Estimation Error and the "Fundamental Law of Active Management"

By

Richard O. Michaud, David N. Esch, and Robert O. Michaud *

Draft: September 2018

* Richard O. Michaud is CEO of New Frontier Advisors, LLC, Boston, MA 02110; David N. Esch is Managing Director of Research of New Frontier Advisors LLC; Robert O. Michaud is CIO of New Frontier Advisors, LLC.

Abstract

According to applications of Grinold's "Fundamental Law," simply adding securities to an optimization universe, adding factors to forecast return, trading more frequently, and reducing constraints will add investment value to an investment strategy. We show with intuitive discussion followed by a novel simulation study that the proposals are unreliable and often self-defeating. This is because the Grinold formula ignores real world investment issues of estimation error and necessary constraints for practice. A substantial fraction of globally professionally managed funds are estimated to employ optimized portfolio design principles that are applications of Grinold's "Fundamental Law."

Benchmarks arise naturally in judging asset manager competence and for meeting investment goals. An active investment manager typically claims to provide enhanced return on average relative to a given benchmark or index for a given level of residual risk. The information ratio (IR) – estimated return relative to benchmark per unit of residual risk or tracking error – is a convenient and ubiquitous framework for measuring the value of active investment strategies.

The Grinold (1989) "Fundamental Law of Active Management" asserts that the maximum attainable IR is approximately the product of the Information Coefficient (IC) times the square root of the breadth (BR) of the strategy.[1] The IC represents the manager's estimated correlation of forecast with ex post residual return while the BR represents the number of independent bets or factors associated with the strategy. Grinold and Kahn (1995, 1999) assert that the "law" provides a simple framework for enhancing active investment strategies. While a manager may have a relatively small amount of information or IC for a given strategy, performance can be enhanced by increasing BR or the number of independent bets in the strategy. In particular they state: "The message is clear: It takes only a modest amount of skill to win as long as that skill is deployed frequently and across a large number of stocks."[2] Their recommendations include increasing trading frequency, size of the optimization universe, and factors to models for forecasting return. Assumptions include: independent sources of information and IC the same for each added bet or increase in BR.

Clarke, de Silva and Thorley (2002, 2006) (CST) generalize the Grinold formula by introducing the "transfer coefficient" (TC) to the Grinold formula. TC is a scaling factor that measures how information in individual securities is "transferred" into optimized portfolios. TC represents a measure of the reduction in investment value from optimization constraints. This widely influential article has been used to promote many variations of hedge fund, long-short, alternative, and unconstrained investment strategies.[3]

A significant literature exists on applying the Grinold law and variations for rationalizing various active equity management strategies. Extensions include: Buckle (2004), Qian and Hua (2004), Zhou (2008), Gorman et al (2010), Ding (2010), Huiz and Derwall (2011), Ding and Martin (2017). Industry tutorials and perspectives include Kahn (1997), Kroll et al (2005), Utermann (2013), Darnell and Ferguson (2014), Menchero (2017). Teachings include the Chartered Financial Analyst (CFA) Institute Level 2, the Chartered Alternative Investment Analyst (CAIA) Level 1 and many conferences and academic courses in finance. Texts discussing the formula and applications include Focardi and Fabozzi (2004), Jacobs and Levy (2008), Diderich (2009), Anson et al (2012), Schulmerich et al (2015). A very substantial fraction of globally professionally managed funds are estimated to employ optimized portfolio design principles that are applications of Grinold's "Fundamental Law."

We show with both qualitative discussion and novel simulation studies that the GK and CST proposals for optimized portfolio design are unreliable and may often be self-defeating. This is because estimation error and investment realistic constraints are ignored in the Grinold formula.[4] These principles have been largely unchallenged for more than twenty-five years and are estimated to impact a substantial fraction of globally professionally managed active equity funds.

The outline of the paper is as follows. Section 1 presents the Grinold formula, the GK and CST prescriptions for active management with reference to the GK casino management rationale. Section 2 discusses the limitations of the GK and CST prescriptions from an intuitive investment perspective. Section 3 provides a discussion of properties of index-relative mean-variance (MV) optimization and previous simulation studies relevant to our results. Section 4 presents our Monte Carlo simulation study that demonstrates that applications associated with the fundamental law are likely invalid and often self-defeating. Section 5 provides a summary and conclusions.

