than the Galton.

Table 17: 500 random horse races

This table shows the outcome of 500 random horse races. At the beginning of each year, I randomly select 50 companies among the 100 with the highest market capitalization. The first column represents the average annualized Sharpe ratio. The second column indicates the percentage of times in which an investment strategy delivers an annualized Sharpe ratio greater than the one offered by Galton GMV or Empirical Bayes. The third column shows the percentage of times in which an investment strategy delivers an annualized Sharpe ratio greater than the one offered by Galton MV. The fourth column indicates the average ratio between the realized volatility and ex-ante expectation. In parenthesis, I show the t-statistic where the null hypothesis is that investment strategy delivers at least the same Sharpe ratio of Galton portfolios in at least half of the races. A *** indicates that the proportion is statistically different from 0.5 at 99% confidence level.

	Sharpe	Sharpe ≥ Galton GMV	Sharpe ≥ Galton MV	Ratio
Galton MV	0.51	0.01	1.00	1.00
		(-21.91)***		
MV	0.00	0.02	0.05	23.43
		(-21.65)***	(-20.30)***	
Galton GMV	0.64	1.00	0.99	0.99
			(21.91)***	
GMV	0.20	0.03	0.10	7.11
		(-21.11)***	(-17.89)***	
Empirical Bayes	0.64	1.00	0.99	0.99
			(21.91)***	
Value-Weighted	0.44	0.08	0.31	
		(-18.78)***	(-8.68)***	
1/N	0.46	0.01	0.29	
		(-21.73)***	(-9.57)***	
Ledoit Wolf	0.55	0.26	0.61	1.68
		(-10.64)***	(4.74)***	

The annualized Sharpe ratio of the simple MV and GMV allocations differs significantly from the values presented in Table 2, in contrast to the other portfolios. Among all the investment strategies, the Galton GMV still holds the highest average Sharpe ratio (0.64), followed by Ledoit Wolf (0.55) and the Galton MV (0.51). Not surprisingly, the 1/N and the value-weighted allocations outperform the Galton GMV in less than 8% of cases (31% when compared to the Sharpe ratio of Galton MV). No strategy produces an annualized Sharpe ratio higher than the one offered by Galton GMV in at least half of horse races and this result is statistically significant at 99% confidence level. However, the Ledoit Wolf strategy outperforms the Galton MV in 61% of cases demonstrating a consistently stronger performance during the time span analyzed.

In terms of ratio between volatility and ex-ante expectation, Galton allocations show a ratio closer to 1, indicating a more accurate volatility predictions while the Ledoit Wolf strategy confirms the presence of some optimism in forecasting future volatility. This exercise confirms that Galton portfolios easily beat the benchmark and slightly struggle in outperforming the Ledoit Wolf allocation in terms of Sharpe ratio. As regard forecasting volatility, the Galton portfolios are still the best in class.

4.4 The Alpha estimation

Markowitz (1959) has the merit of dividing the risk of a portfolio along two directions: the idiosyncratic and the systematic component. The former refers to the specific risk of individual companies which can be mitigated through diversification by investing in stocks with low correlation. For this reason, scholars expect that the market should not give any reward to investors to bear this type of risk. In contrast, the latter stems from an exposure to common macroeconomic factors, potentially harming all the assets and it cannot be diversified. If a portfolio generates a performance which exceeds what would be predicted considering its level of systematic risk, it has a positive alpha. There are several factor models that attempt to predict the ex-ante performance of a portfolio by examining its exposure to some risk-factors.

