INTERNATIONAL JOURNAL OF BUSINESS, 20(1), 2015 ISSN: 1083-4346

Mutual Fund Performance: Luck or Skill?

Ajay Bhootra[a], Zvi Drezner[b], Christopher Schwarz[c], Mark Hoven Stohs[d]

[a]*Mihaylo College of Business and Economics, California State University Fullerton*
Fullerton, California 92831
abhootra@fullerton.edu
[b]*Mihaylo College of Business and Economics, California State University Fullerton*
Fullerton, California 92831
zdrezner@fullerton.edu
[c]*Paul Merage School of Business, University of California at Irvine*
Irvine, California, 92697
cschwarz@uci.edu
[d]*Mihaylo College of Business and Economics, California State University Fullerton*
Fullerton, California 92831
mstohs@fullerton.edu

ABSTRACT

Should individuals include actively managed mutual funds in their investment portfolios? They should if and only if the result of active management is superior performance due to skill. This paper employs a previously ignored statistical technique to detect whether skill drives the superior performance of some mutual funds. This technique, the generalized binomial distribution, models a sequence of *n* Bernoulli events in which the result of each event is either success or failure (successive quarters during which funds outperform or do not outperform the market). Results display a statistically significant proportion of mutual funds, though small in number, outperform their peers on a risk–adjusted basis and do so as a result of skill, not luck. This result signifies the rationality of entrusting one's wealth to successful and skillfully managed mutual funds. Hence, a well–designed portfolio that includes actively managed funds may trump a wholly passive index fund strategy.

JEL Classifications: G10, G11, G12

Keywords: mutual fund performance; risk–adjusted abnormal returns; skill; luck; portfolio choice; investment decisions; risk return.

I. INTRODUCTION

Fundamental finance principles, along with common sense, dictate that investors take account of both risk and return, yielding some measure of positive risk–adjusted returns that also account for all transaction costs and/or management fees. In other words, investors seek positive alphas (α's), i.e., risk and cost–adjusted positive returns. Many investors turn to professionals in search of such gains. According to the Investment Company Fact Book (2011), almost 30% of world's wealth is invested in mutual funds: $25 trillion of the $85 trillion total world–wide investible wealth. About 44% of individual wealth in the U.S., $12 trillion, is invested in mutual funds. Mutual funds charge management fees that appear to reduce total return to investors. Many mutual funds either do not beat the market or fail to outperform their chosen benchmarks. Indeed, while Fama and French (2010) note "the aggregate portfolio of actively managed U.S. equity mutual funds is close to the market portfolio, but the high costs of active management show up intact as lower returns to investors," they document existence of genuine inferior and superior performance at the negative and positive extremes.

Results herein mirror those of Fama and French (2010), though a probability approach leads to the results herein versus the bootstrap method used by them. Experiments involving many successive trials with multiple subjects, demonstrate a positive probability that a subject will win several times in a row merely due to luck, not skill. For example, while the probability that a person flips a coin and ends up with 10 heads in a row is less than one in a thousand (1/1032 = 0.098%), such strings of luck do occur, and when they do, it is not skill on the part of the person flipping the coin. Some of the thousands of investment funds may end up the "winners" for prolonged periods of time. Naturally, investors need to know whether streaks of superior performance are due to skill. The analysis herein demonstrates that such streaks exist within the family of mutual funds and are due to skill, not luck.

The main contention of this paper relies on detecting the existence of skill with the generalized binomial distribution, using an established analogy with winning in sports. Money managers are usually well–educated and even skilled, as are professional athletes. Competition alone in sports dictates that not all professional teams can win; though as noted in more detail below, the most highly skilled teams appear to win more often. In contrast, a good deal of prior evidence indicates that a majority of actively managed mutual funds fail to outperform passively managed index funds.

According to Berk and Green (2004), the apparent lack of investment skill has "led researchers to raise questions about the rationality of investors who place money with active managers despite their apparent inability to outperform passive strategies and who appear to devote considerable resources to evaluating past performance of managers." This paper contributes to this important debate about luck vs. skill. Using a novel statistical approach based on the Generalized Binomial Distribution (Drezner and Farnum, 1993), this paper investigates whether some mutual funds provide positive α's over time as a result of skill. The approach demonstrates, in contrast to a large degree of the extant evidence, that skilled fund managers do exist – persistence in superior performance cannot be attributed to luck.

