NEW FRONTIER

Estimation Error and the Fundamental Law of Active Management: Technical Companion

By David Esch, Richard Michaud, and Robert Michaud

Introduction

This document provides supplementary technical and explanatory details of the simulation experiment in "The Fundamental Law of Active Management: Is Quant Fundamentally Flawed," forthcoming in the Journal of Investing, and covered also in the CFA Presentation "Is Quant Fundamentally Flawed?" given on April 16, 2020, which can also be viewed online. In order to provide proper context, some of the material from the forthcoming paper is included in modified form here, but it is generally assumed that the reader is familiar with the main article.

Our simulation experiment is meant to bring together results from previous experiments including Jobson and Korkie (1981) and Frost and Savarino (1988). They studied the simulated out-of-sample performance of (1) budget-only constrained Mean-Variance optimization, (2) positivity-constrained (long-only) optimization and (3) equal-weighted portfolios using various estimation procedures on plausible simulations from real data. In particular, these simulation studies investigated which method has the best average referee-scored performance (information ratios from the simulation parameters) under various conditions of estimation error (controlled by having each method operate on blocks of data simulated from the referee's inputs) and available information (controlled by varying the number of simulation periods in the estimation). While these two studies still stand as important milestones in understanding the impact of estimation error, they did not unite all of the treatments in a single study design. Jobson and Korkie (1981) showed that equal weighting outperformed budget-only optimization under a wide range of treatments. Frost and Savarino (1988) showed that constraints are helpful to out-of-sample optimization performance, to a point, because constraints may minimize the greatest errors in the simulated inputs. Our study examines all of these treatments together in one experiment.

Simulating Adding Breadth while Maintaining Information Levels

In the Grinold and Kahn (GK) application of the fundamental law, each spin of the metaphorical roulette wheel adds one unit of breadth to the investment game. The more spins, the more on average the house wins even for a very small advantage, i. e. discrepancy between the odds and the payout. In our simulations, the critical deviation from the GK roulette wheel framework is that the probability of a win for the investment house is not known or constant but is unstable with estimation error. Our task is to construct a simulation experiment including estimation error where each additive asset adds a realistic unit of breadth for a given IC level.

Simulation Framework

We begin with a sample of historical market return data[1] which will be the basis for all our simulations. The particular dataset is immaterial to our argument. What is essential is that the master dataset represents a realistic vector of expected returns and full-rank covariance matrix for the largest sample size of the experiment.[2]

Simulating Breadth

We propose simulation framework that consists of drawing random permutations (sampling without replacement) grouped into increasing size subsets of the referee's master set of risk-return estimates from the master optimization universe. The averaging of the results of thousands of random permutations from the master stock universe provides a realistic simulacrum of the theoretical concept of linearly increasing breadth for a sequence of investment universes with increasing number of assets. Due to the average linear relationship of universe size to breadth the functional form of average out-of-sample simulation performance can be realistically compared to a monotonic increasing concave function prediction. Without estimation error, the optimizations should bear some resemblance to the Grinold and Kahn square root function of the Fundamental Law. A detailed description of why our estimation error-free deviates slightly from a perfect square root relationship appears in a footnote in the last section of this document.

We draw returns for each simulation from a multivariate normal distribution with parameters corresponding to the referee's master mean and covariance matrix. Each simulation is built from a permutation of 500 stocks, arranged into increasing size portfolios in steps of five to 50 assets and then steps of 50 to 500 assets. Each increase in size adds 5 or 50 new assets to the portfolio, and the referee's truth is the same as the simulation parameters, i. e. the corresponding elements of the master expected return and covariance matrix.

---

[1] We use a recent history of US market data (1994-2013) of publically available data to create our master asset list and corresponding mean and variance parameters. We selected all the assets from the largest 1000 in market capitalization with contiguous data from the period, excluding returns greater than 50% or less than -50% per month. We were able to find 544 stocks that met our criteria. Parallel experiments with shorter histories were also run to investigate if selection bias affects results, with no positive findings, so we present the twenty-year history here. Readers wishing to replicate our experiment can access our data at www.newfrontieradvisors.com/research/data.

