When Michaud Optimization Fails

By

Richard O. Michaud and David N. Esch[1]
New Frontier Advisors
Boston, MA 02110

Draft: December 2017

Abstract

The Markowitz and Usmen (MU) (2003) simulation study reported Michaud (1998) mean-variance (MV) portfolio optimization superior to Markowitz (1952, 1959) out-of-sample on average in each of thirty cases examined. However, a simplified replication of the MU test found thirty percent failures of Michaud relative to Markowitz. Instances of Michaud failures were associated with asset risk and return characteristics inconsistent with diversified portfolio risk management. Risk-return properties in professional asset allocations and large universe portfolio optimizations may often be similarly perverse. The simulation framework in Michaud (1998) can be a valuable diagnostic for risk-return estimate diversification perversity when appropriately applied. Our results underscore the necessity of investment sense oversight and validation for successful application of quantitative methods.

For more than sixty years, Markowitz (1952, 1959) mean-variance (MV) portfolio optimization has been the theoretical standard for defining portfolio optimality in modern finance. However, Michaud (1989) noted that Markowitz optimized portfolios in practice are unstable and ambiguous. Michaud (1998) introduced Monte Carlo resampling to address estimation uncertainty in MV portfolio optimization. Simulation studies in Michaud (1998, 2008a, b) showed that Michaud optimized portfolios are often superior to Markowitz on average out-of-sample. The Michaud procedure produces more diversified better risk managed portfolios with stable, often investment intuitive, allocations across the efficient risk spectrum.

The Markowitz and Usmen (MU) (2003) simulation test of Michaud versus Markowitz optimization found Michaud superior out-of-sample on average in each of thirty cases. However, in spite of the authority of MU, the superiority of Michaud optimized portfolios relative to Markowitz is not without controversy.[2] We present a simplified replication of the MU study and find approximately thirty percent failures of Michaud relative to Markowitz. Instances of Michaud failures were associated with simulated risk-return properties inconsistent with diversified portfolio risk management. Risk-return characteristics for portfolios of real market securities may often be similarly perverse, both in asset allocation and large-universe optimizations. A simulation framework as in Michaud (1998) can be useful for identifying perverse risk-return estimates when customized to investor mandates. Our study indicates that quantitative methods generally require investment sense financial intermediation for successful implementation.

The structure of this paper is as follows: Section 1 reviews the optimization simulation test framework and gives an example with Michaud versus Markowitz MV optimization. Section 2 describes the MU simulation test of Michaud versus Markowitz and presents the results of a replication that contradicts their findings. Section 3 examines instances of Michaud failures and presents a methodology for identifying diversification perverse risk-return estimates. Section 4 provides a summary and conclusions.

1.0 Optimization Simulation Tests

In an optimization simulation study, the referee is assumed to know the true means, standard deviations, and correlations of K assets. The steps of the Monte Carlo simulation procedure are: 1) Simulate T multivariate returns N times; 2) Compute estimated means, standard deviations and correlations from the N simulations of T multivariate returns; 3) Compute Markowitz and Michaud MV in-sample efficient frontiers for the simulated risk-return estimates; 4) Compute the out-of-sample Markowitz and Michaud MV frontier risk and return scores from the referee's true risk-return estimates. Repeat steps 1 to 4 many times. Compute average performance for the two sets of efficient frontiers.[3] Display the in-sample and out-of-sample efficient frontiers.

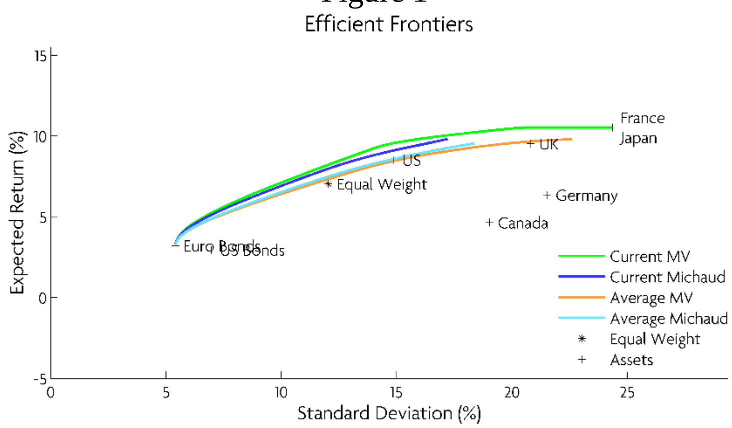
1.1 Simulation Test Example

Figure 1 provides an example of an optimization simulation test of the in-sample and out-of-sample Markowitz and Michaud MV efficient frontiers. It is based on the historical returns of eighteen years of monthly data from 1978-1995 as described in Michaud (1998) of eight asset classes consisting of two bond indices – U.S. and Euro bonds, and six country equity indices – U.S., U.K. France, Germany, Canada, Japan.[4] The display shows the in-sample Markowitz and Michaud MV efficient frontiers based on the historical data and average of the out-of-sample

Markowitz and Michaud frontiers. While the Markowitz frontier is superior in-sample for given level of risk, out-of-sample the Michaud frontier is superior.