Sharpe (1964) develops the first single factor model: the Capital Asset Pricing Model (CAPM) which considers as the main source of systematic risk, the excess return of the market portfolio which is regarded the most possible diversified allocation. I use as proxy of the market portfolio the FTSE all share index. I avoid any issue of survivorship bias by including in the dataset also the companies that were delisted from the FTSE All share index. To get an estimate of the alpha by using CAPM, I run the following regression:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \theta_{i,t} \quad (34)$$

Where $R_{i,t}$ represents the excess return of strategy i at time t, $R_{m,t}$ denotes the excess return of the market portfolio, β_i measures the sensitivity of the returns of portfolio i with the market portfolio and α_i corresponds to the alpha of the allocation i. To test whether the estimated alphas are statistically different from 0, I employ White's standards errors (1980) which account for the presence of heteroskedasticity. Table 18 displays the annualized alpha for each investment strategy. Only the 1/N and the value-weighted allocations record negative alphas. While Galton portfolios exhibit positive alpha, they are not statistically significant, likely due to the shorter out-of-sample time horizon. Despite the fact that simple MV portfolio has huge alphas, none of these are statistically significant due to high standard errors.

Table 18: Alpha estimation with CAPM and FTSE All share index

This table showcases the annualized alpha in percentage points for each investment allocation assuming that the Capital Asset Pricing Model holds. I show the White t-statistic in parenthesis.

	N=10	N=30	N=50
Galton MV	2.62	1.60	1.24
	(1.24)	(0.89)	(0.63)
MV	6.51	45.50	157.08
	(0.77)	(1.45)	(0.78)
Galton GMV	2.38	1.91	2.38
	(1.12)	(1.18)	(1.27)
GMV	3.57	0.70	6.97
	(1.06)	(0.16)	(0.85)
Empirical Bayes	2.38	1.91	2.38
	(1.12)	(1.18)	(1.27)
Value Weighted	-0.12	-0.85	-0.66
	(-0.06)	(-0.80)	(-0.88)
1/N	1.30	-0.43	-0.11
	(0.62)	(-0.51)	(-0.15)
Ledoit Wolf	1.06	1.90	4.01
	(0.35)	(0.54)	(1.08)

Since the value-weighted allocation has a negative alpha, I propose to regress the excess returns of each investment strategy on the excess returns of the value-weighted allocation. In this way, I check if the low level of alpha gained by the other investment strategies is due to the overall performance of the largest companies sorted each year or these allocations in this time span do not produce excess return higher than the ones predicted by the CAPM. In Table 19, I perform this analysis.

Table 19: Alpha estimation with CAPM and value-weighted allocation

This table shows the annualized alpha in percentage points obtained by regressing the excess return of each investment strategy on the excess return of value-weighted allocation. I display the White t-statistic in parenthesis. A * and ** means respectively that the alpha is statistically different from 0 with a 90% and 95% confidence level.

	N=10	N=30	N=50
Galton MV	2.74	2.49	1.87
	(2.57)**	(1.38)	(0.95)
MV	5.95	45.13	151.76
	(0.73)	(1.46)	(0.75)
Galton GMV	2.52	2.83	3.08
	(2.10)**	(1.67)*	(1.55)
GMV	3.67	1.24	7.35
	(1.22)	(0.29)	(0.92)
Empirical Bayes	2.52	2.83	3.08
	(2.10)**	(1.67)*	(1.55)
Ledoit Wolf	1.37	2.63	4.55
	(0.48)	(0.72)	(1.21)

As it is possible to appreciate from Table 19, nearly all the allocations have increased their annualized alpha no matter the size of the portfolio. This finding underscores that focusing solely on the 50 largest companies, rather than the whole market, these allocations produce higher risk-adjusted returns than what is predicted by CAPM. However, only the Galton portfolios record statistically different from 0 alphas when the investment universe is composed by 10 and 30 stocks, albeit MV allocation has still the highest alpha.

When estimating the alpha according to the Capital Asset Pricing Model, I also consider Roll's critique in 1977 which was mainly based on the impossibility in observing the true market portfolio because it should include all the conceivable assets and on the possibility that there are more factors which may affect the systematic risk of a portfolio. To partially address this issue and avoid model misspecification, I also use the 3 Factor models elaborated by Fama and French (1992) to estimate alphas. The researchers add the value and size factors to the market risk because empirical data suggests that, on average, small companies in terms of market capitalization exhibit higher returns than bigger firms. Similarly, distressed companies (high book-to-market ratio) manifest large returns than financially health firms (low book-to market ratio). Since this higher performance may be addressed to a higher exposure to systematic risk, I run the following regression to obtain an estimate of alpha using the 3 Factor Models:

$$R_{i,t} = \alpha_i + \beta_{i,m} R_{m,t} + \beta_{i,SMB} SMB_t + \beta_{i,HML} HML_t + \vartheta_{i,t} \tag{35}$$