1.0 Grinold's Fundamental Law of Active Management

The Grinold (1989) formula is an approximate decomposition of the information ratio (IR) generally associated with active equity investment management. Grinold shows that the MV optimization of an inequality unconstrained residual return investment strategy is approximately proportional to the product of the square root of the breadth (BR) and the information correlation (IC).[5] Mathematically,

$$IR \cong IC * \sqrt{BR}$$

where IR = information ratio = (alpha) / (residual or active risk)
IC = information correlation (ex ante, ex post return correlation)
BR = breadth or number of independent sources of information.

The formula teaches that successful active management depends on both the information level of the forecasts and the breadth associated with the optimization strategy. However, Grinold and Kahn (GK) (1995, 1999) and Clarke, deSliva, and Thorley (CST) (2002, 2006) go further. They apply the Grinold formula to assert that only a modest amount of information (IC) is necessary to win the investment game simply by sufficiently increasing the number of assets in the optimization universe, the number of factors in a multiple valuation framework, more frequent trading and reducing optimization constraints.

GK use a casino roulette game to rationalize applications of the Grinold formula to asset management in practice.[6] The probability or IC of a winning play (for the casino) of the roulette game is small but more plays (breadth) lead to the likelihood of more wealth. However, there are important differences between the play of a roulette game in a casino and the play of an investment game in practice. In the casino context, all probabilities are known, therefore the IC is stable, known, positive, and constant across plays. In an investment game the IC is unstable, has estimation error, and may be insignificantly or even negatively related to return. Plays of the investment game may not be independent and increasing the number may be undesirable. While interesting, the casino game rationale for rationalizing applications of the Grinold formula to investment practice is invalid.

There are two fundamental reasons for limitations of the principles associated with the Grinold law for practical asset management: 1) the formula ignores the impact of estimation error in investment information on out-of-sample optimized investment performance; 2) the formula assumes a quadratic utility unconstrained optimization framework that ignores the necessity of including investment realistic constraints required for defining portfolio optimality in practice.

2.0 Discussion of GK and CST Prescriptions

GK and CST propose four principles of optimized portfolio design for enhanced investment value in an index-relative MV optimization framework. We discuss the limitations of each prescription in turn from an intuitive point of view.

2.1 Large Optimization Universe Fallacy

GK argue that investment value increases with the size of the optimization universe conditional that the IC is roughly equal for all securities in the optimization universe. How realistic is this assumption?

For a small universe of securities, the assumption of uniform average IC may be tenable. Small universes may be fairly homogeneous in character. However, for a large and expanding optimization universe, it seems untenable to assume uniform average IC across all subsets. Any manager will naturally use the securities with the best information first. While, theoretically, adding more assets may add marginally to breadth, all other things the same, it is also likely to result in less predictable securities and reduce the overall average IC level of the universe. A lower average IC may cancel any gains made from increasing breadth.

The issue can be framed in a more common practical setting. Consider an analyst suddenly asked to cover twice as many stocks. Given limitations of time and resources, it is highly unlikely that the analyst's average IC is the same for the expanded set of stocks. Issues of resources and time are the primary reasons why analysts tend to specialize in areas of the market or use managers in investment strategies that limit the number of securities that they cover. In practice many traditional managers limit the number of securities they include in their active portfolio to not much more than twenty or fifty. Except for relatively small asset universes, the average IC and overall level of IR may often be a decreasing function of the number of stocks in the optimization universe, all other things the same. Grinold and Kahn seem to be aware of these limitations, for example as suggested by their statement "The fundamental law says that more breadth is better, provided the skill can be maintained." Nevertheless, average IC and optimization universe size are often negatively correlated in practical applications.

2.2 Multiple Factor Model Fallacy

Large stock universe optimizations typically use indices such as the S&P500, Russell 1000 or even a global stock index as benchmarks. In this case individual analysis of each stock is generally infeasible and analysts typically rely on factor valuation frameworks for forecasting alpha. For example, stock rankings or valuations may be based in part on an earnings yield factor.[7] As GK note, if earnings yield is the only factor for ranking stocks, there is only one independent source of information and breadth equals one.

In the Grinold formula, the IR increases with the number of independent positive significant factors in the multiple valuation forecast model. However, in practice, asset valuation factors are often highly correlated and may often be statistically insignificant, providing dubious out-of-sample forecast value.[8] Finding factors that are reasonably uncorrelated and significantly positive relative to ex post return is no simple task.