Note that Michaud optimization does not increase average return out-of-sample relative to Markowitz. Michaud optimization represents a near rigid left shift of the average out-of-sample Michaud frontier relative to the average out-of-sample Markowitz, with less risk for a similar level of out-of-sample average return.

Figure 1



In Figure 1, the "Average MV" (Markowitz MV) and "Average Michaud" (Michaud MV) frontiers, represent the average means and square roots of average variances for the two players. The players can be scored visually by examining the relative positions of the two curves. Figure 1 shows the in-sample the superiority of Markowitz (green) versus Michaud (dark blue) frontiers. However, out-of-sample, the Michaud frontier (cyan) is above the Markowitz (orange) frontier for the length of available levels of resampled risk. In effect, the Michaud procedure pushes the out-of-sample frontier to the left of the out-of-sample Markowitz frontier. While the Michaud procedure cannot enhance return out-of-sample, it reduces risk relative to a given level of available return out-of-sample, a desirable property for risk-managed investment strategies in practice.

2.0 Markowitz and Usmen (2003) (MU) Experiment

The MU optimization simulation experiment is a more comprehensive test framework of out-of-sample performance of Michaud versus Markowitz MV optimization. The MU study is based on the same eight asset class historical return data from Michaud (1998) as in Figure 1. In the MU experiment, Monte Carlo simulation is applied to the "uber" historical risk-return data to spawn ten sets of risk-return pairs, each to be used as a referee truth. These ten truths are used as parameters in Step 1 of the simulation experiment framework as described in the prior section. They are also used to compute out-of-sample performances of the two optimization procedures in Step 4. More referee truths are intended to avoid dependence of the results on any particular set of risks and returns by testing over a dispersed region of possible inputs.

In the MU framework, a Bayesian procedure is used to modify the simulated risk-return optimization inputs for computing simulated Markowitz frontiers. Also, MU use three utility functions – low, medium, and high risk aversion – representing a range of investor risk preferences to compute the average out-of-sample performance of the simulated frontiers. The MU study reports that Michaud outperformed Markowitz in all thirty cases examined.

2.1 Replication of the Markowitz-Usmen Experiment

In our replication of the MU framework we make three changes: 1) We Monte Carlo simulate twenty, rather than ten, referee truths starting from the same uber eight asset historical risk-return data set; 2) For simplicity, we dispense with the MU Bayes procedure;[5] 3) We use the frontier-rank procedure in Michaud (1998) instead of three utility functions for averaging the simulated frontiers to compare out-of-sample performance as in Figure 1.[6]

2.2 Replication Results

The results of all twenty referee truths of the MU study, presented in the graphical display format of Figure 1, are available online.[7] There are thirteen referee truths that were mostly wins for Michaud while six were clear wins for Markowitz, and one indeterminate, with the lower part of the frontier winning for Markowitz and the upper part winning for Michaud. Tests are numbered according to Referee truth for reference. Trials numbered 2, 5, 9, 11, 12, and 13 are definite failures of Michaud relative to Markowitz; trial 16 is indeterminate, and the rest are wins for Michaud. Due to space limitations and the repetitive character of the results we examine two archetypal examples of the Michaud failures here. Readers are invited to review all cases online.

Figure 2 displays the results of the simulation experiment for referee truth 2. The left-hand side panel displays the in-sample and average out-of-sample efficient frontiers in the experiment in the format of Figure 1. The right-hand side panel in Figure 2 includes composition maps associated with in-sample Markowitz (top) and Michaud (bottom) efficient frontiers. Composition maps show in color-coded format the optimized allocations of the Markowitz and Michaud in-sample efficient frontiers from the left or minimum variance portfolio to the right or maximum return efficient frontier portfolio. Composition maps provide a valuable analysis of efficient frontier portfolios.