This paper's results suggest important practical implications. The lack of convincing evidence that mutual funds provide superior performance dramatically influences the investing habits of individuals. The lack of such evidence along with such straightforward factors as the ease of trading ETFs and lower management costs for ETFs have most likely led to the growth of ETFs and a relative decline in mutual funds in recent years. For example, from 2005–10 the number of ETFs grew from 204 to 950, a growth of nearly 400%, with Weinberg (2012) reporting a total of over $1.4T invested in ETFs by 2012 from a zero base in 1994. Between 2005 and 2010 the number of mutual funds grew from 8,449 to 8,545, a growth of just over 1% with the apparent peak of 8,884 mutual funds in 2008 according to the Investment Company Fact Book (2011). Indeed, from the peak in 2008 to the more recent number of 8,545, the decline in the number of mutual funds is almost 4%. The research herein focuses on the relative numbers of these types of funds a as symptom of broader movements in the investment world and refrain from a more in depth analysis of such funds, given the aim of addressing the luck vs. skill issue.

Doubt surrounding the genuine performance of mutual funds is probably a factor in the relative lack of growth of mutual funds in recent years. However, if the (higher) risk–adjusted performance of some mutual funds is due to skill, then investors will not abandon them altogether. Aside from the crucial impact on financial theory, the long–term fate of a $25 trillion industry rests in part on whether that industry succeeds on the basis of skill or luck. Results herein provide one additional reason for believing that skill is involved, and thus that the mutual fund industry should have a secure future.

While the paper shines a spotlight on one particular area of contention in the broader debate, the analysis of mutual fund performance remains enveloped in controversy. At least four substantial debates lie at the center of any such controversy about the risk–return tradeoff. The perspective offered herein does not aim to resolve these debates, rather it indicates which approach to these debates keeps the spotlight on the paper's central contribution. The paper's secondary contribution lies is delineating the major debates, given the overall complexity of differentiating skill from luck. The aspiration to identify the nature of the skill involved in achieving extraordinary returns with a mutual fund investment remains. The conclusion addresses this natural aspiration.

The first debate centers on how to measure risk appropriately. It is largely related to the validity of CAPM or its alternatives. The standard approach in the analysis of mutual fund performance side–steps these controversies by measuring performance given any reasonable metric, i.e., the Fama–French 3–factor and the Carhart 4–factor models. In addition to the choice of a one–period measure of risk, Welch and Levi (2012) propose that longer–run risk measures have no predictive power. While acknowledging the importance of measuring risk, the analysis herein utilizes the standard approach.

A second debate, tied to the issue of measuring risk, concerns the appropriate time period for measuring risk–adjusted returns. Fama and French (2010) note this issue by distinguishing their use of "long histories of individual fund returns" from the alternative approach of persistence tests such as those conducted by Grinblatt and Titman (1989) and Carhart (1997), i.e., "that is, whether past winners continue to produce high returns and losers continue to underperform." The analysis herein fits within the "persistence" tests category.

The third debate concerns the choice of appropriate comparison samples. An instructive example is the Wall Street Journal's "Investment Dartboard Contest," which has been held for almost 50 years. The contest appears to addresses the "luck vs. skill" debate. Can luck, disguised as stock portfolios "chosen" by darts, outperform the skill of individual investors making their own reasoned portfolio choices? Ensign (2012) reports that in this well–received contest, luck wins out, in which darts beat "the readers, with 28 wins out of 46 contests." However, such contests use samples that are inappropriate for answering the question posed, about whether superior mutual fund performance is due to skill or luck. Simply, investment professionals are not permitted to select the portfolios submitted by readers. Hence, the WSJ contest is really luck vs. the uninformed, not luck vs. skill. Their contest is fun, sells papers, and provides some useful information to the casual investor, though it remains silent about whether superior mutual fund performance is due to skill.

How to choose an appropriate sample may remain a concern for years to come, given that minimally, performance may be pegged to the: (1) mutual fund itself, (2) manager of the fund, including appropriate training or education, (3) manager and team, (4) mutual fund company which selects managers, and even (5) the stock picking and investment procedures used or permitted (e.g., whether funds allow short–selling). This paper focuses on (1), the performance of mutual funds themselves.