Where SMB_t refers to the excess returns of a portfolio at time t, which is long in the small stocks and short in big ones and HML_t denotes the excess returns of a portfolio at time t, which is long in value assets and short in growth ones. Size and Value factors are downloaded from Jensen, Kelly and Pedersen website. Since the database reports the excess returns of factors in USD, for each month I adjust them for the exchange rate between USD and GBP (during the time span, the average exchange rate, considering only the out-of-sample period, was 1 USD for 0.71 Pound). By increasing the number of risk factors, the main implication should be that it becomes more challenging for portfolios to generate alpha.

Nonetheless, the Galton GMV allocation continues to exhibit positive alphas, although none of them is statistically significant across all sizes of the investable universe. In contrast, the benchmark portfolios are still the only ones that demonstrate negative alphas. While Ledoit and Wolf allocation indicates slightly higher estimated alphas compared to the Galton strategies when N is equal to 30 or 50, these differences are still insufficient to reach statistical significance.

Table 20: Alpha estimation with 3FF and FTSE All share index

This table showcases the annualized alpha in percentage points for each investment allocation assuming that the 3 Factor Model holds. I show the White t-statistic in parenthesis

	N=10	N=30	N=50
Galton MV	2.43	1.50	1.13
	(1.11)	(0.84)	(0.58)
MV	6.76	43.81	119.42
	(0.80)	(1.43)	(0.70)
Galton GMV	2.20	1.86	2.46
	(1.00)	(1.16)	(1.31)
GMV	3.42	0.18	8.02
	(1.01)	(0.04)	(0.97)
Empirical Bayes	2.20	1.86	2.46
	(1.00)	(1.16)	(1.31)
Value Weighted	-0.31	-0.93	-0.71
	(-0.15)	(-0.82)	(-0.90)
1/N	1.05	-0.50	-0.17
	(0.47)	(-0.57)	(-0.23)
Ledoit Wolf	1.07	2.10	4.72
	(0.34)	(0.60)	(1.36)

I repeat the same exercise replacing the excess returns of the FTSE All share index with the ones of the value-weighted portfolio. Table 21 displays the annualized alpha and the relative White t-statistic. Once again, the alphas are incredibly increased with some Galton allocations recording statistically significant alphas at 90%, 95% and 99% confidence levels. Also, after the introduction of size and value factors, the alphas of these investment strategies are poorly explained by these factors. Finally, after this adjustment, all the other allocations do not have any significant alpha.

Table 21: Alpha estimation with 3FF and value-weighted allocation

This table 21 the annualized alpha obtained by regressing the excess return of each investment strategy on the excess return of value weighted allocation, size and value factors. I display the White t-statistic in parenthesis. A *, ** and *** means respectively that the alpha is statistically different from 0 with a 90%, 95% and 99% confidence level.

	N=10	N=30	N=50
Galton MV	2.73	2.47	1.80
	(2.58)***	(1.37)	(0.92)
MV	6.16	43.90	114.58
	(0.76)	(1.44)	(0.66)
Galton GMV	2.51	2.87	3.20
	(2.08)**	(1.67)*	(1.60)
GMV	3.63	0.80	8.41
	(1.20)	(0.19)	(1.04)
Empirical Bayes	2.51	2.87	3.20
	(2.08)**	(1.67)*	(1.60)
Ledoit Wolf	1.53	2.91	5.31
	(0.52)	(0.80)	(1.51)

Conclusions

I replicate the Galton strategy proposed by Barroso and Saxena (2021) for a different stock exchange and for an investable universe composed by the 10^{th}, 30^{th} and 50^{th} stocks with the highest market capitalization. I find that also in the United Kingdom stock market, Galton allocations produce higher risk-adjusted measures than the benchmarks considering both the standard deviation and the downside risk as proxies of risk. This investment strategy seems to be optimal also for investors who have CRRA utility functions, having a higher Morningstar Rate Adjusted Return in a context where most of the competitors display a low metric.

