Is Quant Fundamentally Flawed?[1]

Webinar Transcript (Edited)

April 16, 2020

Presented to:

CFA Society Boston

Gary Sarkissian, Moderator

Speakers:

Richard Michaud, David Esch, and Robert Michaud

New Frontier Advisors, LLC

[1] Michaud, R., D. Esch, R. Michaud, 2020. "Estimation Error and the Fundamental Law of Active Management: Is Quant Fundamentally Flawed?" *Journal of Investing*, June.

Precis

This document is an edited and annotated transcription of the webinar "Is Quant Fundamentally Flawed?" presented April 16, 2020 by Richard Michaud, David Esch, and Robert Michaud, New Frontier Advisors, with Gary Sarkissian, CFA Society Boston, as moderator.

The presentation is based on the publication: Michaud, R., D. Esch, R. Michaud, 2020. "Estimation Error and the Fundamental Law of Active Management: Is Quant Fundamentally Flawed?" *Journal of Investing*, June.

The narrative holds to the tone of the presentation as it was given as much as possible. However, it was necessary to refine and clarify a number of issues especially in the nine unrehearsed questions in Q&A. Some new material is included to complement important issues. Readers may find that the format provides a less formal and more accessible introduction to understanding of the issues in the published article.

Gary Sarkissian:

Thank you everyone for joining us for today's webinar – "Is Quant Fundamentally Flawed?" My name is Gary Sarkissian and I'm the director of Education and Content at CFA Society Boston. In a moment I'll be introducing you to today's speakers. But first I wanted to make some quick comments regarding the current environment as it relates to educational programming here at the Society.

First off, our thoughts and prayers go out to those who have been directly affected by the COVID- 19 disease as well as the medical professionals who tirelessly work on the frontlines of this epic battle. We hope all of you are staying safe and healthy.

Up until recently, live in-person programming was a primary delivery vehicle of education here at the Society. Now we're in a different environment and must adapt to virtual forms of communication for the time being. That said, we're hoping our technology will enable us to reach a much wider audience than we had with live programming.

As a case in point, today's presentation was originally organized as a live event in a venue that would have supported roughly 50 attendees. Now, thanks to the virtual format, we have almost 400 people registered to watch this livestream, some of whom are located outside the US. In addition, we welcome members of other CFA societies who join us on our live webinars. Registration to our webinars will be free to both members of CFA Society Boston as well as members of other CFA societies.

Finally, I can't emphasize enough that our members are one of the most important resources for the Society. We appreciate your continued feedback regarding our educational programs. We welcome new ideas from you and encourage you to maintain communication with your fellow members as well as society staff and volunteers via our online community, Connect.

Now I'd like to move on and introduce you to today's presentation.

Over the last few decades, quantitative investing has attracted both a sizable following among investment professionals as well as assets under management. In fact, according to hedge fund research in 2018, assets invested in quantitative hedge fund strategies approach the one trillion dollar mark. Furthermore, traditional active managers facing stiff fee competition and weak fund flows have also grown in this area and in some cases have replaced their human investment decision makers with automated rules-based trading.

Perhaps coincidentally, during the last 30 years, whether through textbooks, university courses, or written research, the academic world has also touted claims that the applications of the Grinold (1989) Fundamental Law of Active Management theory may add investment value to optimized strategies even when only a small informational advantage is available.

However, in a forthcoming paper in the *Journal of Investing*, our speakers present important new research on the likely limitations of various active quant investment strategies. They show that many popular strategies associated with Grinold's (1989) Theory are likely unreliable and self-defeating because of estimation error and constraints required in practical application. They further show that far more information than commonly assumed and appropriate portfolio structure are likely necessary for outperformance.

It is my honor to introduce our speakers today, Dr. Richard Michaud and Dr. David Esch of New Frontier Advisors. Dr. Michaud is the President and CEO of New Frontier. He has earned his Ph.D. in mathematics from Boston University and has taught investment management at Columbia University. He is the co-author of the book *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation* (1998, 2008). In addition, he has authored over 60 published journal articles manuscripts and white papers. He is co-holder of four US patents in portfolio optimization and asset management and a Graham and Dodd scroll winner for his work in portfolio optimization. He is a former editorial board member of the *Financial Analysts Journal* and associate editor of the *Journal of Investment Management* and former director of the Q group.

Dr. David Esch is the Managing Director of Research at New Frontier. Doctor Esch completed his PhD in statistics at Harvard University in 2004 and earned his master's degree in mathematics and statistics from Boston University. His specialties include mathematical statistics, numerical analysis, and computational Bayesian statistics and econometrics. He is the author of the article "Non-Normality Facts and Fallacies" which appeared in the *Journal of Investment Management* and was selected as one of the best papers of that journal for 2010. He is also co-author of many other peer-reviewed journal articles.

Now before I turn over the presentation to our speakers, please note that your lines will be muted for the duration of the presentation. At the end of the presentation we will be taking your questions for Q&A should you have any questions for the speakers. Please feel free to submit them at any time during the presentation using the chat feature for GoToWebinar. So without further ado, Dick please take it away.

Richard Michaud:

Thank you very much, Gary, for your kind words. We at New Frontier are very proud to be pioneering with the CFA Society Boston’s online presentations that are normally available for local members but also, in this case, nationally and beyond.

My purpose is to present the results of a new paper: “Estimation Error and the Fundamental Law of Active Management: Is Quant Fundamentally Flawed?" The paper is co-authored with David Esch and Robert Michaud and will be published this June in the Journal of Investing. It is currently available in preprint on ResearchGate.com and SSRN.com.