The fourth debate focuses on the statistical approach and/or methodology used to distinguish luck from skill. Naïve and perhaps even sophisticated statisticians or econometricians are likely to agree that statistics should be able to provide a definitive approach. Yet, no consensus has yet been achieved. Fama and French (2010) use bootstrapping. Barras, Scaillet and Wermers (2010) refine a standard approach by employing a control for zero–alpha funds. The primary contribution herein centers on using the Generalized Binomial Distribution (GBD) to provide a new and important perspective about the luck versus skill mutual fund debate. Drezner and Farnum (1993) first used the GBD to analyze an analogous debate about of skill vs. luck in sports and in student success.

The data section describes the procedures for creating the sample of monthly returns for 981 mutual funds for a 15-year period from 1995 to 2009. Intuitively, the analysis examines how many funds consistently performed "at the top" across the whole sample period. For example, a natural and important question is, how many funds consistently performed within the top 25% of all mutual funds over the whole sample period of 15 years? The funds in this top 25% are the winners. The basic argument is that the GBD technique implies that they are not winners by chance (luck). They win as a result of skill. The results provide an estimate of the number of funds expected to win as a result of skill using the standard binomial distribution. If fewer such funds exist than expected, perhaps because funds regularly move into and out of a winning position, then mutual fund performance cannot be distinguished from luck. However, if significantly more such funds exist in the sample than expected, skill is involved. The number of such funds must be significantly higher by an appropriate statistical test to reach such a conclusion.

The percentage constraint is p in the analysis (the percentage of funds in the top p of all funds across the sample period), which can be set to any level of probability. Percentages used include $p = 0.25$, 0.5 and 0.75 which means that the mutual fund's rate of return was at the top one quarter of the returns, at the top one half, or the top

75% (alternatively not in the bottom 25%) for the entire period. The results over 180 months display statistical significance at an extreme level.

I. DATA

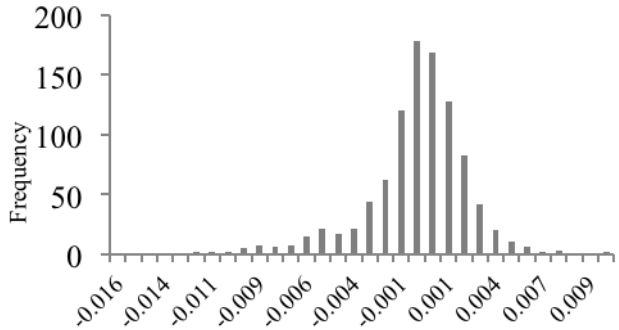
The data analyzed in this paper draws on the data employed and described in detail in Schwarz (2011). Briefly, three sources combine to generate the data. First is the Mutual Fund database as of 2007 from the Center for Research in Security Prices (CRSP), which provides the standard information about mutual funds, most importantly the type of fund, NAV, returns, and asset data. Second, the Thomson Financial Mutual Funds Holding Dataset (TFDS) as of 2007 is used to filter the first set of funds to permit only stock funds (50% or more of holdings in stocks) and those with total assets of at least $10 million are in the final dataset. Finally, the analysis requires linking the dataset of mutual funds to the CRSP Mutual Fund dataset using the MF Links files available via WRDS. This link file focuses largely on equity funds. Returns in the final set of mutual funds may be calculated beyond the set "cutoff" date of 2007, which is why returns extending to December 2009 are indicated below.

The basic dataset generates rates of return for 5,134 mutual funds over the 180 month period from January 1995 to December 2009, including 553,973 return "observations." The GBD based approach requires that only funds with all available observations over the entire sample period be included in the analysis. Many of the mutual funds did not operate over the whole period of 15 years and thus were dropped from consideration. The sample includes only those mutual funds that operated throughout the whole period, yielding a final set of monthly returns for 981 mutual funds over 180 months, from January 1995 to December 2009. On average, these "survivor" funds are likely to be more skilled as compared with the ones that did not survive the entire period. However, not all funds within this group are equally skilled. As demonstrated later, the statistical approach based on GBD allows for identification of a small number of skilled funds within these survivors. As a result, the analysis herein does not suffer from survivorship bias as documented by Brown, Goetzmann, Ibbotson, and Ross (1992) and others.

Finally, time-series regressions of monthly fund returns on contemporaneous risk factors each year (using 12 observations for each year in the regressions) produce the risk-adjusted returns (alphas $\equiv \alpha$s) for all funds. The regressions employ the three standard risk-adjustment approaches: (1) CAPM, (2) the Fama–French 3–factor model and (3) the Carhart 4–factor model, yielding three sets of 14,715 annual αs (15 annual α observations for each of the 981 sample funds). For details on calculating the alphas, see Appendix. All three approaches yield consistent and robust results, though the discussion below focuses only on αs resulting from the Carhart 4–factor model.