Furthermore, Galton allocations result as the best in class in forecasting future risk. This strategy is not too optimistic unlike the competitor portfolios with a realized over ex-ante volatility ratio close to the unity for the different size of allocations. This primacy is also reaffirmed from a risk management perspective with not extreme VaR hit rates both at 95% and 99% confidence levels. This strategy delivers positive, but not statistically significant alpha using as source of systematic risk both the Capital Asset Pricing Model and the 3 Factors Models. Replacing the FTSE All share index with the value-weighted allocation, the annualized alpha generated by the portfolios increased, meaning that the largest stocks sorted each year in that time span record on average low and negative level of alpha.

Moreover, I conduct a simulation consisting of 500 random horse races. At the beginning of each year, I select randomly 50 stocks among the largest 100. This exercise enables me to compute 71,500 (143*500) OOS returns for each investment strategy. The outcome of this exercise is clear: all the Galton allocations clearly outperform the benchmark, the naïve 1/N and their simple versions. Furthermore, the Galton GMV and the Empirical Bayes version statistically outperform the Ledoit and Wolf allocation in 74% of cases. The Galton MV portfolio records a higher risk-adjusted performance than the Ledoit and Wolf in 39% of cases.

Regarding the main limitations of this work, it is important to point out the relatively small length of the sample size for the out-of-sample test, although partially mitigated by the 500 random horse races within an investment universe of 50 companies. Additionally, I do not take into account the turnover of the investment strategies and thus the transaction costs that an investor may face. However, in the appendix, I display additional tables relative to the dynamics of the weights for each portfolio involved in the horse race. Galton allocations require the usage of less extreme and variable weights as opposed to their simple versions.

Learning from past OOS errors seems to offer a better alternative than the classical plug-in approach where the true and unknown population parameters are replaced with their sample estimates. In a demanding out-of-sample context where investors can use only available information, an enough long time series is the only necessary requirement to learn from the not perfect randomness of such errors to construct more robust and performing allocations.

Appendix

In this appendix, I show some tables regarding the dynamics of weights for each investment allocation across the different size of the investment universe. Barroso and Saxena (2021) conduct these exercises in their main paper.

Table 22: Short Positions

This table shows the average sum of short positions expressed in percentage points for each investment strategy.

	N=10	N=30	N=50
Galton MV	-0.97	-11.28	-25.73
MV	-300.72	-1143.37	-5152.28
Galton GMV	-0.02	-5.79	-15.59
GMV	-97.48	-173.45	-458.62
Empirical Bayes	-0.02	-5.79	-15.59
Ledoit Wolf	-12.40	-38.50	-54.30

In table 22 I do not consider the value-weighted and the 1/N allocations since their weights are always positive. For each month, I compute the sum of all the negative weights and at the end of the out-of-sample period, I calculate the average. Simple MV and GMV portfolios are based on a huge amount of short selling as opposed to the Galton investment strategies.

In tables 23 and 24, I compute in each month the minimum and the maximum weight for each portfolio and at the end of the horse race, I take the average. These tables display that after the Galton correction, the weights of the allocations on average vary in a smaller range. It is crucial to point out that Galton allocations provide fewer extreme weights than all the competitors in the horse race.

Table 23: Minimum weights

This table shows the average of minimum weights expressed in percentage points for each investment strategy.

	N=10	N=30	N=50
Galton MV	2.30	-3.84	-4.33
MV	-207.82	-418.34	-1254.79
Galton GMV	3.16	-2.10	-3.11
GMV	-74.42	-60.74	-101.94
Empirical Bayes	3.16	-2.10	-3.11
Value Weighted	6.17	1.04	0.39
1/N	10.00	3.33	2.00
Ledoit Wolf	-6.68	-8.33	-7.55

Table 24: Maximum weights

This table shows the average of maximum weights expressed in percentage points for each investment strategy.