Introduction

Grinold's (1989) "Fundamental Law of Active Management" theory has been called the "canon of active management." It is often used as a framework for rationalizing and marketing various optimized investment strategies. Iconic examples of published applications of the theory include Grinold and Kahn (1995 and 1999) and Clarke, de Silva, and Thorley (2002 and 2006).

In their texts, Grinold and Kahn claim that adding stocks, factors, and/or trading frequently can often add investment value to an optimized or quantitative investment strategy even when a modest amount of information is available. Clarke, de Silva, and Thorley claim reducing constraints can often add investment value to an optimized investment strategy.

Grinold Theory has spawned an enormous body of published research rationalizing various institutional investment strategies for more than 25 years. These principles have been used to market many institutional investment strategies. A partial list is included on the slide and in the paper. These extensions include industry tutorials, university courses, academic and professional conferences, CFA Institute and CAIA group, and textbooks and publications rationalizing many contemporary investment strategies.

In this presentation we will show that many applications of the Grinold "Fundamental Law" Theory, including hedge funds, long-short, leveraged, and alternative strategies, are not useful as a framework for asset management in practice. We start with intuitive discussion followed by a comprehensive simulation study. We conclude that applications are likely unreliable and self-defeating. This is because Grinold theory ignores estimation error in optimized portfolios (Michaud 1989) and constraints required in practical application. Markowitz (2005) is a recent reference for the importance of constraints. The limitations of the "Law" may have been a very significant factor in the inability of many quantitative investment strategies to show superiority relative to passive management.

Grinold (1989) Formula

Index-relative mean-variance optimization is the framework of choice for designing active optimized strategies. Benchmarks arise naturally in active management for defining investment goals and for judging manager competence. Managers typically claim enhanced return relative to a given level of residual or index-tracking error risk. The information ratio (IR) is the objective of choice. It is the estimated return relative to a unit of residual risk. It is a very convenient framework for defining active investment management.

The Grinold formula describes a decomposition of the mean-variance optimization of the information ratio. The maximum of the information ratio is approximately the information correlation, or IC, between ex ante and ex post return times the square root of the breadth (BR), where breadth represents the number of independent sources of information in a sum-to-one or budget constrained optimization. The result is hardly controversial. Active management is a combination of skill and investment opportunities.

In the Grinold and Kahn text, they apply the Grinold formula to show that even if a manager may have a relatively small amount of IC or information level, the performance may be enhanced simply by increasing BR. They claim 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. That quote appears in both editions of their text. The implication is that increasing the size of the optimization universe may enhance performance independent of any need for additional information. These results have encouraged managers to optimize 500, or a thousand stocks or even 3,000 and 10,000 stocks or more. Also, managers have been counseled to add factors and trade frequently. The key assumption to applications of the theory is independent information and the IC constant for additional breadth.

Clarke, de Silva, Thorley (CST) (2002, 2006) is closely related. They generalize the Grinold formula to introduce what they call the “transfer coefficient" or TC. TC is a number between 0 and 1. It is a measure of how much information is not transferred into the optimized portfolio. A large transfer coefficient reflects a reduction of the information in alpha due to the presence of constraints in the optimization. They counsel reducing constraints to let the information flow or “transfer” directly to the optimized portfolio. This paper has been widely influential to promote quant, hedge funds, long-short, absolute return, and alternative fund strategies.

Road Map

Our roadmap will begin with intuitive discussion followed by Monte Carlo simulation optimization with estimation error. We consider optimization universe size as a function of equal, budget, and sign constrained portfolios. We will note differences between theory and practice. Our simulation results generalize classic estimation error studies by Jobson and Korkie (1981) and Frost and Savarino (1988). Our results help rationalize the empirical study by deMiguel et al (2007). We conclude applications of Grinold theory are likely unreliable and self-defeating in practice.

Roulette Rationale

The Grinold and Kahn claim of needing only a modest amount of skill to win as long as that skill is deployed frequently and across a large number of stocks is rationalized in terms of a roulette game. Roulette in a casino has a positive probability of the casino winning for each play of the game. The probability of winning is small but positive. The more plays lead to the likelihood of more wealth for the casino all else the same.

The Grinold and Kahn use of roulette for rationalizing applications of the law has serious limitations. The example goes to the heart of the limitations of the theory. The probability of the casino winning is known, positive, and constant. In contrast, the probability of an investment game-winning play is unknown, unstable, and may often be negative. Increasing investment plays may often be undesirable and the rationale invalid. Perhaps the illustration was not meant as substantive by the authors. We continue with more serious issues.

Practical Issues

The key issue in Grinold and Kahn applications is the assumption that skill level is constant. Consider an analyst covering 20 stocks now suddenly asked to cover 40. Consequently, the analyst’s overall skill or average IC is likely to decline. The example is not irrelevant. Average IC and universe size may often be negatively correlated in practice. The bigger the universe size, the smaller the average IC, all other things the same. Naively enlarging the optimization universe may often be self-defeating.

In a large stock universe index-relative optimization, an analyst will likely rely on multiple factor forecasts of index-relative return. Valuation of individual securities is likely to be impractical. Grinold and Kahn propose increasing the number of factors to add breadth. But finding forecast factors that are reasonably uncorrelated and significantly positive that satisfy the necessary condition of constant increase in breadth is no simple task.

Michaud (1999) is a limited but relevant study of factor investing in actual practice. Many factors that may have statistically significant relationships over long time periods may often prove insignificant or negative in a sequence of three to six yearly periods. Factors are typically chosen from a small number of categories such as value, momentum, quality, dividend yield. It is typically difficult to find time periods where factors are significantly positive and additive. Breadth related to the number of factors is often limited independent of stock universe size. Naively increasing factors in a forecast of index-relative return may reduce, not increase, performance. I knew of a manager in Boston who had 50 factors and wondered why his forecasts were not working well.