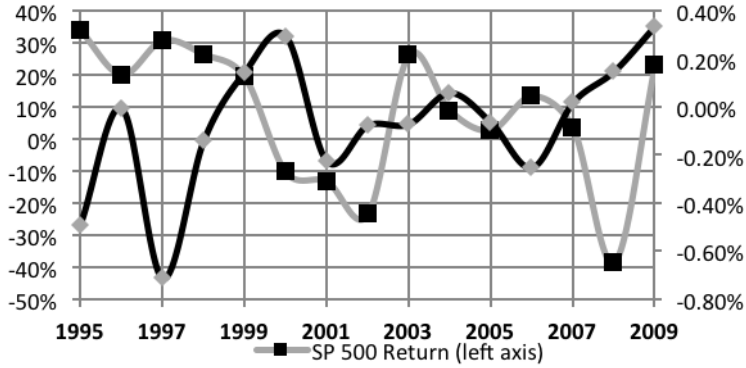
Before the central analysis proceeds, briefly consider the descriptive statistics for the annual fund αs. As noted above, the data set includes 14,715 αs consisting of the excess annual returns for 981 mutual funds over 15 years. Figure 1 displays a histogram of the average excess return by fund across all 15 years. As expected, the graph approaches the shape of a normal curve, with a slight negative skew.

Figure 1
Histogram of the average α by mutual fund during the sample period 1985–2009



Notes: The bars represent average αs (x–axis) for the 981 mutual funds during the sample period, 1995–2009.

Figure 2
The average annual α of all 981 sample mutual funds by year during the sample period 1985–2009



Notes: The graph includes connecting lines in order to relate returns to one another. The lines are not intended to be indicative of mid–period (e.g., monthly) returns.

Next, consider the all fund (aggregate) average annual alpha by year, largely to explore whether any obvious "trend," or lack thereof, exists. The graph in Figure 2 above compares the average annual mutual fund α to the average annual return of the S&P 500. It should be noted that while none of the sample mutual funds are index funds, one would expect that in the aggregate their performance matches the whole market, while also keeping in mind that one cannot calculate αs for the S&P 500. Figure 2 displays the relatively random nature of annual (aggregate) mutual fund αs.

II. THE GENERALIZED BINOMIAL DISTRIBUTION

Suppose that a sequence of n Bernoulli events (the result of each event is either "success" or "failure") is observed. Each outcome has a probability p of success and 1–

p of failure. If the events are random and not correlated, the distribution of the number of successes is a Binomial distribution. There are many situations where the events are related. Drezner and Farnum (1993) proposed the Generalized Binomial Distribution (GBD) to model such related events. Examples of such situations are grades on a multiple choice exams, standing of teams at the end of a baseball season (Drezner and Farnum, 1993) or the standing at the end of the NBA season (Drezner, 2006). A strong analogy exists between skill in sports and the current case of whether mutual fund performance displays skill over time.

In all these cases it is quite evident that skill plays a role in longer term (repeated) outcomes. If a student answers the first question on an exam correctly, the probability that he/she answers the second question correctly increases; while if he/she answers the first question incorrectly, the probability that he/she answers the second question correctly decreases. The same reasoning holds in sports where a first game win increases the probability of winning the second game due to skill. The reason for this observation is that skill is involved and the outcome is not random.

Consider a correlation factor θ describing the skill level. The value of θ is determined by fitting the resulting frequency distribution to the data by a χ^2 goodness of fit test. Suppose that in the first k events there are r successes recorded. The "rate" of success is r/k. The probability of success for the next event equals:

$$(1-q)p+q\frac{r}{k} \tag{1}$$

When events are independent and no skill is involved, then θ = 0 and the probability of success is p regardless of the history. When θ = 1, then if the first event is a success, all events are successes; and if the first event is a failure, then all k events fail. An increasing θ between these two extremes depicts an increasing level of skill. Specific values of θ for exam grades (0.5921), baseball games (0.397) and NBA games (0.5765) were derived so that the GBD fits the actual results of these data sets very well, while the binomial distribution does not fit the actual data.