[2] A principal components decomposition of our referee's covariance matrix confirms that none of the independent dimensions of the system vanish. All of the eigenvectors are needed to replicate our forecast to reasonable precision. If some of the eigenvalues were vanishingly small, the practical answer to the question of breadth would be quite different from the mathematically rigorous one. However, the full covariance matrix of 500 assets in our dataset has a smallest eigenvalue of over 10 basis points, which is likely significant for most definitions of statistical significance. This would correspond to an annualized standard deviation of approximately 11%, which is substantial by most measures. The submatrices of smaller portfolios tend to have even greater values for the smallest eigenvalue. This line of reasoning confirms that the effective breadth of a sample of size N from our universe is identically N in a practical sense as well as the theoretical one.

Covariance Estimation Issues

We avoid the problem of ill-conditioned or non-full-rank covariance estimation by assuming the referee's truth. This also avoids arbitrariness in covariance estimator choice, on which experts disagree as to best practices, and avoids potential blame for underperformance on badly conditioned covariance matrices or suboptimal estimation procedures.[3] It also means that our results represent a generous upper bound of any practical estimation of the covariance matrix on out-of-sample performance in actual practice. Our results are averages from 16,000 simulations of the process of simulating returns for each of the 19 nested subsets relative to the referee's truth.

Simulating IC

We examine three levels of IC: 0.10, 0.20, and 0.30. The IC levels in the experiment are attained by varying the number of periods of simulated returns in the data blocks used for estimation for optimization inputs. The best numbers of return periods were determined as follows: for a range of numbers, average correlations were computed between the mean of each number of returns and one independently drawn quarter of returns, averaging over many repetitions to eliminate Monte Carlo error. Then, the numbers of return periods most closely matching the target ICs were chosen. Errors were selected to overshoot the target IC, again generously to the optimization methods, in order to eliminate information shortfall as an explanation for underperformance. For our dataset, ICs of approximately 0.10, 0.20, and 0.30 corresponded to 4, 13, and 30 simulation periods of returns. These numbers may be surprising to readers with experience in econometrics because such small sample sizes lead to what are conventionally thought of as very good ICs. The explanation for such good information levels is that in our simulation we have a truly independent and identically distributed (iid) sample from the referee's master parameters. In practice, both independence and identical distribution of historical data are unattainable because of overlapping historical data and other estimation information (affecting independence) as well as the highly dynamic nature of capital markets (changing the return distribution). Given a set of non-identical assets it is not hard to believe that four truly iid samples might lead to a 10% correlation between sample mean and true parameters.

Because of the Monte Carlo nature of our experiment, the average realized ICs for each sample size are not precisely equal to their target values. We present the average realized IC for each target IC and portfolio size in Table 1. The observation sizes for each target IC were determined by calibrating the largest portfolio size (500) for the experiment.[4] While IC levels greater than 0.10 are not formally applicable to predictions from the Grinold formula, our simulations transcend assumptions in the law and may have important teachings in other investment applications.

[3] In particular, this assumption avoids the issues in Fan et al (2008).

[4] Because of the positively skewed distribution of the sample standard deviation in the denominator of the sample correlation formula, the averages tend to be slightly lower for smaller sample sizes, although the realized ICs are still fairly close to their targets. Of course different datasets would probably require different numbers of return periods to attain similar average ICs.

Table 1
Realized IC by Universe Size

IC	5	10	15	20	25	30	35	40	45	50
0.1 (N=4)	0.0982	0.1035	0.1075	0.0978	0.1038	0.1076	0.1129	0.1205	0.1119	0.1127
0.2 (N=13)	0.1810	0.1850	0.1882	0.1948	0.1964	0.1931	0.1980	0.1952	0.1978	0.2005
0.3 (N=30)	0.2597	0.2688	0.2797	0.2848	0.2883	0.2892	0.2876	0.2930	0.2912	0.2935
IC	100	150	200	250	300	350	400	450	500	
0.1 (N=4)	0.1138	0.1120	0.1117	0.1156	0.1140	0.1135	0.1134	0.1125	0.1146	
0.2 (N=13)	0.1996	0.2029	0.2029	0.2017	0.2039	0.2023	0.2010	0.2021	0.2045	
0.3 (N=30)	0.2984	0.3010	0.3020	0.3005	0.2998	0.2997	0.2984	0.3026	0.3019	

The purpose of averaging many permutations of the stock universes is to estimate, on average, the impact of realistic additive breadth with additional stocks in a realistic framework. While each of the five to 500 stock subsets without replacement will necessarily reflect the random vagaries of additive stocks for a particular selection on the results, an average of 16,000 such simulations represents a realistic estimate of linearly additive breadth based on optimization size for a realistic data set of historical returns. While another historical dataset set will exhibit differences, the characteristics of the results we present provides convincing evidence for many cases of practical interest.