Factors are often chosen from a small number of categories considered to be relatively uncorrelated and positively related to return such as value, momentum, quality, dividends, and discounted cash flow.[9] In experience, breadth of multiple valuation models is typically very limited and unlikely to be very much greater than five independent of the size of the optimization universe.[10,11] As in adding stocks to an optimization universe, adding factors at some point is likely to include increasingly unreliable factors that are likely to reduce, not increase, the average IC of an investment strategy.

Michaud (1990) provides a simple illustration of the impact of adding factors to a multiple valuation model. While adding investment significant factors related to return can be additive to IC, it can also be detrimental in practice. There is no free lunch. Adding factors can as easily reduce as well as enhance investment value, and the number of factors that can be added while maintaining a desirable total IC is generally limited in practice.

2.3 Invest Often Fallacy

GK recommend increasing trading period frequency or “plays” of the investment game to increase the BR, and thus the IR of a MV optimized portfolio. The Grinold formula assumes trading decision period independence and constant IC level. However, almost all investment strategies have natural limits on trading frequency.[12] For example, an asset manager trading on book or earnings to price will have significant limitations increasing trading frequency smaller than a month or quarter. Reducing the trading period below some limit will generally reduce effectiveness while increasing trading costs.

Fundamentally, trading frequency is limited by constraints on the investment process relative to investment style.[13] Deep value managers may often be reluctant to trade much more than once a year while growth stock managers may want to trade multiple times in a given year. Increased trading, to be effective, requires increasing the independence of the trading decision while maintaining the same level of skill. This will generally require increased resources, if feasible, all other things the same. The normal trading decision period should be sufficiently frequent, but not more so, in order to extract relatively independent reliable information for a given investment strategy and market conditions.

It is worth noting that the notion of normal trading period for an investment strategy does not imply strict calendar trading. Portfolio drift and market volatility relative to new optimal may require trading earlier or later than an investment strategy “normal” period. In addition a manager may need to consider trading whenever new information is available or client objectives have changed. Portfolio monitoring relative to a normal trading period including estimation error is further discussed in Michaud et al (2012).

2.4 Remove Constraints Fallacy

Markowitz's (1952, 1959) MV optimization can accommodate linear equality and inequality constraints. In actual investment practice, MV optimized portfolios typically include many linear constraints. This is because unconstrained MV optimized portfolios are often investment unintuitive and impractical. Constraints are often imposed to manage instability, ambiguity, poor diversification characteristics, and enhanced out-of-sample performance.[14] However, constraints added solely for marketing or cosmetic purposes may result in little, if any, investment value and may obstruct the deployment of useful information in risk-return estimates.

In general, inequality constraints are necessary in practice. Inequality constraints reflect the financial fact that even the largest financial institutions have economic shorting and leveraging limitations. Recently, Markowitz (2005) demonstrates the importance of practical linear inequality constraints in defining portfolio optimality for theoretical finance and the validity of many tools of practical investment management. Long-only constraints limit liability risk, a largely unmeasured factor in most portfolio risk models and often an institutional requirement. Regulatory considerations may often mandate the use of no-shorting inequality constraints. Performance benchmarks may often mandate index related sets of constraints for controlling and monitoring investment objectives. Moreover, inequality constraints limit the often negative impact of estimation error in out-of-sample performance (Frost and Savarino 1988).

3.0 Properties of index-relative MV optimization

The Grinold formula concerns properties of the index-relative MV efficient frontier. Index-relative MV optimization is total return MV optimization with an index weight sum to zero constraint. The efficient frontier starts at the origin where the index portfolio has no index relative risk and return by definition. In the unconstrained case the index-relative MV efficient frontier is a straight line starting from the origin with slope equal to IR. In the sign constrained case the efficient frontier is concave as a function of index-relative risk except for a straight line segment rising from the origin with slope equal to IR until the first pivot portfolio on the constrained efficient frontier.

Roll (1992) provides the classic critique of the index-relative active return MV optimization framework. He shows that index-relative MV optimized portfolios are dominated by portfolios in MV total return space unless the index is total return MV efficient.[15] Merton (1987) provides a rational market framework for index-relative MV optimization. Under relatively straightforward conditions consistent with many active asset management strategies, common benchmarks may often be considered total risk and return MV efficient. In the following, presuming economically rational agents, we assume the index chosen is total risk and return MV efficient. This assumption is a best-case scenario for the investment value of an index-relative active MV optimized investment strategy.