Drezner and Farnum (1993) provide proofs for the properties of the GBD. Two important properties for the distribution exist:

Property 1:

$$\text{Mean} = np \tag{2}$$

Property 2:

$$\text{Variance} = (1-p)\frac{1-\sum_{k=0}^{n-1}\frac{k+2q}{k+1}}{1-2q} \tag{3}$$

In the case of mutual fund performance, values for p, such as p = 0.5, define the "success" that occurs for a mutual fund when it performs in the top 50% of all mutual funds in the specified period. If skill is involved in the performance of mutual funds the distribution of number of successes should differ from a binomial distribution because success in one period increases the probability of success in the next period.

The χ^2 goodness of fit test indicates whether the distribution of the returns of mutual funds is significantly different from a binomial distribution. Suppose that data about m periods for a given proportion p is given. The range of the binomial distribution is between 0 and m. The range [0, m] is divided into 10 ranges so that the binomial probability for each range is about 10%. This yields 10 expected values, one for each range. The actual counts in each range constitute the actual data and the χ^2 statistic with 9 degrees of freedom is found with the corresponding p–value for the goodness of fit test. Low p–values, when comparing the data to the binomial distribution, indicate deviation from randomness or, i.e., skill. When comparing the data to the GBD, high p–values suggest no deviation from GBD. Minimizing the χ^2 statistic by employing the solver in Excel provides the best fitting value of θ.

III. ANALYSIS OF THE DATA RESULTS USING αs

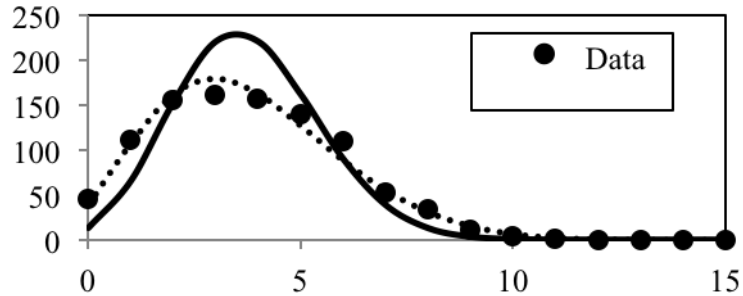
To adjust the data for different types of mutual funds the rates of return were adjusted to provide annual αs for 15 years for these 981 mutual funds. As noted above, the only results presented rely on the αs as calculated using the Carhart 4–factor model. Table 1 summarizes the statistical results, and Figures 3–5 display graphs of the three cases. The analysis shows that to be and remain at the top 25% of mutual fund performance over time (or avoiding the bottom 25%) clearly requires skill, while to be among the top 50% requires minimal skill. Hence the results herein clearly demonstrate that persistently good mutual fund performance requires skill. Performance placing a manager in the top 50% is not statistically significant (the binomial p–value = 0.059), because being slightly above or below the median is a matter of luck, which affects the count of instances below and above the median. However, the cut–off for being in the top 25% or the top 75% (or avoiding the bottom 25%) does not result in many instances in the "gray area" that are the result of luck.

Table 1
Goodness of fit

Binomial			Generalized Binomial Distribution		
p	χ^2	p – value	θ	χ^2	p – value
0.25	220.46	1.7×10^{-42}	0.245	10.54	0.308
0.50	16.40	0.059	0.064	7.42	0.594
0.75	121.82	5.7×10^{-22}	0.262	4.55	0.872

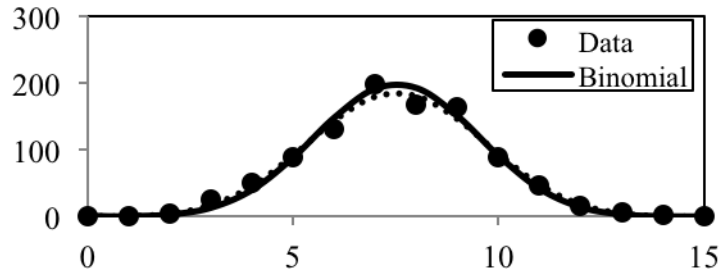
Notes: Table 1 presents statistical results of the analysis. p is the probability that a mutual fund performs in the top p % of all funds during the sample period. As noted above, low p–values, when comparing the data to the binomial distribution, indicate that skill is involved in remaining among the top mutual fund performers. When comparing the data to the GBD, high p–values suggest no deviation from GBD. Row 1 (p = 0.25) corresponds to Figure 1 below, Row 2 to Figure 2 and Row 3 to Figure 3.