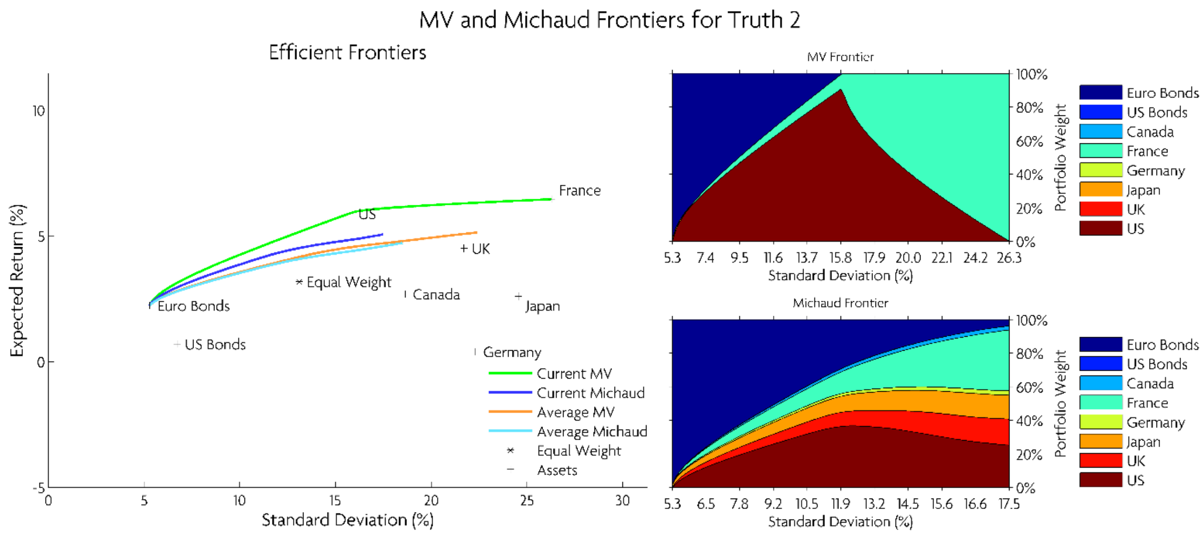
2.3 Analysis of Figure 2

The left-hand panel in Figure 2 demonstrates that the Michaud out-of-sample frontier underperforms Markowitz for essentially the length of the frontiers. Examination of the associated composition maps shows why Michaud failed out-of-sample. The in-sample Markowitz frontier (top-right panel) shows there are two dominant assets: France and US. In contrast, the Michaud frontier (bottom-right panel) displays a far more diversified portfolio across the risk spectrum. The Michaud frontiers computed from Monte Carlo simulated histories also show the same added diversification relative to their Markowitz counterparts. Referring back to the left-hand display, France has the highest true return while the US has almost the same return with much less risk for truth 2. The remaining six assets have far less return for similar levels of risk. In simple parlance, the Michaud frontier has too much diversification relative to the truth 2 risk-return characteristics. Michaud failures for other cases are similar variations on the theme that diversification for a particular set of risk-return characteristics may not be beneficial.

2.4 Analysis of Figure 3

The results in Figure 3, referee truth 9, provide a somewhat different example of Michaud failure relative to Markowitz. Examination shows that Michaud may underperform in some parts of the efficient frontier risk spectrum and outperform in others. Note that US is a low return asset in the middle of the risk spectrum pulling down estimated portfolio return for the more diversified Michaud portfolio. In this example as in others, the Michaud frontier may be inferior on average out-of-sample relative to Markowitz when diversification fails to be beneficial for given risk-return properties.[8]