Assuming Merton MV efficiency for the index, all index-relative MV optimized portfolios are also MV total return efficient. Each max IR portfolio is conditional relative to a specified level of tracking error to the index. There is no in-sample index-relative MV efficient portfolio independent of a specified tracking error that uniquely represents the out-of-sample investment characteristics of an index-relative optimized investment strategy. However, note that the max Sharpe ratio (MSR) portfolio is both index-relative and total return efficient under our rationality assumptions. It is also a convenient identifiable single in-sample MV efficient portfolio to represent the out-of-sample investment value of the strategy relative to the inputs. In the following, the in-sample optimal MSR portfolio is used to represent the out-of-sample investment value of max IR investment strategies for both unconstrained and sign constrained investment strategies.

3.1 Testing GK and CST proposals

Investment managers often use a back test to demonstrate the likely value of a proposed investment strategy. In this procedure a factor or strategy is evaluated on how it performed for some historical data over some time period. While the benefit of a back test may be practicality, no reliable prospective information is possible by definition. Back tests are notorious for misleading investors, resulting in loss of wealth, careers, and dissolution of firms. Investors should be keenly aware of the serious limitations of any back test as evidence of the reliability of any factor relationship or investment strategy.[16]

A simulation study is a far more reliable framework for testing the value of optimized investment strategies. Such a procedure evaluates the likely out-of-sample performance of an in-sample optimized portfolio for many realistic investment scenarios.

In the following sections we explain the summary statistics used to evaluate the out-of-sample performance of investments from following the prescriptions of the fundamental law, describe the simulation test framework in greater detail, and discuss the results of our simulation experiment.

3.2 Portfolio simulation study framework

Our study uses a framework similar to other well-known simulation studies for portfolio construction methods.[17] In this framework, a referee is assumed to know the true means, standard deviations, and correlations for a set of assets and consequently the true MSR for an optimized portfolio of those assets. The players do not know the referee's true MV parameters. The players receive simulated returns based on the referee's parameters, so they can only observe the truth obscured by estimation error, as is true for all real-world investment managers. The players then compute optimal weights for their inputs and send them back to the referee to score. The referee determines the estimated MSR for that simulation. The procedure is repeated many times for a range of referee simulated scenarios, and averages of MSRs computed for each player. In this way the out-of-sample performance of each strategy can be compared, and the better overall strategy determined.

3.3 Prior MV optimization simulation studies

Jobson and Korkie (1981) provide the classic study of the effect of estimation error on the out-of-sample investment value of inequality unconstrained MV optimized portfolios.[18] In their study the referee's truth is based on historical MV inputs for twenty stocks. They compute simulated MV inputs reflecting five years of monthly return data. They find that the average of the true Sharpe ratios (SRs), as measured by the referee, of simulated MSR optimal portfolios, was twenty-five percent of the MSR of the referee's optimal portfolio. In addition they show that equal weighting substantially outperforms the optimized portfolios.[19] They conclude that unconstrained MV optimization is not recommendable for practice.

Frost and Savarino (1988) perform a related simulation study for long-only MV optimized portfolios. They find that additional constraints may often add investment value to the out-of-sample performance of MV optimized portfolios. Economically realistic constraints may often act like Bayesian priors focused on portfolio structure enforcing rules representing legitimate information not contained in the optimization inputs. Such restrictions can mitigate estimation error in risk-return estimates implicitly by forcing the simulations towards more likely optimal portfolios.

We note that the Jobson and Korkie and Frost and Savarino studies contradict the theoretical results of CST for two fixed size stock universes. Our study confirms their results in the context of optimization universe size on the out-of-sample relationship between positivity-constrained and unconstrained optimization methods.

4.0 Simulating Adding Breadth while Maintaining Information Levels

In the GK application of the Grinold formula, each spin of the 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 amount of information. 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 in the context of estimation error where each additive asset adds a realistic unit of breadth for a given IC level.

4.1 Simulation Methodology

We begin with a sample of historical market return data[20] of 500 stocks which will be the basis for all our simulations. This 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.[21]

We propose a novel simulation framework that consists of random sampling without replacement of increasing size subsets of the 500 stocks of the referee's risk-return estimates from the master optimization universe. The averaging of the results of thousands of samplings without replacement from the given 500 stock universe with increasing size subsets provides a realistic estimate of the theoretical concept of increasing a unit of breadth for increasing number of assets. In this way the functional form of average out-of-sample simulation performance can be realistically compared to a monotonic increasing concave function prediction.

