Examining the Presence of Luck and Skill in Mutual Fund Returns

Joe Matten
Stanford University
jmatten@stanford.edu

Abstract I wrote the following essay for Professor Shoven's class, "A Random Walk Down Wall Street." Inspired by Burton Malkiel's book of the same name, the class featured discussions of the major topics and themes of the work—one such theme being whether or not discernable skill exists in actively managed mutual funds. Quite surprising to me was the book's insistence that actively managed mutual funds underperform passive benchmarks, and even were skill to exist, it would not be discernable from luck to an investor. In my paper, I try to explain why actively managed mutual funds underperform passive benchmarks, and I include research demonstrating evidence for both sides of the argument for the presence of skill in actively managed funds. First, I present William Sharpe's logical argument for why these active funds must underperform passive funds. Next, I provide a literature review detailing some facets of the research into the persistence of mutual fund returns. I conclude by offering a more in-depth look into the seminal 2010 Fama and French paper that compares the true alpha of actively managed mutual funds to a cloned population where skill has been statistically removed.

*Author's Note: This research was done under Professor John B. Shoven in ECON 13SC: A Random Walk Down Wall Street.

Before I get to the focus of my paper, I hope to briefly motivate it by sharing some of my inspirations for choosing this topic. Prior to being introduced to Burton Malkiel's "A Random Walk Down Wall Street," I read Benjamin Graham's "The Intelligent Investor" and Peter Lynch's "One Up on Wall Street." I believed that by following the fairly intuitive advice within these books, such as trying to buy dollars for 50 cents and buying what you know, that one could have a significant advantage over the market. So after reading "A Random Walk Down Wall Street," I was perplexed as to how experienced and intelligent mutual fund managers following sound advice such as looking for arbitrage opportunities and informational asymmetries would, on average, produce returns worse than those of the passive benchmarks.

Indeed, Malkiel's claim that the average actively managed mutual fund is outperformed by the S&P 500 is backed up by numerous studies. For instance, Fama and French (2010), using their famous three factor model, estimate that from 1984-2006, the α (abnormal return) of the average after-cost (net) returns of actively managed mutual funds was -0.81% per year. Even though the literature can sometimes be inconsistent with regard to its views on the persistence of returns and the presence of skill, I did not find a single paper that failed to mention that the average after-costs returns of actively managed mutual funds was significantly less than the returns of passive benchmarks.

Before jumping into the academic research surrounding the persistence of returns and the presence of statistically significant skill, I hope to address a logical argument which informally demonstrates why the average net returns of actively managed mutual funds are lower than those of passive benchmarks. Professor William F. Sharpe (1991) provides a lucid argument that "is embarrassingly simple and uses only the most rudimentary notions of simple arithmetic." He claims that if active and passive management styles are defined in sensible ways, he will be able to prove that: (1) before costs, the return on the average actively managed dollar will equal the return on the average passively managed dollar and (2) after costs, the return on the average actively managed dollar will be less than the return on the average passively managed dollar (Sharpe, 1991). Sharpe (1991) goes about this process by first stating that the market is made up of both passive and active investors. He defines a passive investor as one who holds every security in the market. The weighting of each security in a passive investor's portfolio must be the same as that security's weighting in the market. Sharpe defines an active investor as any investor who is not passive. Because passive investors track the market, they should obtain the market return, before costs. From this fairly obvious statement, Sharpe asserts that it follows that the return on the average actively managed dollar (also before-costs) will also equal the market return. What is the reasoning behind this extraordinary claim? The market is made up of only active and passive investors. Passive investors seek to equal the market return by tracking the benchmark indices and active investors usually try to act on perceived mispricing. Since the market return must equal the weighted average of the returns in the passive and active segments of the markets, and one part (the passive segment before-costs) equals the market return, the other part (the active segment before-costs) must also equal the market return. A more mathematical explanation could be presented with the help of the