Figure 2

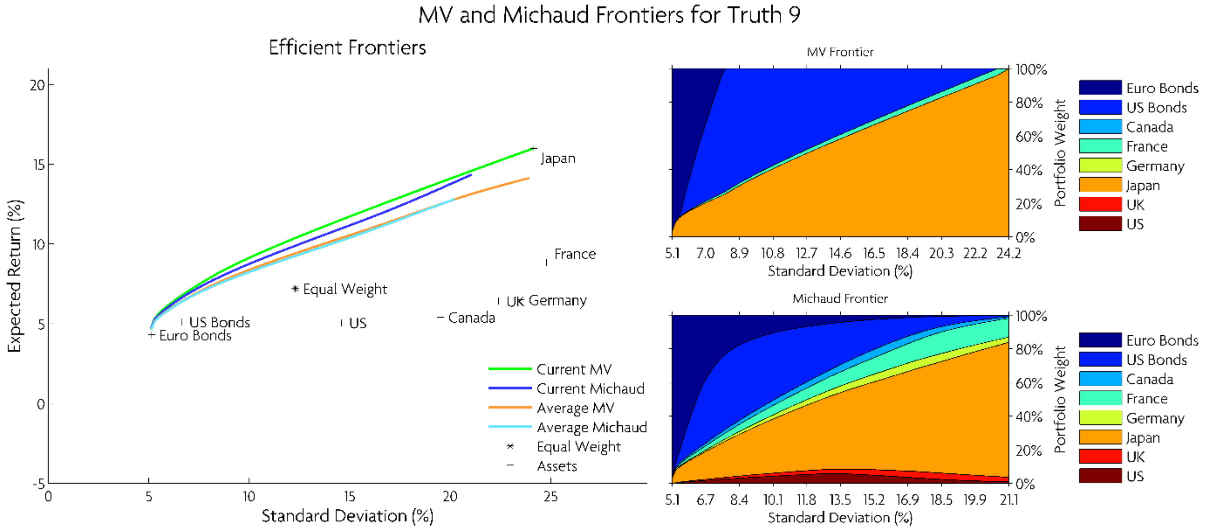


3.0 Discussion and Resolution

Our simulated experiments show that Michaud MV optimization fails relative to Markowitz on average out-of-sample when diversification is not beneficial relative to risk-return properties of assets used in the optimization program. In some cases, as in Figure 2, the risks and returns are inconsistent with a diversified investment program over the entire length of the efficient risk spectrum. In other cases, as in Figure 3, risks and returns may be inconsistent with diversification over some sections of efficient frontier risk.

Our results are not confined to the MU framework of simulated risk-return estimates. Risk-return estimates from historical data may often exhibit Figure 2 or 9 characteristics. More generally, investors may often be unaware of the underlying perverse diversification characteristics of their optimization universes. A major concern of asset management in practice should be the effort to avoid perverse cases, which typically manifest as assets for which exposure will harm performance, thereby making diversification within the investment universe undesirable.

Figure 3



In the simple case of the eight well-known asset classes in Figure 2, the risk-return estimates for France and US clearly dominate the others. An experienced investor will likely question why the six inferior return assets belong in a procedure designed to provide a well-diversified investment program. Similarly, in Figure 3, the middle part of the efficient frontier contains inferior return assets. Such assets are likely candidates for review in a procedure with the objective of defining an optimal diversification program. In practice, however, professional asset allocation strategies often include many lesser known securities and large scale portfolio optimizations may include literally thousands of securities. The challenge of identifying perverse risk-return estimates is magnified many times and should be a major concern.

3.1. A Novel Proposal

The simulation framework in Figure 1 can be a very useful diagnostic for identifying perverse asset risk-returns for constructing a well-diversified optimized portfolio.[9] Michaud failure relative to Markowitz as in Figures 2 and 3 may be indicative of risk-return estimates inconsistent with a well-defined diversification program. However, implementation requires a number of further considerations.

3.2 Some Properties of Michaud Optimization

In this report, the number of simulation periods used to estimate level of information associated with computed asset risk-returns was the number of monthly periods in the historical data, in this case 216 months. In practice, reliable historical monthly data may not exist. More generally, the amount of information associated with estimated risk-returns appropriate for a particular investment strategy will seldom be as simple as counting the number of historical periods in a data set.

Michaud optimization requires the analyst to input a "forecast certainty" parameter that represents the analyst's confidence level in the information in the optimization inputs. The parameter represents the range from certainty to complete uncertainty. Such a parameter can be defined many ways.[10]

A second important caveat is in order. A simulation framework such as Figure 1 for identifying diversification perverse assets has no referee truth available. An analyst may have a high, though not Markowitz, level of confidence in risk-return estimates but be confronted out-of-sample with a high volatility market, and conversely. There is no fail-safe process and performance clearly benefits from the analyst's quantitative investment expertise.

4.0 Summary and Conclusions

We review optimization simulation technology as in Michaud (1998, Ch. 6) and in Markowitz and Usmen (2003) (MU) for demonstrating average out-of-sample benefits of Michaud versus Markowitz MV optimization. We introduce a simplified replication of the MU simulation test framework and find, in contradiction to the results reported in MU, that Michaud failed relative to Markowitz in about thirty percent of cases examined.

Detailed examinations of cases where Michaud failed relative to Markowitz were found to be examples of risk-return characteristics inconsistent with a well-diversified investment program. The problem of diversification-perverse risk-return properties can easily occur with historically estimated data and is likely common in practice in asset allocation strategies and large universe portfolio optimizations. We note that the Figure 1 simulation framework may often be a useful diagnostic tool for identifying perverse risk-return estimates. However, important caveats for application include the absence of a referee truth and inappropriate Michaud MV optimization certainty conditioning.