We Monte Carlo simulate returns assuming a multivariate normal distribution for the referee's mean and covariance matrix. Each simulation consists of sampling without replacement of the 500 stocks of increasing size subsets in steps of five to 50 assets and steps of 50 to 500 assets. In each subset sample of stocks, the referee's truth is computed by independently adding assets to the referee's expected return vector, and a new row and column to the referee's covariance matrix. From this we simulate sample means for each subset universe. We avoid the problem of computing a sample covariance from the simulated returns that is not full-rank by assuming the referee's truth. This assumption eliminates non-full-rank covariance estimation from simulated returns as a plausible explanation of our results.[22] 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 randomly chosen without replacement subsets relative to the referee's truth.

We examine three levels of IC: 0.10, 0.20, and 0.30. The IC level of a simulated set of returns is computed by varying the number of periods of simulated returns for each size universe. For our dataset, ICs of approximately 0.10, 0.20, and 0.30 corresponded to 4, 13, and 30 simulation periods of returns. Because of the Monte Carlo nature of our experiment, the average realized ICs for each sample size is not exactly equal to 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.[23] 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.

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 many samplings without replacement of the process of building up subsets of the 500 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, the consequence of an average of 16,000 such simulations represent a realistic estimate of additive breadth based on optimization size for a realistic data set of historical returns. While some other historical data set will reflect differences, the characteristics of the results we present provides convincing evidence for many cases of practical interest.

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.[24] 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.[25]

4.2 Simulation Results

Figure 1 consists of three panels of simulation results for 0.1, 0.2, and 0.3 IC levels of estimation error, for sizes of optimization universes ranging from 5 to 500 assets. Each value presented on the graph is averaged from 16,000 samplings without replacement optimizations. The three graphed series 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 graph reflects the average Sharpe ratios for unconstrained MV optimization for the no estimation error case as a function of optimization universe size.

Our simulations confirm the pioneering 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. Improvement 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.[26] In particular, note that unconstrained optimized portfolios may dramatically underperform both sign constrained and equal weighting out-of-sample for small optimization universes. Further, 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.[27] 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.

4.3 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 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.

5.0 Summary and conclusions

Our narrative does not contradict the simple intuition that investment performance is a function of skill and breadth. It is always true that it is better to have more reliable information (IC) and more additional investment opportunities to apply it (BR). But Grinold's formula is no scientific law in the sense of Newton's second law of motion. Naively adding assets or factors will not necessarily add portfolio performance. There is a rich interaction between quality of information, breadth, and estimation error that goes well beyond the simplicity of the Fundamental Law.

Applications of the Grinold formula are based on the false premise in practice that estimation error and realistic constraints are immaterial and that IC and BR can be thought of independently in order to formulate optimization design. Section 2 qualitatively discusses why it's not reasonable to assume a manager can increase BR without affecting IC in practice. Section 4 quantitatively demonstrates that even under the highly idealized conditions of our simulation study where BR can be additively applied while holding IC level constant, the results fall dramatically short of predicted relationships.

Our discussions and simulation studies indicate that the popular four principles of optimized portfolio design proposed by GK and CST from applications of the Grinold formula – frequent trading, adding securities, adding forecast factors, and removing constraints – are not reliably beneficial and may often be self-defeating. Our simulations of the out-of-sample average max Sharpe ratio improvement as a function of breadth modeled as additive size of the optimization universe contradicts anything similar to a square root improvement.

Our results have important implications for contemporary investment practice. For more than twenty-five years, a substantial fraction of professionally managed quantitative equity funds are estimated to employ optimized portfolio design principles and marketed investment strategies that are applications of Grinold's "Fundamental Law." In particular, many rationales for investing in hedge fund, long-short, and unconstrained strategies may be impacted.

The necessary conditions for reliably winning the investment game remain the fundamental principles of reliable long-term asset management: 1) investment significant information and high quality investible assets relevant to a given size optimization universe; 2) economically meaningful constraints; and 3) properly implemented estimation error sensitive portfolio optimization technology.