Figure 3
Mutual fund performance data versus the binomial and the GBD distributions
$p = 0.25$



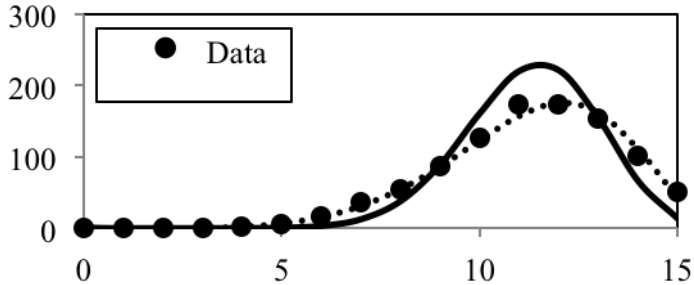
Notes: Figure 3 shows the number of times a mutual fund's annual performance consistently ranks among the highest 25% ($p = 0.25$) during the sample period, as compared to the theoretically predicted number of times according to both the Binomial and the GBD (Generalized Binomial Distribution) distributions. Actual mutual fund performance data (•) clearly indicate a stronger fit with the GBD versus the Binomial distribution, yielding the likelihood that skill is required to be among the top 25%.

$p = 0.50$



Notes: This figure shows the number of times a mutual fund's annual performance consistently ranks among top half (p = 0.50) during the sample period, as compared to the theoretically predicted number of times according to both the Binomial and the GBD (Generalized Binomial Distribution) distributions. Actual mutual fund performance data (•) is virtually indistinguishable from either the Binomial or GBD distributions, indicating that performing among the top half of mutual funds is unlikely to require skill.

$p = 0.75$



Notes: This figure shows the number of times a mutual fund's annual performance consistently ranks among top 75% during the sample period, as compared to the theoretically predicted number of times according to both the Binomial and the GBD (Generalized Binomial Distribution) distributions. Actual mutual fund performance data (•) clearly indicate a stronger fit with the GBD versus the Binomial distribution, yielding the likelihood that skill is required to avoid being among the bottom 25%.

IV. PRACTICAL APPLICATIONS

Results demonstrate that mutual funds that consistently outperform do so as a result of skill. One remaining question is whether it is possible to translate this knowledge into winning strategies for individuals investing in mutual funds. Many such strategies may exist. One rather intuitive approach follows.

Simply, individuals may identify funds that remain in the top 25% of all funds for five consecutive years (using the techniques described herein). Recall that the motivating factor is that such funds "won" as a result of skill, and that skill remains for a period of time. In sports, team owners and sports fans rely on the persistence of skill. By analogy and as shown by the analysis herein, the same holds true for mutual fund performance. Out of the 981 mutual funds, five remained in the top 25% for the 1995–99 period, and four of these five outperformed the market. Their average was higher than the average for the remaining 976 returns with a p–value of 0.08. Note also that the expected number of funds in the top 25% for the first five years is 981/1024 < 1. The fact that five exist indicates skill, not luck.

To explore the subsequent performance of the winners in more detail, consider whether the top performing funds in an early period maintain that superior performance throughout the sample period. Table 2 presents the (statistical) significance of that performance, given the "winners" during various years. At the end of one year, the top 25% (245 funds) of the 981 funds are automatically the top performers. After two years, only 50 funds remain in the top 25% for both years, and so on as indicated in row two.

The results presented in Table 2 reinforce the primary conclusion and constitute the main contribution to the literature, i.e., some superior mutual fund performance occurs through skill. Indeed, it is likely that individual investors who invest in managed mutual funds with appropriate superior performance make rational investment decisions.

Table 2
Top mutual fund performance

# of Years in Top 25%	1	2	3	4	5
# of Mutual Funds in Top 25%	245	50	24	11	5
Statistical Significance	2.7×10^{-07}	2.9×10^{-06}	6.3×10^{-04}	6.6×10^{-05}	0.08

Notes: This table shows the number of mutual funds (row 2) in the top 25% of all fund performance for the number of years given in row one. For example, after the first year's performance, 245 funds out of 981 funds total performed among the top 25. The statistical significance of their superior performance (based upon the average $\boldsymbol{\alpha}$), compared to the lower performing group for the remainder of the sample period, is given in row three.

V. CONCLUSION

This paper demonstrates that skill is involved in a mutual fund remaining among the top 25% of performers of all mutual funds over a 15 year period. Since the analysis herein does not attempt to identify the source of that skill, the exact nature of that skill remains elusive. The desire to identify the nature of investment skill appears perfectly rational. If the needed skill can be defined and "packaged," then more funds would be winners.