Experimental Treatments and Display

In each case of simulated mean and variance inputs, we create MV optimized portfolios via three methods: unconstrained maximum Sharpe ratio, maximum Sharpe ratio with positivity constraints, and equal weighting.[5] Average out-of-sample Sharpe ratios are then calculated for each method using the referee's parameters.

Our displays cover two ranges of optimization universe size in practice: asset allocation and equity portfolio optimization. Asset allocation strategies typically include five to thirty securities and rarely more than fifty. On the other hand, equity portfolio optimization strategies may include hundreds or even thousands of assets in the investment universe.

The no-estimation-error case shown in green in the three panels of Figure 1 is the same in all panels, since the assets used in the system are derived from the same real data with the same means and variances. In simple terms, the green curve reflects the average Sharpe ratio of the referee's return distribution for given optimization universe size free of estimation error, as in the roulette wheel game. Alternatively, it represents increasing the simulation parameter N, which defines the level of estimation error in the IC, as it approaches infinity representing perfect certainty. The additional noise added through estimation error in panels 1, 2, and 3 serve to dilute the signal and are calibrated to attain correlations of 0.1, 0.2, and 0.3 with respect to the true return distribution of the assets. Of course perfect estimation is never attainable in practice.[6]

---

[5] Other portfolio construction methods are possible but not part of the scope of our study. One obvious case is to compute Michaud and Michaud (2008a, b) optimized portfolios with positivity constraints.

[6] If IC is considered a measure of the signal-to-noise ratio, it is important to distinguish two types of noise which dilute the signal and lower the IC for a manager. The first type is determined by the random variation of the returns

Simulation Results

The three panels of Figure 1 show our simulation results for 0.1, 0.2, and 0.3 IC, with sizes of optimization universes ranging from 5 to 500 assets on each panel. Each value presented on the graph is averaged from 16,000 referee-scored simulations. Three curves in each panel show progressions of average Sharpe ratios resulting from three different optimization methods. The "unconstrained" series displays the out-of-sample averages of the simulated unconstrained MSR portfolios, the "equal weight" series displays the average Sharpe ratios of equal weighted portfolios, and the "constrained" series reflects the average Sharpe ratios of out-of-sample simulated long-only MSR portfolios. The fourth curve shows the average Sharpe ratios for unconstrained MV optimization for the no estimation error case.

---

themselves, analogous to the roulette wheel described in Section 1.0. The second type of noise is estimation error, i.e. imperfect estimation of the probabilities associated with the first type. Our simulation framework includes both types of noise in the simulation of IC, but also includes the perfect estimation scenario with no estimation error. Having both allows a simulation of the impact of estimation error on portfolio value, as expressed by the maximum Sharpe ratio.