The root of the failure of applications of the theoretical Grinold formula goes well beyond the assumption that investment management can be modeled with the operation of a casino game. It is one of many examples of the fundamental and ubiquitous fallacy in many areas of modern science of regarding inference from in-sample statistics and fixed probability models as the full measure of uncertainty.[28]

Figure 1
Average Maximum Sharpe Ratios by Information Level, Referee Covariance

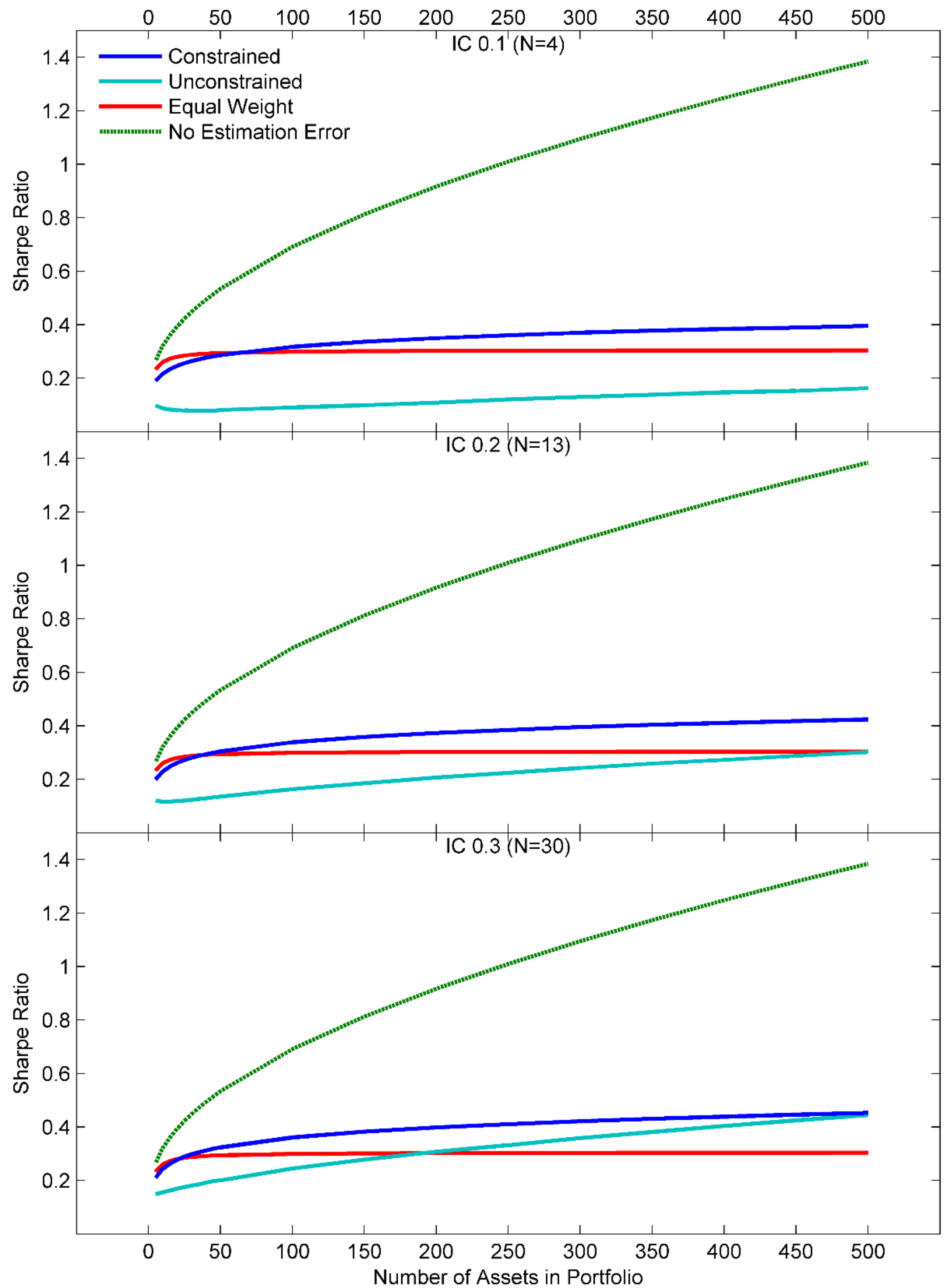

Figure 1: Average Sharpe Ratios for three different portfolio construction methods and three different information coefficients for the equity optimization case, using the referee's covariance matrix. Target information coefficients are not precisely attained by the simulations and realized ICs are shown 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.