following expression: Weighted Avga+p = (wa*a + wp*p) = m, where a and p are respectively the before-cost returns of active investors and passive investors and m is the market return. The variables wa and wp respectively represent the percent weighting of a and p in the market. If p = m, then a needs to equal p in order for the weighted average of a and p to equal m. Therefore, Sharpe has proven the first assertion that before costs, the return on the average actively managed dollar will equal the return on the average passively managed dollar. In order to prove the second assertion, he relies on the assumption that the expenses needed to actively manage a fund are greater than the expenses needed to track a benchmark index. An actively managed mutual fund needs to pay for skilled analysts, brokers, and traders. Additionally, these funds typically incur more trading costs due to higher turnover. Because active and passive returns are equal before cost, and because active managers bear greater costs, it follows that the after-cost return from active management must be lower than that from passive management,

Sharpe's reasoning forms the basis of equilibrium accounting which is featured heavily in Fama and French's 2010 paper "Luck versus Skill in the Cross Section of Mutual Fund Returns." The Fama and French paper builds off the idea of equilibrium accounting and uses fairly advanced statistical procedures, such as bootstrapping, to discern skill from luck. A simpler, yet less precise, method of analyzing the skill of active mutual fund managers is to look at the persistence of their returns. The rationale behind analyzing persistence is that skilled managers (true $\alpha > 0$) will have high returns and will be able to maintain such returns through their skill. A mutual fund manager who does not have skill (true $\alpha = 0$) may get lucky, but he will not have enough skill to maintain his high returns. Equivalently, a manager with true α equal to 0 may have negative returns for a few years if he is unlucky, but he won't have enough negative skill for his poor returns to persist. Persistence also applies to managers with negative skill (true $\alpha < 0$). Due to the negative skill of these managers, negative returns also supposedly persist.

The literature surrounding persistence has not yet found a clear consensus. However, there is some evidence of short and longer-term persistence of returns. Looking at funds that existed from December 31, 1974 to December 31, 1984, Grinblatt and Titman (1992) found evidence for the persistence of both positive and negative returns. They performed a time-series regression of the returns in the last 60 months on the returns from the first 60 months. From this regression, they found a slope coefficient of 0.281, which they deemed to be statistically significant at the 0.05 level. This means that mutual funds in the second five year period are expected to realize a 0.28% higher abnormal return in the second five years for every 1% achieved in the first five years. This relationship also holds in the negative direction (Grinblatt and Titman, 1992).

In their 2002 paper,"Do Winners Repeat with Style?", Ibbotson and Patel look at returns from January 1975 to December 2000 to see if they can find evidence of short-term persistence. They define winners based off of yearly returns to see if there is persistence of returns in the next one-year period. They find that when winners are defined more stringently in the initial period, the percentage of winners repeating increases as does the average α of the subsequent period. Reproduced below is a graph showing the positive relationship between α in the initial period with α in the subsequent period. As the definition of a winner becomes more stringent (the minimum

α requirement is increased), the subsequent year's average α also increases. Two weighting schemes are used in the evaluation. The first equally weights all funds in the sample while the second equally weights the years. The latter acknowledges that the draws are not independent and are conditional on the year that is being examined (Ibbotson and Patel, 2002) [See Figure 1].

Carhart (1997) also finds short-run mutual fund return persistence. He states that the "net gain in returns from buying the decile of past winners and selling the decile of losers is 8 percent per year." Even though these results are terrific evidence for persistence, Carhart does not think that persistence in returns demonstrates skill. He instead believes that most of persistence can be explained by expenses and transaction costs. As such, he believes that evidence of persistence only occurs because expenses and transaction costs persist in the short run. The table represented below shows estimates which are derived from time-series averages of monthly cross-sectional regression slope estimates [See Figure 2]. The table shows a strong relationship between performance and expense ratios and performance and turnover. The -1.54 coefficient on expense ratios