However, with more "winners," the perceived gains from investing with some skilled investment approach end up being arbitraged away, especially when the skill becomes widely known, accessible, and implemented. An analogy with sports teams provides insight again. Consider a professional football team that devises an innovative strategy for defense and ends up with a perfect season. As long as other teams can eventually replicate the strategy and the skill spreads, the success of the new team remains limited. Eventually, the competitive nature of sports eliminates the advantage arising from the new strategy. Of course, the fundamental principles of arbitrage and efficient markets enforce the same basic result for investment strategies. If a winning investment strategy can be replicated, it will not long remain a winner. Longer–term winning strategies also require skill at concealing the winning strategy. Thus, the public identification of winning investment strategies is self–defeating.

In addition, the small percentage of winning investment strategies indicates that the attempt to discover some immutable nature of the skill involved cannot be successful. There are likely to be several distinct skilled strategies such that identifying one or even several winning strategies from among a small sample of them is doomed to failure given that meaningful statistically significant results could not be achieved (due to the small sample).

Nonetheless, the persistence of successful mutual fund performance is due to skill, and this conclusion has important ramifications for the worldwide multi–trillion dollar mutual fund industry. While not a direct confirmation nor an endorsement of such measures, the primary result of this paper lends some support to guides such as Morningstar's five–star rating system for mutual funds.

APPENDIX

Alphas from factor models

Alphas are obtained from time-series regressions of fund excess returns on factors, as specified below. The factors are obtained from Ken French's website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

CAPM: $r_{it} - r_{ft} = \alpha_i + \beta_{i,mkt}(r_{mt} - r_{ft}) + \varepsilon_{it}$

Fama-French: $r_{it} - r_{ft} = \alpha_i + \beta_{i,mkt}(r_{mt} - r_{ft}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \varepsilon_{it}$

Carhart: $r_{it} - r_{ft} = \alpha_i + \beta_{i,mkt}(r_{mt} - r_{ft}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,MOM}MOM_t + \varepsilon_{it}$

where:
α_i = alpha of fund i,
r_{it} = return of fund i in month t,
r_{mt} = return on market portfolio in month t,
r_{ft} = risk-free rate in month t,
SMB_t = size factor in month t,
HML_t = book-to-market factor in month t,
MOM_t = momentum factor in month t, and
β's are regression coefficients on market, SMB, HML, and MOM factors.

REFERENCES

Barras, L., O. Scaillet, and R. Wermers, 2010, "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas," *Journal of Finance* 65(1), 179–216.

Berk, J.B., and R.C. Green, 2004, "Mutual Fund Flows in Rational Markets," *Journal of Political Economy* 112, 69–1295.

Brown, S.J., W. Goetzmann, R.G. Ibbotson, and S.A. Ross, 1992, "Survivorship Bias in Performance Studies," *Review of Financial Studies* 5, 553-580.

Carhart, M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance* 52(1), 57–82.

Drezner, Z., and N. Farnum, 1993, "A Generalized Binomial Distribution," *Communications in Statistics: Theory and Methods* 22, 3051–3063.

Drezner, Z., 2006, "On the Limit of the Generalized Binomial Distribution," *Communications in Statistics: Theory and Methods* 35, 209–221.

Ensign, R.L., 2012, "Darts Top Readers Again, Extend Overall Lead," *The Wall Street Journal*, July 14.

Fama, E.F. and K.R. French, 2010, "Luck versus Skill in the Cross–Section of Mutual Fund Returns," *Journal of Finance* 65(5), 1915–1947.

Fulkerson, J.A., 2013, "Is Timing Everything? The Value of Mutual Fund Manager Trades," *Financial Management* 42(2), 243–261.

Grinblatt, M., and S. Titman, 1989, "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," *Journal of Business* 62(3), 394–415.

Investment Company Institute. *2011 Investment Company Fact Book*.

Morningstar, morningstar.com

Schwarz, C.G., 2012, "Mutual Fund Tournaments: The Sorting Bias and New Evidence," *Review of Financial Studies* 25(3), 913–936.

Welch, I., and Y. Levi, 2012, "The Case for Risk–Neutral Capital Budgeting," *Working Paper*, UCLA.

Weinberg, A., 2012, "What ETFs' Next Act Will Look Like," *The Wall Street Journal*, October 23.