Grinold and Kahn also propose trading more often to add breadth. With the exception of high-frequency-trading strategies, which are pattern recognition based algorithms, investment strategies in general have a natural trade frequency. In the typical case, a book-to-price or earnings-to-price value manager is likely to be reluctant to trade much more often than a month or a quarter, because the information is not very fresh until then. Many do not trade more than once or twice a year. While a growth stock manager may trade more often, the number of trades in a year are often still relatively limited. More frequent trading than natural for a given strategy is often suboptimal or even infeasible. The upshot would likely reduce effectiveness while increasing trading costs. Trading frequency is in general limited by constraints relative to investment style mandates.

Clarke, de Silva, and Thorley have also been very influential proponents of the Grinold "Law." The transfer coefficient they introduce into the formula measures the loss of information transferred into the optimized portfolio due to the influence of optimization constraints. Optimized portfolios in practice typically include many linear constraints. Now, some constraints may be cosmetic for marketing purposes, but many are necessary in practice. As Markowitz (2005) notes, sign constraints limit liability risk and is often a regulatory requirement. Constraints may often manage optimizer instability, ambiguous solutions, and poorly diversified optimized portfolios. Even the largest financial institutions have leveraging limitations while performance mandates often impose risk constraints.

For many practical reasons, the Grinold and Kahn and Clarke et al prescriptions may not be operative. But they do not consist of a proof of the limitations of applications of Grinold theory in practice. The essential issue is how best to decide whether an optimization procedure is useful or not.

Back Tests

Back tests are typically used to demonstrate the value of a proposed investment strategy. In this procedure you use some historical data and impose the rules for the proposed strategy and statistically measure how it performed over the historical period. Another version of a back test is called a horse race. In this case there is a starting point and multiple strategies, and their performance is assessed over some period of time and a winning strategy determined.

Back tests are ubiquitous in asset management and financial theory. They have the virtue of practicality. They provide a procedure where no other is available. But they all have the problem of providing no reliable prospective information. The results are always period dependent. They can only show what has happened over some time period.

In my personal experience on Wall Street and in asset management, back tests are notorious for misleading investors. Such procedures often lead to loss of wealth, the loss of careers, and the loss of firms. The limitations of back tests were well understood in Knight (1921). As Paul Samuelson has said, a hundred years is just another hundred years. No one can say what's going to happen in any future period.

Simulation Test Framework

In this study, we turn to a mathematically rigorous testing framework – Monte Carlo simulation – for understanding the investment value of optimized investment strategies. In the following, we describe the Monte Carlo simulation test framework. I recommend that teachings in modern finance should become ubiquitous.

In a simulation test, the referee knows the true risk-returns for a given set of securities. The referee computes the Sharpe ratio (SR) of the maximum Sharpe ratio (MSR) optimal portfolio for the referee's data. In the simulation test, a player receives simulated returns for the referee's data. Note that simulated returns require an assumption of a multivariate return distribution. The most common assumption, and the one we shall use, is the multivariate normal distribution. But a simulation study could be based on other multivariate distributions. The player computes the means, variances, and correlations for the player's simulated returns and computes the MSR optimal portfolio for the simulated data. The player then reports the simulated MSR optimal portfolio back to the referee. The referee scores the true Sharpe ratio for the simulated MSR portfolio. The procedure is repeated many times. The referee scores and averages the SRs of the simulated MSR portfolios that can then be compared to other strategies. This is a very powerful method for understanding the performance of optimized portfolios.

Classic Simulation Tests

The classic simulation test of estimation error in budget-constrained mean-variance portfolio optimization is Jobson and Korkie (1981). I believe they were the first to provide a simulation test in modern finance. They Monte Carlo simulate five years of multivariate normal returns from historical data from a universe of 20 stocks. The process is repeated many times and the referee reports the average of the SRs for the strategy. The referee also scores the true Sharpe ratio for the data and the SR of an equal weighted portfolio.

The results for this famous study are as follows: MSR for the referee's data is 0.32. The Sharpe ratio for the equally weighted portfolio is 0.27. The average of the Sharpe ratios of the optimized portfolios is 0.08. The average Sharpe ratio of budget constrained optimized portfolios is 25% of the true value! Because the 0.08 value is an average, many of the optimized portfolios are literally buried in the X-axis. Moreover, the equally weighted portfolio is far closer to optimal than optimized. They conclude that budget constrained mean-variance optimization is not recommendable for practice.

This result is far more profound for modern finance than for simply how optimized portfolios work in practice with real data. Consider that the Sharpe (1964) paper, which is the foundation of the Capital Asset Pricing Model (CAPM), has been the dominant financial theory of modern finance for the last fifty years. The optimization in the CAPM is expected utility maximization of a budget-constrained mean-variance utility function. The result in the simple Jobson and Korkie study indicates that CAPM theory is based on an optimization process that is worse than equal weighting. While a small universe study, the Jobson and Korkie results will be seen shortly as robust.

The classic Frost and Savarino (1988) simulation study is also material to the prescriptions associated with Grinold theory applications. In their study, they compare sign and increasingly constrained optimized portfolios to no more than 5% in any security, no more than 4%, no more than 3%, all the way to equal weight for the 200 stock universe. In contrast to Clarke, deSilva, and Thorley, they find that the more restrictive the constraints, the larger the average Sharpe ratios up to a certain point. Their results are completely the reverse of Clarke et al. From an intuitive point their result is easily rationalized. Constraints are like Bayesian priors on portfolio structure that mitigate estimation error by forcing the simulations towards more likely optimal portfolios.