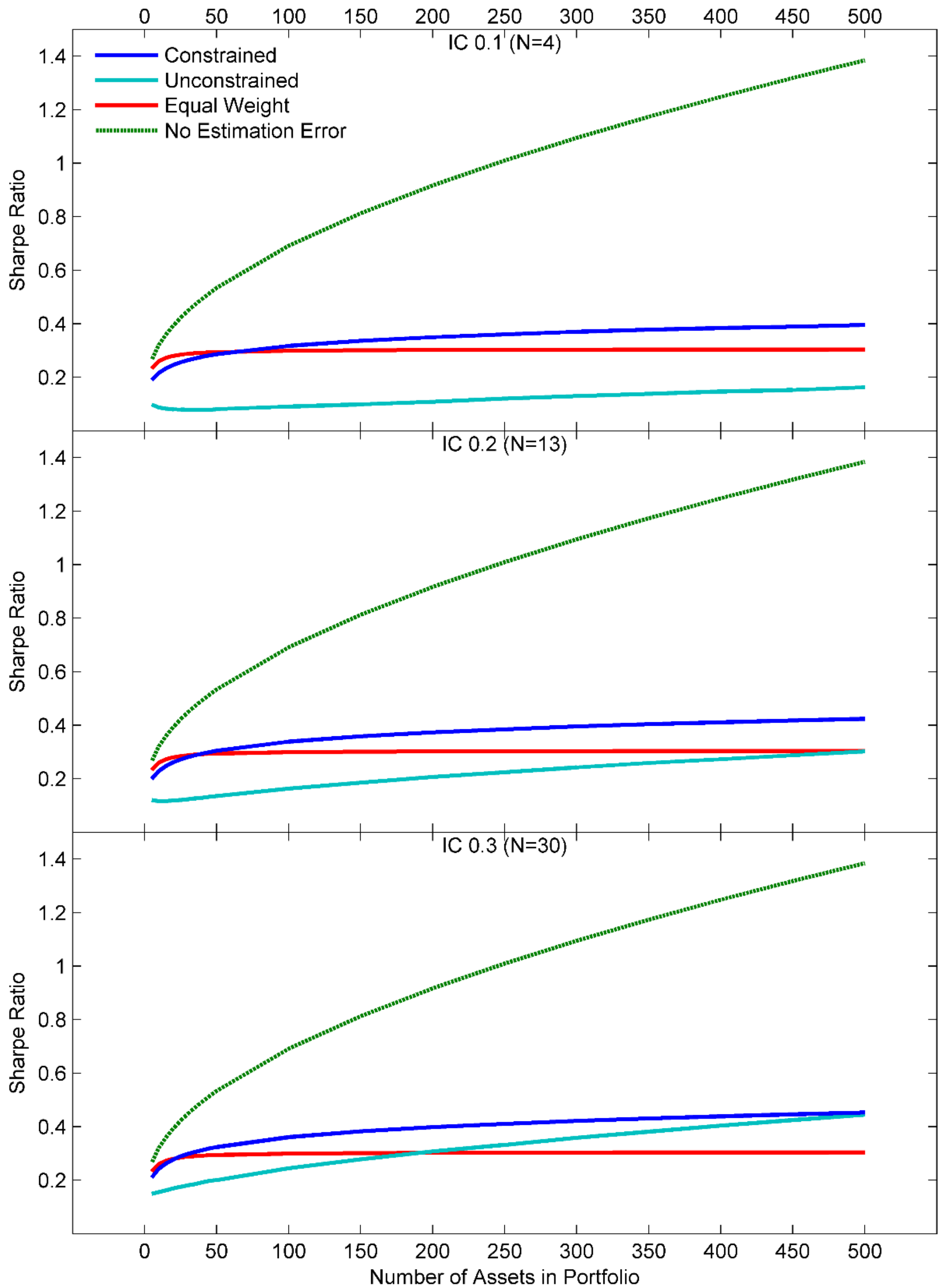

Figure 1: Average Sharpe Ratios for three different portfolio construction methods and three different information coefficients for the equity optimization case, using the referee's covariance matrix. Target information coefficients are not precisely attained by the simulations and realized ICs are shown in Table 1. This experiment was run on many simulations of up to 500 U. S. stocks which had at least 20 years of contiguous monthly price data ending in December 2013.

Our simulations confirm the results in Jobson and Korkie (1981) and Frost and Savarino (1988). On the other hand, our experiments in Figure 1 are stark and dramatically at odds from principles of optimization portfolio design associated with applications of the Grinold formula from GK and CST and others. Growth in average Sharpe ratios is far less than the no estimation error relationship as a function of universe size posited in GK or in the value of unconstrained MV optimization posited in CST.[7] In particular, note that unconstrained optimized portfolios may dramatically underperform both sign constrained and equal weighting out-of-sample for small optimization universes. Furthermore, note how positivity constraints depend on the quality of information and universe size. For larger portfolio sizes, the optimized cases often outperform the equal weighted case, with better performance for greater information levels and for positivity constraints.