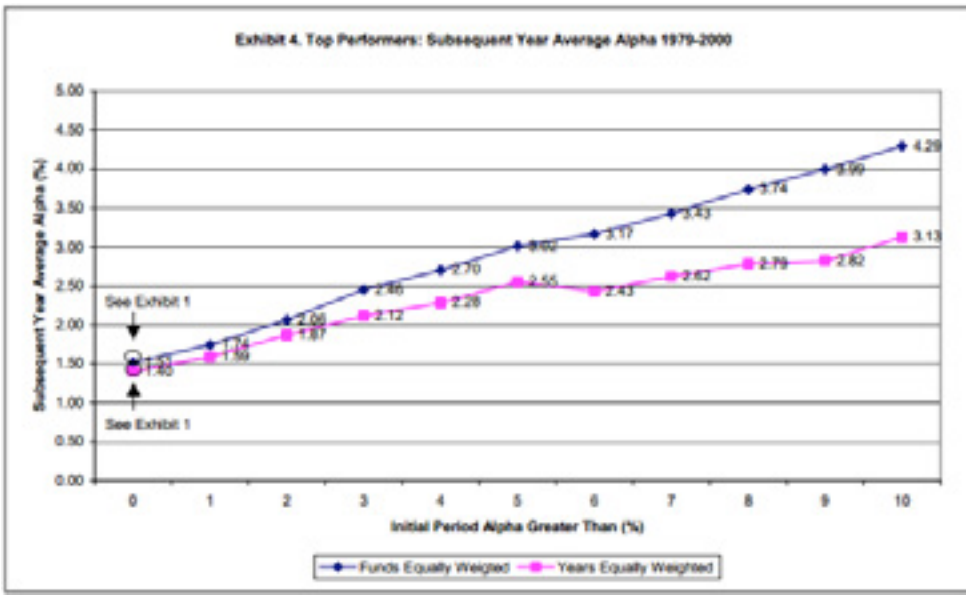

Figure 1

Independent Variables (Coefficients × 100)	Estimate	t-statistic
Expense ratio (t)	−1.54	(−5.99)
Turnover (t) (Mturn)	−0.95	(−2.36)

Figure 2

implies that for every 100 basis point increase in expense rations, α drops by about 154 basis points per year. The turnover coefficient of -0.95 implies that for every 100 basis point increase in turnover, α drops by about 95 basis points per year.

Thus far, the researchers looking at persistence have searched for maintained returns in mutual funds, but they have failed to consider the role of the fund manager specifically. One criticism of this approach is that even if persistence of returns is demonstrated, this does not show that any one particular mutual fund manager has positive or negative skill. This is because mutual funds can be managed by multiple people and the management of any mutual fund is bound to change over time. In response to this criticism, Porter and Trifts (2012) in "The Best Mutual Fund Managers: Testing the Impact of Experience Using a Survivorship-bias Free Dataset" examine 289 solo managers of mutual funds. Looking at the best solo managers (with tenures of at least 10 years) over the past 80 years, the study seeks to find a relationship between tenure and performance. The results (represented below) are not particularly

Table V. Comparing Performance in the First Three Years to Subsequent Years for Managers with Ten or More Years of Solo Tenure

The Change in Performance column shows difference between the average annual market-adjusted return in the first three years and all subsequent years, the middle value is the *t*-statistic, and the lower value is the significance level.

Panel A. All Managers

	Average annual market-adjusted returns		
	First three years	Subsequent years	Change in Performance with Experience
All Managers	1.18%	0.46%	-0.72% (-1.73) 0.0836
Best quartile	6.76%	4.56%	-2.20% (-2.10) 0.0361
Best half	4.24%	2.75%	-1.49% (-2.34) 0.0196
Worst half	-1.90%	-2.01%	-0.11% (-0.23) 0.8187
Worst quartile	-3.14%	-3.56%	-0.42% (-0.58) 0.5599

Panel B. Best 50 Managers

Best 50 Managers	8.27%	5.96%	-2.31% (-1.49) 0.1366
Best 25 Managers	11.37%	8.19%	-3.18% (-1.23) 0.2213
Best 10 Managers	11.70%	11.64%	-0.06% (-0.01) 0.9910

Figure 3

flattering for mutual fund managers. The table shows the average annual market-adjusted returns for the first three years and subsequent returns for mutual fund managers with ten or more years of experience at the same fund. In all cases, solo fund managers performed better during the first three years of their tenure than the latter part of their career. One would expect that if skill were involved, the returns in the latter part of one's career would be higher as a result of more experience. The cases where the market-adjusted returns did not change very much occurred in the very high and very low percentiles of the distribution. This might show the presence of some skill at high levels and negative skill at low levels, but as a whole the data present an outlook that is critical of skill as a major determinant of returns [See Figure 3].