Jobson and Korkie and Frost and Savarino directly contradict the Clarke, de Silva, Thorley results for two very different sized stock universes. We will return to these results at the end of the paper.

Now I am very pleased to turn over the presentation to Dr. David Esch as he addresses in much more generality the optimization dimensionality issue.

David Esch:

Thank you and good afternoon. Welcome. I’m going to present some details of our simulation study which follows in the footsteps of the landmark Jobson and Korkie and Frost and Savarino studies and examines dimensionality issues related to optimization.

Introduction

So, what are the objectives of our study? We wish to study the impact of estimation error on several optimization procedures relevant to the Fundamental Law of course across a range of different breadths and different ICs. So specifically, we are interested following in the footsteps of Jobson and Korkie of examining situations under which optimization is a worthwhile pursuit. What we mean by that is that it increases the expected information ratio over a baseline such as an equal weighting procedure. Because if you can't beat equal weighting, then optimization has no investment value and is not advisable.

So, of course, we're interested in the functional relationship of breadth and information ratio. Is it a square root rule like in the Fundamental Law, or does the curve level off a little bit as breadth is added to the system? We're also interested in studying constraints because we have these two conflicting results from Clarke, de Silva, and Thorley who claimed that constraints damaged the average performance against the Jobson and Korkie results where constraints to a point were helpful to the out-of-sample performance in terms of the information ratio. And, of course, we're interested in studying these things across various levels of IC.

I'm going to say right now that the concept of information coefficient or IC is related to estimation error in a certain way. Because if you think of estimation error as noise, increasing the noise of the system has a tendency, all else equal, to reduce the information because it's lowering the signal-to-noise ratio in the system. So we're going to study the full matrix of different treatments experimentally and how they interact with each other.

Master Dataset

Our master dataset was taken from listed stocks on the New York Stock Exchange. We required a 20-year history of contiguous monthly returns spanning from 1994 to 2013. We found 544 suitable stocks. This may be optimistic in terms of restricting ourselves to these stocks. This may create a little bit of optimism in our results. And you'll see this is a theme of our experiment. We would rather be optimistic about all of these procedures than to have a critical reader be able to assign blame for underperformance to some other cause than just the differences among the different optimization procedures.

Covariance Matrix

So, what we're looking for out of our master data set is the referee's table of expected returns and a full rank, well-conditioned covariance matrix. To get a full rank well-conditioned covariance matrix we use a famous procedure, the Ledoit and Wolf (2004), which mixes together a factor model with the sample covariance optimally and produces a good estimate of covariance that is full rank and well-conditioned but also has reasonably well targeted and plausible estimates for our list of stocks. So we believe that our master table fulfills its mission in terms of being representative of what real equity optimizing looks like, as well as matching the theoretical considerations for studying the operational characteristics of the Fundamental Law. This data set is available online, and we will be publishing the URL for that in the paper. So anybody who's interested can play around with the data and perhaps conduct another experiment as a follow-up just to make sure we're on the right track.

Simulation Framework

So let me describe our simulation framework a little bit. This is a sort of recipe or a simple version of the algorithm that we followed for doing our experiment. Steps 1 to 3 on this page are each simulation, so steps 1 through 3 are repeated many times. The first step is to draw a list of 500 names from our master data set, and those are ordered so we create lists of, say, the first 5, the first 10, first 15, and steps of 5 up to 50 and then in steps of 50 up to 500, so that we have nested optimization universes where each size universe is contained inside the next larger universe but from a full list of 500 names.

We simulate some returns from a multivariate normal distribution, and so we pass that block of data along to each of the players in our simulation framework. The players represent 3 different optimization treatments. So trivially there's the equal weight player who just passes back 1 over the sample size for each of the sample sizes that I just detailed. Then there are two optimizing players. One of them is budget only constrained, so they're allowed to have long and short positions in their portfolios, and the other one is sign-constrained so that's long only, because we're interested in studying the impact of sign constraints on the referee-scored or out-of-sample performance of each of these players.

So we are adding assets without replacement that’s in sequence of nested groups that produces a nice increase in breadth, because we're adding stock-specific information as the size increases. So after the players all create their best portfolios maximum Sharpe ratio, at least what they think is the maximum Sharpe ratio portfolios, they pass them back to the referee, and they get scored with the true parameters under the simulation framework. So that it's an honest score and reflects the out-of-sample performance in the simulation framework.

We repeat this thousands of times and thousands of different permutations of all of the names that are in the master data set and we average across those, and then we can make a chart and plot them on a graph and compare how the procedures did.

Referee's Covariance Matrix

One hurdle in the study design for this experiment that we had to overcome was the covariance estimation problem. There you may have noted that our master dataset has two hundred and eighty-eight time periods but many more than that stocks and many of our simulation treatments for the players. The block of data has many more assets in the universe than time periods with which to estimate the covariance. So the sample covariance matrix is out, but fortunately that's not a problem since practitioners of equity optimization generally use factor models to estimate their covariances. However, there's a lot of disagreement about the right factors to use, the right numbers of factors, which models exactly, how to do that, and we wanted to avoid any possibility of blame for underperformance on the covariance estimation procedure because people that don't want to believe that one optimization method is doing worse than equal weighting or doing worse than another one. They are going to say, here's a possible cause for that underperformance. So we wanted to avoid that, so our workaround was just to give the players the referee's covariance matrices. What that means is that we may be overestimating performance. But you can think of our results as upper bounds. We tend to err on the side of optimism for the simulation experiment, because we don't want to be able to assign blame to other causes than just the treatment differences for each combination of breadth and constraint treatment.