The rhetorical question of the title of the paper can now be answered. Our results show the importance of financial sense intermediation for successful implementation of quantitative portfolio management. While Michaud optimization always provides better diversified portfolios relative to classical, the best performing portfolio on average out-of-sample is not always one that is well-diversified. Failure of the Michaud optimizer is not the fault of the algorithm but the consequence of risk-return characteristics inconsistent with the objective of a well-diversified risk-managed investment program.

Perverse optimization universes may often be appropriate for many active investment strategies where risk-management may limit performance. However, as Knight (1921) reminds us, uncertainty is endemic in investment practice.[11] Michaud MV optimization is simply a generalization of Markowitz optimization that allows the investor to control level of uncertainty in risk-return estimates in the optimization process.[12] Consequently, Michaud superiority in practice is not decidable since true return distributions are uncertain. Michaud fails when more diversification is not beneficial relative to an investor's outlook and/or subsequent market performance. Effective diversification and risk management are neither more nor less than an axiom of professional management often considered desirable for meeting many long-term investment objectives.

References (incomplete)

Harvey, C., R. Liechty, R. Liechty 2008a. "Bayes vs. Resampled efficiency: A Rematch." *Journal of Investment Management* 6(1):29-45.

Harvey, C., R. Liechty, R. Liechty 2008b. "A Response to Richard Michaud and Robert Michaud: Letter to the Editor." *Journal Of Investment Management*. 6(3): 114.

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. Boston: Houghton-Mifflin.

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. and N. Usmen, 2003. "Diffuse Priors vs. Resampled Frontiers: An Experiment." *Journal Of Investment Management* 1(4): 9-25.

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

Michaud, R., 1998. *Efficient Asset Management*. New York: Oxford University Press, 2001. First published by Harvard Business School Press.

Michaud, R. and R. Michaud. 2005. *Scherer's Errors.* In New Frontier Newsletter, December 2005. Available at www.newfrontieradvisors.com/research/articles/optimization-and-portfolio-construction/.

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. and R. Michaud 2008c. "Discussion on article by Campbell R. Harvey, John C. Liechty and Merill W. Liechty, Bayes vs. Resampling: A Rematch (2008, Volume 6, Number 1), *Journal Of Investment Management* 6(3):1-2.

Scherer, B. 2002. "Portfolio Resampling: Review and Critique." *Financial Analysts Journal* 58(6).

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

[1] Richard O. Michaud is President and Chief Executive Officer at New Frontier Advisors. David N. Esch is Managing Director and Research Director at New Frontier Advisors.

[2] Some examples and rebuttals include: Harvey, et al (2008a) and Michaud and Michaud (2008c); Scherer (2002) and Michaud and Michaud (2005).

[3] The return-rank averaging procedure of simulated efficient frontiers in Michaud (1998, Ch. 6) is one of a number of methods that could be used. Frontier portfolios are equally spaced with respect to standard deviation in both MU and in our replication, and frontiers are associated with others by pairing equally-ranked portfolios.

[4] Similar tests are also given in Michaud (2008a, b).

[5] We are fairly certain that the Bayesian procedure used in MU created more dispersed risk-return estimates and is perhaps a factor why their experiment found no failures in Michaud.

[6] We consider our graphical display for comparing the two procedures more transparent and reliable. This is because the MU utility function averaging process is efficient frontier curvature dependent. The portfolios chosen with the same utility function may exist near the bottom of one simulated efficient frontier and near the top of another depending on curvature.

[7] Readers can review all the results of all the simulation tests in the replication study on our website: https://newfrontieradvisors.com/media/1487/truths-1-20.pdf .

[8] Scherer (2002) who claimed to have found a flaw in Michaud optimization is actually simply another case of perverse risk-return estimates relative to a well-diversified optimization investment program.

[9] The "Simulator" option available to New Frontier software subscribers has been available for many years.

[10] New Frontier uses a logarithmic scale to define information levels for implementation.

[11] See also Weisberg (2014).

[12] The definition of Michaud MV optimization is dependent on the assessment of the level of uncertainty of information in risk-return estimates. While it is common practice to use the number of periods in a historical data set, such as in Michaud (1998) based on 216 monthly returns, there is no theoretical imperative to do so. In fact, in practice, managers should always impose their own perceptions of uncertainty independent of the data used in estimation. In practice, New Frontier has defined a logarithmic scale from 0 to 10 of "forecast certainty" as a heuristic for applications.