Now that we have covered the usage of persistence analysis in the study of skill in actively managed mutual funds, we will shift our focus to another method of determining skill from luck. This second method is to look for statistical discrepancies between actual returns and simulated returns from funds in which true α is set to zero. By setting the benchmark's true α to zero, we are able to remove all traces of skill. The advantages of this design are that it is more precise than the persistence approach. Fama and French state that the weakness of persistence tests is that "because they rank funds on short-term past performance, there may be little evidence of persistence because the allocation of funds to winner and loser portfolios is largely based on noise" (Fama and French, 2010).

Fama and French (2010) looked at the performance of US mutual funds during the period of 1984-2006. Using their three factor model, they estimate that the average α of the mutual funds net of costs was -0.81% per year. Their estimate for the average α for the gross returns for these funds was 0.13% per year. These estimates closely match up with Sharpe's (1991) proof showing that the α of average gross returns will be 0 and the α of the average net returns will be negative. Fama and French (2010) build on Sharpe's (1991) "The Arithmetic of Active Management" when they introduce the idea of equilibrium accounting. Essentially, equilibrium accounting assumes that Sharpe's (1991) assumptions are correct and adds propositions that logically follow the assumptions. These new propositions basically entail the idea that active investment is a zero sum game before costs and a negative sum

game after costs. If the α of the average gross returns is positive, this simply means that the α of some other active investment is negative. If the α of some active mutual funds is positive, this just means that the α of some other active mutual funds is negative. The introduction of equilibrium accounting is essential for the paper because it means that if funds with negative true α are found, there are also firms with positive true alpha and vice versa.

In their 2010 paper, "Luck versus Skill in the Cross-Section of Mutual Fund Returns," Fama and French compare actual estimates (using their three-factor model) of the average α of US actively managed mutual funds to a cloned population of funds that have the return characteristics (fat tails, correlation, etc.) of the actual population of funds, except that in the cloned population, true α and thus true t(α) are zero for every fund. The symbol t(α) is the t-statistic of α, or the ratio of an α estimate to its standard error. Dividing each α estimate by its standard error gives Fama and French (2010) precision-adjusted α estimates that allow meaningful comparisons across funds. The statement that the cloned funds have true α equal to zero means that only chance and not skill is affecting the distribution of the funds with regard to t(α). For example, even though a fund may get lucky and have a higher than average t(α), we know that this result is due to chance alone. In order to create the cloned population of funds, Fama and French (2010) first estimate the three-factor model on each fund's actual returns and then subtract the resulting α estimate from the fund's returns. This creates funds that have all the properties of the actual fund's returns, except that true α and t(α) for the cloned returns set to zero. To generate a chance distribution of α and t(α) estimates, Fama and French (2010) draw a random sample (with replacement) of the 273 months from the cloned population of fund returns. This procedure is also known as bootstrapping in statistics. This yields one chance distribution of α and t(α) estimates. To have many such samples on which to base inferences, they repeat the bootstrap simulation 10,000 times (Fama and French, 2010). For every fund, they then estimate the three-factor model on the random sample of returns.