No Estimation Error Case

We also present a fourth curve. So we have the equal weight, the two optimizations, and then we'll have the reference curve which is the no estimation error case. This is what the referee knows to be the maximum Sharpe ratio. So, Dick was saying before, there's an analogy of a roulette game where there's a known probability distribution of the outcomes. When you know the real return distribution for optimization and so the best strategy for that is well known. It's simply budget-constrained mean-variance optimization. This particular case is equivalent to the Grinold assumptions in the Fundamental Law, but it's also highly unrealistic in practice. It's unattainable in practice.

IC Levels

One thing that I haven't addressed is how we targeted particular IC in the experiments. So, as I mentioned before, the information coefficient is identified somewhat with estimation error itself. So by varying the amount of estimation error for each trial in the simulation experiment, we can vary the realized IC for that. And so the way we calibrated this was by taking the entire master dataset pool and simulating from that and computing an IC by computing the correlations with the referee and then averaging over many simulations of that to get a Monte Carlo estimate of IC by sample size for the dataset, and then we choose the number of simulation periods or number of time periods in each simulation that most closely matches the target information coefficient. So the target information coefficients were not attained exactly, but they're very closely approximated. They are generally within 1% even for the small optimization universes.

Speaking of averaging, I want to also mention that averaging over every permutation of the assets in the master list of stocks in our data set means that every asset has an equal chance of occurring first or last or not at all in every simulated data set. So some assets may have more information in their simulated data than others - a highly volatile asset may have very little information in its simulation, for example - but because we're averaging across every permutation of the data, we get a nice smooth linear model of additive breadth for our experiment. Some people may be concerned that you're adding more and more assets to the universe and that the breadth is not increasing as you're adding more assets, but for our experiment that is not the case. We have examined this very closely and we're quite confident that we have a smooth linear input of breadth for the experiment.

One final note from me before I turn the floor back over is that ICs greater than 10% are not formally consistent with the proof presented in the original Grinold paper. In the Grinold and Kahn text there's one step of the proof where there's a series that has to be approximated, and so they make that assumption, but for our experiment we found it extremely illuminating to examine ICs greater than 10% because we wanted to know just how much of a boost you need to give, how much more information or how much less estimation error is needed to boost the different optimizations, both with respect to equal weighting and with respect to the no estimation error case but also with respect to each other. So there is a crossover point where the budget-constrained optimization will actually outperform the sign-constrained optimization. We know there probably should be a cross over point if you add huge amounts of information to the system, but those may not be realistic for practice. We found it very interesting and illuminating to look at those things, and so this concludes my summary of some of the methods in our experiment. And with that, I am going to pass it back to Dick.

Richard:

Thank you very much, David, for this detailed presentation of our simulation test. So now let’s discuss the results of our tests.

Figure 1

Our simulation results are displayed in Figure 1. The Figure has three panels corresponding to three increasing levels of information or IC values: 0.1, 0.2, 0.3. The graphs display the average Sharpe ratio as a function of the size of the optimizer universe. The y-axis in each panel is the average Sharpe ratio for the particular optimization strategy as a function of the number of stocks in the optimizer universe, which in our case is the number of independent sources of information, or breadth. Each graph is an average of the Sharpe ratios for universes up to 500 assets without replacement each value from 16,000 samplings.

Results

The green curve (highest in the panel) is the Grinold theory case. It is entirely consistent with the roulette wheel game rationale in Grinold and Kahn. This is the no estimation error sum-to-one constrained mean-variance optimization case. What is displayed is exactly what is predicted by Grinold theory. As you increase the breadth of the strategy, it increases the amount of information in the average Sharpe ratio. The curve is a pure function of the number of independent sources of information or, in this case, the size of the optimization universe.

The cyan curve (lowest in the panel) is exactly the Grinold framework theory with estimation error. The simulation experiment captures the value of the optimized portfolios when estimation error in the mean is included. As is clear, the results are a total contradiction to Grinold and Kahn. The effect of additive stocks in the optimization has almost no effect of the average Sharpe ratio of the optimization as a function of the size of the optimization universe. The results represent complete contradiction to predictions of the theory for practice for the conditions in the theory.

The red curve represents the average Sharpe ratios of equal weighted portfolios. As the simulations show, an equally weighted portfolio is superior to the budget-constrained optimized portfolios for all values in the experiment. In particular, the results of the Jobson and Korkie small size optimization study are entirely consistent with those in Figure 1.

The purple curve represents a sign-constrained optimization. The results show that sign-constrained optimized portfolios are far superior to the budget-constrained solutions for the entire length of the portfolios in the experiment. The results also confirm the results in Frost and Savarino of superiority of sign to equal weight for sufficiently large optimization universes.

Information Level

A recent paper in the *Financial Analysts Journal* (Allen et al 2019) asks the question: What if we can forecast? Is the problem of estimation error in optimized portfolios the result of too little information? If we have enough information, will estimation error become moot?

In Panels 2 and 3 we consider the cases where the IC is 0.2 and 0.3, information levels usually considered not sustainable in investment practice. In the case for an IC of 0.2, the budget- and sign-constrained curves are closer than in Panel 1. However, the budget-constrained case is inferior to equal weighting for all stocks in the optimization universe and the sign-constrained solution is very superior.

In Panel 3, where IC is 0.3, the budget-constrained optimization succeeds at being superior to equal weighting, but only for part of the spectrum of number of stocks and barely reaches parity with sign-constrained at the very end of the optimization universe size. Even for the limitations of the experiment that considered only optimization universes of up to 500 stocks, the Grinold theory fails spectacularly. Moreover, a sustainable IC of 0.3 hardly seems likely in practice.