	N=10	N=30	N=50
Galton MV	17.91	10.94	8.95
MV	211.89	411.95	1155.14
Galton GMV	17.25	10.46	8.20
GMV	86.34	70.40	105.03
Empirical Bayes	17.25	10.46	8.20
Value Weighted	16.13	9.12	7.84
1/N	10.00	3.33	2.00
Ledoit Wolf	37.25	25.25	21.68

In tables 25 and 26, I display other two metrics related to the stability of weights which correspond to the average and the standard deviation of the standard deviation of weights. As it is possible to notice, the Galton allocations have lower variability with respect to all the other competitors.

Table 25: Average of the standard deviation of weights

This table shows the average of the standard deviation of weights expressed in percentage points for each investment strategy.

	N=10	N=30	N=50
Galton MV	4.89	3.56	3.14
MV	108.49	132.64	328.56
Galton GMV	4.67	3.08	2.57
GMV	42.34	22.84	30.77
Empirical Bayes	4.67	3.08	2.57
Value Weighted	3.14	2.05	1.74
1/N	0.00	0.00	0.00
Ledoit Wolf	14.28	7.65	5.70

Table 26: Standard deviation of the standard deviation of weights

This table shows the standard deviation of the standard deviation of weights expressed in percentage points for each investment strategy.

	N=10	N=30	N=50
Galton MV	1.78	1.10	0.78
MV	59.20	193.53	1380.81
Galton GMV	1.51	0.83	0.52
GMV	17.76	13.36	12.58
Empirical Bayes	1.51	0.83	0.52
Value Weighted	0.53	0.19	0.14
1/N	0.00	0.00	0.00
Ledoit Wolf	3.61	1.74	0.83

References

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen. "Value and momentum everywhere." *The journal of finance* 68.3 (2013): 929-985.

Barnea, Amir, and David H. Downes. "A reexamination of the empirical distribution of stock price changes." *Journal of the American Statistical Association* 68.342 (1973): 348-350.

Barroso, Pedro, and Konark Saxena. "Lest we forget: Learn from out-of-sample forecast errors when optimizing portfolios." *The Review of Financial Studies* 35.3 (2021): 1222-1278.

Bodie, Zvi, and Alex Kane. "Investments." (2020).

Bollerslev, Tim, and Hao Zhou. "Volatility puzzles: a simple framework for gauging return-volatility regressions." *Journal of Econometrics* 131.1-2 (2006): 123-150.

Cont, Rama. "Empirical properties of asset returns: stylized facts and statistical issues." *Quantitative finance* 1.2 (2001): 223.

Daniel, Kahneman. Thinking, fast and slow. 2017.

DeMiguel, Victor, Alberto Martin-Utrera, and Francisco J. Nogales. "Size matters: Optimal calibration of shrinkage estimators for portfolio selection." *Journal of Banking & Finance* 37.8 (2013): 3018-3034.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?." *The review of Financial studies* 22.5 (2009): 1915-1953.

Engle, Robert F., and Victor K. Ng. "Measuring and testing the impact of news on volatility." *The journal of finance* 48.5 (1993): 1749-1778.

Fama, Eugene F., and Kenneth R. French. "The cross-section of expected stock returns." *the Journal of Finance* 47.2 (1992): 427-465.

Fang, Hsing, and Tsong-Yue Lai. "Co-kurtosis and capital asset pricing." *Financial Review* 32.2 (1997): 293-307.

Galton, Francis. "Regression towards mediocrity in hereditary stature." *The Journal of the Anthropological Institute of Great Britain and Ireland* 15 (1886): 246-263.

Goyal, Amit, and Sunil Wahal. "Is momentum an echo?." *Journal of Financial and Quantitative Analysis* 50.6 (2015): 1237-1267.

Jacobs, Heiko, and Sebastian Müller. "Anomalies across the globe: Once public, no longer existent?." *Journal of Financial Economics* 135.1 (2020): 213-230.