Having introduced the procedure used to compare actual returns to simulated returns, let us now analyze Fama and French's (2010) results. Represented below Table 1 are the values of three-factor t(α) estimates at selected percentiles for actual net fund returns and from the simulated returns where true α is set to zero. To the right of this table is a cumulative distribution function (CDF) plot which visually represents the data in the table. CDF plots show the cumulative sum of the probabilities up to a given point. If actual net returns are below the returns of the simulated funds (always displayed before-costs), this shows us that the mutual fund managers do not possess enough skill to cover their costs. If actual net returns are equal to the returns of the simulated benchmark, this shows that managers have enough skill to cover their costs. And, if actual net returns are greater than the simulated benchmark, this is evidence that managers have more than enough skill to cover costs. Essentially the idea behind looking at net returns is to look at skill from the investor's point of view. The investor does not care whether or not the manager has true α greater than zero, rather, the investor cares whether or not the manager has enough skill to cover his costs. For this reason, comparing net returns does not measure the true skill of the manager. In order to determine true skill, we must compare the gross returns

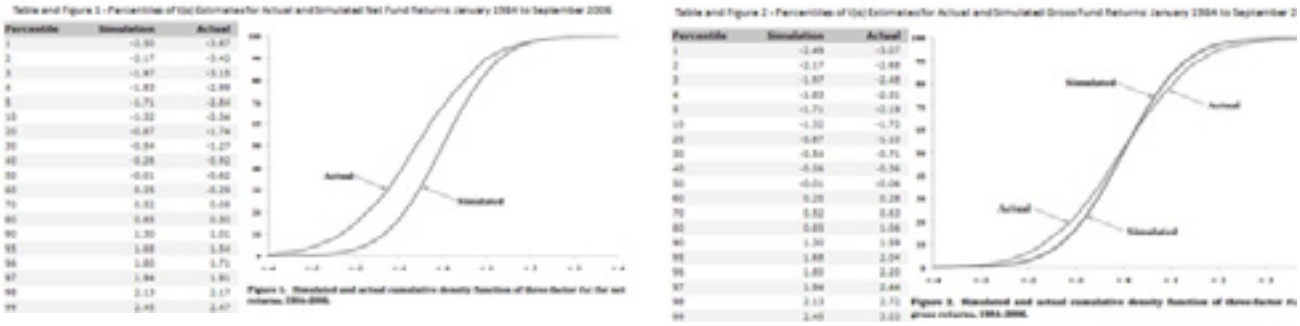

Figure 4

of the actual funds to the returns of the simulated benchmark. As can be seen from Table 1 (left), actual funds underperform the simulation all the way up to the 97th percentile. This means that only the top three percent of active mutual fund managers have enough skill to cover their costs. Looking at Table 2 (gross returns), we see that actual funds underperform the simulation all the way up to the 50th percentile. However, after the 50th percentile, they perform better than the simulation [See Figure 4]. Thus, we can conclude that there is some negative skill (true $\alpha < 0$) in the left-tail of the distribution and some positive skill in the right tail of the distribution.

In our journey through the research on the presence of skill and luck in mutual fund returns, we have come across information that is both very interesting and practicable. We have learned that investing in the average actively managed mutual fund is not a wise decision. We have also seen that there is in fact a significant amount of skill (positive and negative) at the tails of the distribution. Unfortunately for us, we are unable to discern whether better than average performance is the result of skill, luck, or the persistence of costs. In spite of this, we can take advantage of the research performed on persistence to achieve higher returns. Whether it be due to skill or the persistence of costs, there is significant evidence of persistence in the short and long run. Therefore, my advice for those unwilling to invest in an index fund is to choose a fund with low expenses and turnover which has a steady track record of solid performance (an example with which we are familiar is Dodge & Cox).

References:

Carhart, Mark M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57–82.

Fama, Eugene F. and French, Kenneth R., 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, The Journal of Finance 65, 1915-1947.

Grinblatt, Mark, and Sheridan Titman, 1992, Performance persistence in mutual funds, Journal of Finance 47, 1977–1984.

Ibbotson, Roger G. and Patel, Amita K., 2002, Do Winners Repeat with Style?, Yale ICF Working Paper No. 00-70.

Porter, Gary E. and Trifts, Jack W., 2012, The Best Mutual Fund Managers: Testing the Impact of Experience Using a Survivorship-bias Free Dataset, Journal of Applied Finance 1, 1-13.

Sharpe, William F., 1991, The arithmetic of active management, Financial Analysts Journal 47, 7–9.