Larger Universe

Given that the simulations of the budget-constrained optimizations slowly rise as a function of the size of the universe, it may provide some the hope that the Grinold and Kahn theory could be rescued with many more stocks or many more sources of independent information. But this is unlikely.

The slowly rising level of budget-constrained simulations is a necessary artifact for the simulation framework. Each point on the curve represents the average of thousands of simulations without replacement. So every additional stock or source of independent information, by definition of the simulation framework, adds some value. In addition, the simulations are based on the assumption that the covariance matrix is known with certainty. But a risk model that is perfect is unknown in practice. Indeed, there are many controversies on how to build a risk model suitable for asset management in practice. The upshot is that our results are a very significant upper bound of the true value of optimized investment strategies in practice. In addition, we did not include any cost of managing larger stock universes. Larger universes are likely to add additional costs in practice.

CAPM Theory

We now consider the implications of our results from a very different perspective. Grinold theory is an application of CAPM theory. Sharpe (1964) CAPM theory is based on a budget-constrained mean-variance optimization of expected utility of the mean and variance. The framework is identical to that used in Grinold theory. It is also exactly the optimization framework in the Jobson and Korkie simulation study and in the results we present in Figure 1. This implies that CAPM, the dominant theory of modern finance for the last fifty years, has been based on an optimization framework that is worse than equal weighting out-of-sample under standard assumptions. As a corollary, it also implies that the framework for quantitative asset management largely based on CAPM is as likely as Grinold theory to be fundamentally flawed as a framework for asset management technology.

This result indeed demonstrates the enormous power of simulation methods for 21st century finance. It also demonstrates the incredible neglect of the importance of the work of Jobson and Korkie and Frost and Savarino by the investment and academic community. It is not just Grinold theory that hampered the value of institutional quantitative management in practice.

Summary

We do not contradict the simple intuition that more information and more investment opportunities are useful in active management. However, what we show is that it takes considerably more than a modest amount of skill to win the investment game however frequently the skill is deployed or however many stocks or forecast factors that are being used. Under the highly idealized conditions of our simulation studies, the results dramatically contradict Grinold theory and practice prescriptions. Our simulation studies very nicely generalize the classical simulation studies in the literature Jobson and Korkie and Frost and Savarino, as well as rationalizing the 1/N empirical results in deMiguel et al.

Implications

For many years, the limitations of the law may have negatively impacted many professionally managed quantitatively managed optimized investment strategies and may have been a significant factor in the lack of persuasive evidence of superiority of active quantitative management versus passive. Active management can beat passive, but active done wrong can't for sure. The law may have often encouraged investing in unproductive minimally constrained and overly risky leverage strategies.

The Fundamental Error

The fundamental error is the misunderstanding of the importance of estimation error in optimized investment strategies, as well as the nature of investment uncertainty. The ghost of Knightian (1921) uncertainty emerges. Investment management is no simple roulette game. Monte Carlo simulation studies rightly should be the framework of choice for understanding the value of optimized quant strategies and practice for the future. And probability is no simple summary of uncertainty.

The Reliable Conditions

The reliable conditions for winning the investment game include high-quality investable assets, investment-significant information, economically relevant constraints, and properly implemented estimation error sensitive portfolio optimization technology.

So at this point I will send it back to Gary, and hopefully we will be able to answer some of your questions. Thank you very much.

Gary:

Well thank you, Dick, thank you gentlemen for that excellent presentation. We're going to move into Q&A. Now just as a reminder for everyone who's listening in, on the right hand side or wherever on your screen there is a control panel for GoToWebinar and there's a chat section and in that chat box you can type in your questions so that we can present them to our speakers.

So we have a few that came through during the presentation and the first question here, it seems more of a comment, so we'd like to kind of hear everyone's response on the panel regarding this statement: this is a nice extension of the Michaud’s 1998 work that illuminated the hypersensitivity of quadratic programming or just another term for mean-variance optimization to estimation errors. However, it may not be fair to Grinold and Kahn to imply that simply adding securities or factors results in superior portfolios. I doubt they ever said or believed that- what is your response to that comment?

Richard:

Well, they did state their belief clearly and repeatedly. This is a quote: "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." I made a special effort to be certain that both editions of their texts had the same quote. Now, you need to caveat that it's conditional that the IC is constant as you increase the number of stocks. But our simulations also make that assumption and, in practice, as we observed earlier, it is not that easy to keep the IC constant as the size of the universe increases.

David:

So they basically try to justify very large optimization universes, and that has happened to be true for many quant strategies. And the problem is that you really can't do that, the portfolios are far too sensitive to estimation error and the end result is that you can't use sum-to-one optimization that would allow you to think that, for example, a long short strategy might be better than a sign-only constraint strategy. So I think our simulation studies try to make the case as clearly as possible. And keep in mind that we're trying to make the solution that we describe here as much as possible tilted towards proving the hypothesis, and of course we're assuming a covariance matrix that's perfect, which, of course, makes no sense in practice.

I would add, also, that Grinold and Kahn present the Fundamental Law as a mathematical proof, and we don't dispute any of that proof at all. It's a completely valid mathematical proof, but it's just based on assumptions which are highly unrealistic in practice. So our study is not really contradicting the Fundamental Law. If you can manage to find a perfect frictionless environment and have perfect information, then go ahead and add as much breadth as you can. But that's just not realistic in practice, and so our study is meant to simulate a more realistic environment, and we have bent over backwards to be fair to all of the methods here by giving them the benefit of the doubt and giving them a perfect covariance matrix, as Dick just said. So I think that we have made every effort to be fair, but to get to the spirit of the question, I think a lot of people do understand that with IC and breath you can't just keep adding assets forever and assume that it's all constant.