Jagannathan, Ravi, and Tongshu Ma. "Risk reduction in large portfolios: Why imposing the wrong constraints helps." *The Journal of Finance* 58.4 (2003): 1651-1683.

James, William, and Charles Stein. "Estimation with quadratic loss." *Breakthroughs in statistics: Foundations and basic theory* (1992): 443-460.

Jobson, J. David, and Bob Korkie. "Estimation for Markowitz efficient portfolios." *Journal of the American Statistical Association* 75.371 (1980): 544-554.

Kan, Raymond, and Guofu Zhou. "Optimal portfolio choice with parameter uncertainty." *Journal of Financial and Quantitative Analysis* 42.3 (2007): 621-656.

Ledoit, Olivier, and Michael Wolf. "A well-conditioned estimator for large-dimensional covariance matrices." *Journal of multivariate analysis* 88.2 (2004): 365-411.

Ledoit, Olivier, and Michael Wolf. "Honey, I shrunk the sample covariance matrix." *UPF economics and business working paper* 691 (2003).

Lo, Andrew W. "The Statistics of Sharpe Ratios."

Mandelbrot, Benoit. "New methods in statistical economics." *Journal of political economy* 71.5 (1963): 421-440.

Markowitz, Harry M. *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, 1959. *JSTOR*.

Merton, Robert C. "On estimating the expected return on the market: An exploratory investigation." *Journal of financial economics* 8.4 (1980): 323-361.

Michaud, Richard O. "The Markowitz optimization enigma: Is ‘optimized’optimal?." *Financial analysts journal* 45.1 (1989): 31-42.

Morgan, John Pierpont. *RiskMetrics*. 1996.

Newbold, Paul. *Statistics for business and economics*. Pearson, 2013.

Roll, Richard. "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory." *Journal of financial economics* 4.2 (1977): 129-176.

Roy, Andrew Donald. "Safety first and the holding of assets." *Econometrica: Journal of the econometric society* (1952): 431-449.

Rubinstein, Mark E. "A mean-variance synthesis of corporate financial theory." *The Journal of Finance* 28.1 (1973): 167-181.

Sharpe, William F. "Capital asset prices: A theory of market equilibrium under conditions of risk." *The journal of finance* 19.3 (1964): 425-442.

Sortino, Frank A., and Robert Van Der Meer. "Downside risk." *Journal of portfolio Management* 17.4 (1991): 27.

Stein, Charles. "A necessary and sufficient condition for admissibility." *The Annals of Mathematical Statistics* 26.3 (1955): 518-522.

White, Halbert. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: journal of the Econometric Society* (1980): 817-838.

Acknowledgments

I devote this final section of my work to the people who have contributed to its realization through their support.

First and foremost, I would like to express my gratitude to my advisor, Professor Pedro Barroso, for imparting a vast amount of knowledge, for his availability and for always believing in me.

A special thank you to my parents and my brother, to whom I feel I have at least partially repaid their sacrifices. I hope that you are even prouder of me. The mountain has been conquered and the road is finally downhill.

To Emanuela, despite the immense physical distance that separated us this year, I always felt you close to my heart. Alongside my family, I thank you for always being there during the challenging moments of this journey. I look forward to achieving new milestones with you.

To Alessandra, Marco, Michele and Professor Giovanni Walter Puopolo, with whom I spent five intense months for the CFA Challenge. We came so close to our dream, but I am confident that we will try again next year.

To Giuseppe, who has become a new brother for me, if my Portuguese stay was wonderful, it is especially thanks to you.

To Antonio, Michele, and Antonio Raphael, my lifelong friends from Naples, who have never ceased to manifest our friendship even from a distance. I apologize for not being able to attend your graduation ceremonies, but there will surely be time and opportunity to make up for it.

To Martina, who accompanied me almost throughout the duration of my academic journey, if I chose LMEF and Católica, it was also thanks to you.

Finally, I want to dedicate this achievement to those who are no longer with us and left us too soon. You will always remain in my heart.