References

Anson, M., D. Chambers, K. Black, H. Kazemi. 2009. *An Introduction to Core Topics in Alternative Investments* CAIA Association, 2nd ed. 2012. Wiley.

Buckle, D. 2004. "How to Calculate Breadth: An Evolution of the Fundamental Law of Active Portfolio Management." *Journal of Asset Management* 4(6):393–405.

Clarke, R, H. deSilva, S. Thorley. 2002. "Portfolio Constraints and the Fundamental Law of Active Management." *Financial Analysts Journal* 58(5): 48-66.

Clarke, R., H. deSilva, S. Thorley. 2006. "The Fundamental Law of Active Portfolio Management." *Journal of Investment Management*, 4(3): 54-72.

Darnell, M. and K. Ferguson, 2014. "Thoughts On Grinold & Kahn's Fundamental Law Of Active Management." First Quadrant Perspective.

Diderich, C. 2009. *Positive Alpha Generation: Designing Sound Investment Processes*, Wiley.

Ding, Z. 2010. "The Fundamental Law of Active Management: Time Series Dynamics and Cross-Sectional Properties." Russell and Thaler Asset Management working paper.

Ding, Z and R.D. Martin, 2017. "Fundamental Law of Active Management: Redux." *Journal of Empirical Finance*, 43(September):91-114.

Fama, E. F., & French, K. R. (1992). "The cross-section of expected stock returns." *Journal of Finance*, 47(2): 427-465.

Fan, J., Fan, Y., & Lv, J. (2008). "High dimensional covariance matrix estimation using a factor model." *Journal of Econometrics*, 147(1): 186-197.

Farrell, J. 1983. *Guide to Portfolio Management*. Wiley, NY.

Focardi, S. and F. Fabozzi, 2004. The Mathematics for Financial Modeling and Asset Management, Wiley.

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

Gorman, L., S. Sapra, R. Weigand. 2010. "The role of cross-sectional dispersion in active portfolio management." *Investment Management and Financial Innovations*, 7(3):54-64.

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.

Huij, J. and J. Derwall. 2011. "Global Equity Fund Performance, Portfolio Concentration, and the Fundamental Law of Active Management." *Journal of Banking & Finance*, 35(1):155-165.

Jacobs, B. and K. Levy, 2008. *Investment Management: An Architecture for the Equity Market*. Wiley

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

Kahn, R. 1997. “The Fundamental Law of Active Management: Seven Quantitative Insights into Active Management—Part 3,” BARRA Newsletter Winter.

Knight, F. 1921. *Risk, Uncertainty, and Profit*. Houghton Mifflin Company: Boston.

Kritzman, M., S. Page, D. Turkington, 2010. “In Defense of Optimization: The Fallacy of 1/N,” *Financial Analysts Journal,* 66(2):31-39.

Kroll, B. D. Trichilo, and J. Braun, 2005. “Extending the Fundamental Law of Investment Management (or How I Learned to Stop Worrying and Love Shorting Stocks).” JP Morgan Asset Management white paper.

Ledoit, O., and Wolf, M. 2004. "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices". *Journal of Multivariate Analysis*, 88(2): 365-411.

Markowitz, H. 1952. “Portfolio Selection.” *Journal of Finance* 7(1): 77-91.

Markowitz, H., 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley and Sons. 1991. 2nd ed. Cambridge, MA: Basil Blackwell.

Markowitz, H. 2005. “Market Efficiency: A Theoretical Distinction and So What?” *Financial Analysts Journal* 61(5): 17-30.

Menchero, J. 2017. “Rethinking the Fundamental Law of Active Management.” *Journal of Investment Management* 15(2):92-107.

Merton, R. 1987. “Presidential Address: A Simple Model of Capital Market Equilibrium with Incomplete Information.” *Journal of Finance* 42(3): 483-510.

Michaud, R. 1989. “The Markowitz Optimization Enigma: Is Optimized Optimal?” *Financial Analyst Journal* 45(1):31-42.

Michaud, R. 1990. “Demystifying Multiple Valuation Models.” *Financial Analysts Journal* 46(1):6-8.

Michaud, R. 1993. “Are Long-Short Equity Strategies Superior?” *Financial Analyst Journal* 49(6):44-49.

Michaud, R. 1999. *Investment Styles, Market Anomalies, and Global Stock Selection*. Research Foundation of the Chartered Financial Institute, Charlottesville.