But we're thinking about the prescriptions. What is the applicability of this in terms of actually changing your investment practice? So I think we've already said that most active managers are already investing in as many stocks or factors as they have good information for. So once you're at that point you can't really just keep adding more or keep looking for more places to add more without thinking about how much you're decreasing your information, which I think people are cognizant of, but then adding that extra complexity compounds with the estimation error and very often is going to make you significantly worse off I think is one of the cautionary things we're trying to say.

Gary:

I was wondering if I could follow on that point. You know, when you look at the Grinold and Kahn theory, it's sort of appealing to a smaller size investment manager that lacks the analytical resources and tools that a larger asset manager would have. But based on your conclusions, it seems that it's the opposite, that if you're a smaller manager and you have less informational advantage, simply adding more securities to the portfolio and widening out your universe is not going to help you out. So, if you're a smaller manager or listening to this, what is your approach based off of this conclusion?

Richard:

One of the things that has always impressed me when I was an institutional quantitative analyst on Wall Street, was that stock pickers and traditional managers usually have a relatively small stock universe. When I was writing my 1989 paper, I wondered why they didn't use an optimizer. An optimizer gives you, if you believe the theory, a better way of managing your money and a better way of giving value to your investors. But traditional managers often include no more than 20 or 30 stocks in their portfolios. They don't follow 50 or 200. Just a small group of stocks. And it impressed me that what they were doing was exactly the right thing. Intuitively, they were keeping a constant level of information for every security. They probably should have optimized as long as they had an optimizer they felt worked for them.

So if you're trying to solve problems for much larger universes or for much larger number of factors, the estimation error accumulates dramatically and that's what you see here in our results. I think this should be convincing evidence that the prescriptions in Grinold and Kahn and Clarke, de Silva, and Thorley and associated literature should be suspect if not contradicted by our results. The inverse of what they have been trying to propose is probably much closer to the right answer. You just need to have good investable securities, you need to optimize very carefully trying to avoid all the mistakes that you can make with an optimizer.

Gary:

Great. So, we have a question regarding the simulation results on Figure 1 on slide page 8. The question is (the person is referring to the green line): Shouldn't increasing the IC lead to higher Sharpe ratios all else equal? Actually, they probably are referring to all green lines but if you're increasing the information coefficient, all else equal, shouldn't that help increase your average Sharpe ratio?

Richard:

The IC level in the green curve is assumed to be 1. The green curve has the referee's information level. So the green curves are the same no matter what the IC level for the panel. The three panels describe the effect of three different levels of information on the budget-constrained and signed-constrained optimizations. Note that the equal weight curve does not change from one panel to the other. And the green curve is simply a reference indicating what zero estimation error, the assumption in Grinold theory, implies. So both the green Grinold theory curve and the red equal weight curve are the same from one panel to the other. They represent references relative to budget and sign constrained optimizations. Good luck, if you find an IC-equals-one manager, let me know. I'm happy to invest if you can find one close to an IC of one.

Gary:

That's actually a good transition to the next question. So, an IC of 0.3 seems pretty high according to this person. Do you have an estimate of what IC is required to make active management desirable? What positive message do you have for active managers?

Richard:

Well, I tried to make the case that there's no free lunch in finance. You should be suspicious that it's not just a matter of adding more stocks or adding more factors that's going to add to the performance of the portfolio. That seems in hindsight very naïve. I think we need to be more than just good mathematicians. We need to be good investment professionals as well. Investors need to be reminded there's no magic in quant. You need high-quality investible assets, you need investment-significant information, you need meaningful constraints and you need an optimizer that's going to be sensitive to estimation error.

Robert Michaud:

I think one more positive message may be that constraints, instead of thinking of them as holding you back, are one place where you can think of them as an additional way of adding value. That's very intuitive for a lot of portfolio managers to add information and actually improve performance going forward by using relevant constraints for your portfolio.

Gary:

That's actually an interesting point Robert, and just wanted to add on that. So, you know, you'd mentioned in the paper and in this presentation that equal weighting a portfolio eliminates or reduces some of that noise and helps reduce that estimation error. I think the way you described it is that you use no wrong information to distinguish among assets. How, with that framework in mind of portfolio construction, would this apply to a larger sized asset manager that is going to deal with universal securities that have varying degrees of liquidity?

Richard:

My answer is that equal weighting only works for relatively small stock universes. If you find equal weight better, you are probably doing something you shouldn't have been doing. That's basically what Jobson and Korkie was saying.

What I said in my 1989 paper is that optimizers are error maximizers. By the way, this phrase is very popular among academics. It may be what they write on my tombstone. It tends to be what every academic references whenever they talk about optimization. But we have shown that the sign-constrained solution is better than equally weighted. Equal-weighted is just the benchmark. I think David mentioned in his remarks that if an optimizer can't beat equal weighting, why do you want that optimizer? But unfortunately, that's the optimizer that people have been using at some of the most sophisticated asset managers in the world. A good rationale for why active hasn't beaten passive in many cases.

Gary:

All right. So just as a reminder for the audience, if you have any questions you'd like to submit just remember to use the chat feature in the control pane for GoToWebinar. We've got time I think for another question here. So there's a lot to read through with this question, but basically the question is: this study has enormous implications, as you had mentioned as a matter of fact, and this is just my comments on this you know. You mentioned that a substantial portion of the asset management industry has been built on this academic framework. So now, based off of the conclusions in this study, what do you think is next?