In the case of IC equal to 0.30, the out-of-sample unconstrained performance nearly attains the level of the constrained case for the largest sample size of 500 assets. However, it is essential to note that the Grinold formula does not apply to IC levels greater than 0.1. The Grinold proof would require revision of the functional form of the formula. In addition, these experiments assume clairvoyant forecasts. There is no consideration of financial frictions or investment costs of any kind that would likely severely limit the investment value of large optimization universe asset management. In addition, our assumption of an error free covariance matrix further upward biases our simulations. Our results vividly demonstrate the hazards of ignoring estimation error for optimization design.

Further Discussion

Our deliberate optimism on how additive breadth is modeled when increasing the size of the optimization universe in the simulations has important implications. All of the assets in the simulation universe are assumed to have some investment value. Consequently, an investor is little harmed by putting portfolio weight on a "wrong" asset. In the real world, constraints often limit the harm caused by misinformation. In a truly chaotic world with a lot of estimation error

---

[7] The no-estimation-error/roulette cases in the three panels of Figure 1 deviate somewhat from a precise square root function posited by Grinold for two reasons. Firstly, because the information in the means and variances of returns has a well-known factor structure, with the common factors explaining far more of the total information (variance) of the system than the individual idiosyncratic elements for each asset. Because of these large IC units of breadth, the roulette curve starts at a point greater than the zero intercept of the GK curve. In other words, the units of breadth coming from the first few common factors have an IC that is likely greater than the ICs of the units of breadth informing the idiosyncratic variances of the assets. The random ordering of the individual simulations and averaging guarantees that the idiosyncratic variances contribute to breadth on average linearly with the addition of assets, and have equivalent average IC as well, so the no-estimation-error curves do have the generally concave and monotone increasing shape of a square root law. Secondly, the no-estimation-error curves are not precise square root functions because our curves are measuring maximum Sharpe ratio, i.e. the rise over run on the mean-variance efficient frontier of the absolute weights, rather than the slope of the tangency portfolio or equivalently the slope of the active (benchmark-relative) unconstrained frontier, which is a straight line emanating from the origin. Because of these two considerations, our simulated no-estimation-error curves are not required by GK theory to be precisely square root curves, yet they remain close in shape to square root relationships. Our simulations show that the addition of estimation error to the system drastically changes the response curve of the maximum attainable Sharpe ratio from a vaguely square-root-like function to something else entirely.

and bias, the equal weighted portfolio, which uses no “wrong” information to distinguish among assets, can be hard to beat, for small optimization universes such as in asset allocation strategies.

The consistent slow rising level of unconstrained average maximum Sharpe ratios as universe size increases is a necessary artifact of our simulation framework. This is because, by design, our simulations assume a consistent level on average of realized IC regardless of universe size. In practice, many investment strategies have an optimal universe size. Beyond some point, increasing universe size is likely to be self-defeating in practice.

References

Frost, P. and J. Savarino. 1988. “For Better Performance: Constrain Portfolio Weights.” *Journal of Portfolio Management* 15(1):29-34.

Grinold, R. 1989. “The Fundamental Law of Active Management.” *Journal of Portfolio Management* 15(3): 30-37.

Grinold, R. and R. Kahn. 1995. *Active Portfolio Management: Quantitative Theory and Applications*. Chicago: Irwin.

Grinold, R. and R. Kahn, 1999. *Active Portfolio Management: A Quantitative Approach for Providing Superior Returns and Controlling Risk*. 2nd ed. New York: McGraw-Hill.

Jobson, D. and B. Korkie, 1981. “Putting Markowitz Theory to Work.” *Journal of Portfolio Management* 7(4): 70-74.

Michaud, R., Esch, D. and R. Michaud, 2020. Estimation Error and the Fundamental Law of Active Management: Is Quant Fundamentally Flawed. Forthcoming in the *Journal of Investing*.

Michaud, R. and R. Michaud. 2008a. *Efficient Asset Management*. 2nd ed. Oxford University Press: New York.

Michaud, R. and R. Michaud 2008b. “Estimation Error and Portfolio Optimization: A Resampling Solution.” *Journal of Investment Management* 6(1):8-28.