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.

Michaud, R., D. Esch, and R. Michaud 2012. “Portfolio Monitoring in Theory and Practice*,” Journal Of Investment Management* 10(4):5-18.

Qian, E and R. Hua., 2004. “Active Risk and Information Ratio.” *Journal of Investment Management*, 2(3):1-15.

Roll, R. 1992. "A Mean/Variance Analysis of Tracking Error." *The Journal of Portfolio Management*, 18(4):13-22.

Schulmerich, M., Y. Leporcher, and C. Eu. 2015. *Applied Asset and Risk Management,* Springer: Heidelberg

Utermann, A. 2013. “The Changing Nature of Equity Markets and the Need for More Active Management” Allianz GI Global Capital Markets & Thematic Research.

Weisberg, H. 2014. *Willful Ignorance: The Mismeasure of Uncertainty*. Wiley: New York.

Zhou, G. 2008. “On the Fundamental Law of Active Portfolio Management: What Happens If Our Estimates Are Wrong?” *Journal of Portfolio Management,* 34(4):26-3.

Endnotes

[1] The Grinold formula is analytically derived and based on an inequality unconstrained maximization of quadratic utility. It should not be confused with Markowitz (1952, 1959) which assumes linear (inequality and equality) constrained portfolios and requires quadratic programming techniques to compute the MV efficient frontier. In particular, the Markowitz efficient frontier is generally a concave curve in a total or residual return framework while in Grinold (see e.g., GK 1995, p. 94) it is a straight line emanating from a zero residual risk and return benchmark portfolio. The Grinold derivation also assumes IC to be small, in the order of 0.1.
[2] GK (1995, Ch. 6, p. 130), also GK (1999, Ch. 6, p. 162).
[3] One example is Kroll et al (2005). Michaud (1993) was the first to note possible limitations of the long-short active equity optimization framework.
[4] To be clear, we note that the term "estimation error" in this manuscript refers to Monte Carlo simulation experiments that measure how estimates of optimization parameters impact out-of-sample investment performance, and not, as in Zhou (2008) or Kritzman (2010), to refer to statistical estimation issues.
[5] The detailed derivation is given in GK, Ch. 6, and Technical Appendix.
[6] The casino roulette game framework is very consistent with the assumptions used in the Grinold derivation in GK (1995, 1999, Ch. 6. App.)
[7] Some standard methods for converting rankings to a ratio scale to input to a portfolio optimizer include Farrell (1983) and references.
[8] There is a practical limit to the number of independent investment significant factors even in many commercial risk models, often far less than ten.
[9] Standard methods such as principal component analysis for finding orthogonal risk factors are seldom also reliably related to return over independent periods.
[10] See e.g., Michaud (1999).
[11] While principal component or factor analysis procedures for identifying orthogonal factors in a data set may be used, most studies find no more than five to ten investment significant identifiable factors that are also useful for investment practice.
[12] Special cases may include proprietary trading desk strategies where the information level is maintained at a reasonable level and trading costs are nearly non-existent. Other cases, such as high frequency and algorithmic trading are arguably not investment strategies but very low level IC trading pattern recognition relative to highly sophisticated automated liquidity exchange intermediation.
[13] Trading costs and market volatility are additional considerations.
[14] Michaud (1989).
[15] It is worth noting that the Roll (1992) results assume unconstrained MV optimization.
[16] Even long-term academic studies remain susceptible to the unreliability of results of any back test.
[17] For example, Jobson and Korkie (1981).
[18] Note that the JK study applies equivalently to inequality unconstrained quadratic utility portfolio optimization, a framework widely used in financial theory and for the development of many investment strategies.
[19] An equal weighted portfolio is a simple way to compare the optimality of unconstrained optimized portfolios.
[20] 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 assets with any single-month returns greater than 50% or less than -50% in the twenty-year period. 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 experiment using the twenty-year history here. Readers wishing to replicate our experiment can access our data at www.newfrontieradvisors.com/research/data.
[21] 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 (squared) basis points, which is likely significant for most definitions of statistical significance. Eigenvalues in this context correspond to monthly variance. A monthly 10 squared basis point variance

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.

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

[23] 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.

[24] 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.

[25] 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 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.

[26] 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.

[27] See for example Grossman and Stiglitz (1988).

[28] Knight (1921) and Weisberg (2014).