Richard:

Well, I hope I'm not sounding like a broken record. There's no simple little trick. It seemed like a great solution. You didn't need to know very much information. All you had to do was keep adding stocks or breadth in some fashion. And it was presented as a very sophisticated theory. Managers would go to the big pension plans with this very sophisticated approach and this wonderful rationale for adding value to an actively managed portfolio. But it was based on the assumption that the IC is constant as you increase the size of the universe. Estimation error accumulates as the size of the universe increases, all things the same. The size of an optimization universe and information level are essentially negatively correlated. But influenced by the theory, managers were using very large stock universes.

Managers were good mathematicians using CAPM to justify the kind of things that they were doing, and that's of course what Grinold theory is. Grinold is just CAPM with a very clever proof. Nevertheless, it's a framework that's not taking into account estimation error. The Markowitz optimizer is theoretically correct, but data used to estimate inputs have estimation error. Markowitz requires statistical estimation. How you manage estimation error is to use Monte Carlo simulation which requires that you assume some kind of multivariate normal distribution.

So Markowitz optimization is not what we thought it was. As a pure mathematical algorithm, it is not useful for asset management in practice. In order to be effective, it must convert to a statistical estimation procedure that requires you to pay strict attention to how much information you have. The optimizer is a dumb beast. It will do whatever you want it to do. It needs to be sensitive the level of statistical information in the inputs.

Gary:

Ok. So I think we've got some time to squeeze another question or two. It's a question from another person who is asking does this mean that active managers should be more fundamental than quantitative?

Richard:

I think the answer is probably. Managers should have a much wider net in terms of the things that they care about. Some quantitative managers and academics have a narrow view of information they care about. I was at a conference and the speaker was using various time series of stock market data and complaining that there was not very much information. So I raised my hand and I said you know there are other sources of information that might be useful for managing a portfolio. They asked like what? I answered what about economic data? Does it matter what the interest rates levels are, does it matter whether GDP is going up, down, or sideways, does state of domestic or global economy matter? There is much hope for quantitative active management. But you have to do it thoughtfully. You have to go beyond simple models. You need to better understand your data and to use it well.

Gary:

Okay. Great. We'll try to squeeze in another one here. So I think you had maybe touched on this earlier, but this person's asking why is equal weighting rather than cap weighting used as the naive alternative? Where would cap weighting fall on the chart, and I think they're probably referring to Figure 1 on slide page 8.

Richard:

Our study is basically whether the optimization proposed in Grinold theory - and serves as the basis of applications of the formula - beats equal weighting. If not, what possible interest could it have? But it is possible to have an optimizer do better than equal weighting. For example, a sign-constrained solution turns out to be better than equal weighting for a reasonable amount of information.

The Fundamental Law came from the optimization of the information ratio for a given index. It considers index-relative optimization whatever the index. It could be any reasonable benchmark including capitalization weighting. There is nothing in our research that excludes capitalization-weighted indices. The only issue is that, whatever the benchmark, you optimize the information ratio and the question is how to structure the optimization. I think there really is the possibility of doing so much better than what we have been doing. But one thing for sure, avoid making mistakes.

David:

I'd like to add also that cap weighting would create a lot of challenges in the context of a simulation experiment, because in the experimental design the optimizers are receiving simulated data, and cap weights are something that change over time. With a different set of returns they would be different for every circumstance. We wanted to avoid anything that might create any doubts about our experimental design and open ourselves up to critiques about our experiment. But we really aren't saying anything negative about cap weighting. It's also another completely legitimate benchmark to use for performance.

Richard:

To repeat, the index is independent of our results. We don't have a conclusion whether or not to use capitalization-weighted indices. Managers have many different benchmarks. The optimization in the theory is based on the information ratio that can be used for any well-defined benchmark. We're focused on the effectiveness of the mathematics of the optimization of the information ratio for structuring the portfolio.

Gary:

All right. We're running really short on time so we're just going to take one more question here, and the question is: do machine learning techniques help to select more relevant factors?

Richard:

David you have thought a bit about artificial intelligence.

David:

If you have a good set of training data, machine learning techniques can do almost anything, but it's all dependent on what data you train the machine learning on. If you're training any kind of model - I mean machine learning is just basically a very fancy statistical model - and if you fit any model with garbage data then you'll have garbage results, and if you fit it with highly relevant data to your investment circumstances you might obtain very good results. But it's hard to know.

You know regimes in capital markets can change very quickly, and you don't know if the future is going to match the past. So, I think it's the same set of risks as having a human brain try to do the analysis. It all depends on the circumstance and the training data that you're using for the machine learning.

Richard:

AI training is like a backtest or a horse race. It works for some time period, but what happens with a different time period? The problem is that these procedures don't have prospective information. The right way to think about such issues goes back again to Frank Knight.

Gary:

All right. Well this has definitely been a profound presentation and research. Just to clarify this report will be appearing in the *Journal of Investing* in June. But my understanding is anyone that's interested in reading this as a white paper can access it via SSRN and ResarchGate and via New Frontier's website. Gentlemen, any final words or comments before we wrap it up?

Richard:

Thank you, Gary, for hosting us here, and again, going back to my original comments, we are very pleased to be supporting CFA Society Boston. And thanks, of course, to the Institute as well and the New York Society and many others that are doing such great work in trying to promote good investment practice.

Gary:

Well I appreciate those comments, and thank you everyone for joining us today. Remember to check our website www.cfaboston.org to see what upcoming events we have. Until then, please be well and enjoy the rest of your week.

All:

Thank you.