Can mutual fund stars still pick stocks?: A replication and extension of Kosowski, Timmermann, Wermers, and White (2006)

Timothy B. Riley*
tbriley@uark.edu

Sam M. Walton College of Business
University of Arkansas
Fayetteville, AR 72701

January 27, 2019

Abstract

Kosowski, Timmermann, Wermers, and White (2006) use a novel bootstrap technique to study the performance of domestic equity mutual funds over the period 1975 to 2002. They find that "a sizable minority of managers pick stocks well enough to more than cover their costs." When replicating their analysis during their period of study, I find results similar to theirs. However, if I perform an identical analysis over the period 2003 to 2017, I find no evidence of stock selection ability in excess of costs. Furthermore, the combined 1975 to 2017 period indicates that the alphas of the best funds likely occur solely due to luck.

Keywords: mutual fund; skill; alpha; bootstrap; luck
JEL codes: G00; G11; G14; G20

* I thank Jon Fulkerson, Yuekun Liu, and an anonymous referee for their helpful comments and suggestions. All errors are my own.

Can mutual fund stars still pick stocks?: A replication and extension of Kosowski, Timmermann, Wermers, and White (2006)

Using a novel bootstrap technique, Kosowski, Timmermann, Wermers, and White (2006) (henceforth, KTWW) "find that a sizable minority of managers pick stocks well enough to more than cover their costs" over the period 1975 to 2002. Their work has subsequently been heavily cited as evidence in support of mutual fund manager skill and the value of active management in the mutual fund industry.[1]

Despite KTWW's findings, the years since the publication of their study have seen a large shift away from actively managed mutual funds towards those that are passively managed. In 2007, about 85% of mutual fund assets were held by actively managed funds, compared to only 65% in 2017. In each year from 2014 to 2017, actively managed funds had negative net cash flows. On average, the net cash flow for actively managed funds during those years was about −$222 billion. Over the same period, index funds had an average net cash flow of about $184 billion.[2]

This discordant shift could be occurring because increased market efficiency has made KTWW's findings unreflective of the modern mutual fund industry. Chordia, Roll, and Subrahmanyam (2008, 2011) and Conrad, Wahal, and Xiang (2015) show evidence of increased market efficiency in recent years, which Bernstein (1998) argues will decrease the probability that a given fund manager will outperform the market. If KTWW's findings are outdated because of changes in the market, then it is not surprising that the actions of investors appear to run counter to their findings.

[1] As of November 29, 2018, Google Scholar shows their study has 936 total citations, with 85 citations occurring in 2017. The Wiley Online Library shows their study has been cited by 34 other studies published in the *Journal of Finance*, *Journal of Financial Economics*, or *Review of Financial Studies*.

[2] These statistics are from Figure 2.7 and Table 43 in the 2018 Investment Company Factbook, which is published by the Investment Company Institute (ICI). The report is available at the link below.
https://www.ici.org/pdf/2018_factbook.pdf

Related research since KTWW's study indicates that the number of outperforming funds is decreasing over time. Barras, Scaillet, and Wermers (2010) "observe that the proportion of skilled funds decreases from 14.4% in early 1990 to 0.6% in late 2006," and Fama and French (2010) find that "few funds produce benchmark-adjusted expected returns sufficient to cover their costs" over the period 1984 to 2006.[3] KTWW themselves document a downward trend and note that "outperforming fund managers have become more scarce since 1990." Hence, investors' shift from active to passive funds appears to be supported by real changes in the mutual fund industry.

In this study, I directly test whether KTWW's findings hold outside of their 1975 to 2002 sample period. First, I consider whether their findings hold over the 15 years (2003 to 2017) since their sample ended. This analysis is essential because researchers citing KTWW need to understand whether KTWW's findings (i) give an accurate accounting of the modern mutual fund industry or (ii) only provide a record of that industry in the past. Then, I consider results generated using a combined sample formed from KTWW's original period and my extension (1975 to 2017). The results from this combined sample indicate what KTWW would have found if the full historical record available now was available at the time they completed their study. This analysis is also essential because it indicates whether KTWW's conclusions hold using the most comprehensive sample obtainable today.

When replicating KTWW inside their sample period (1975 to 2002), I find similar results to those presented in their study. They found that funds in the top 10% of performance tend to produce alphas in excess of costs that cannot be explained by luck alone. I also identify such alphas

[3] Like KTWW, Fama and French (2010) generate their results using a bootstrap procedure, but they argue that KTWW's bootstrap procedure is flawed and that KTWW's sample construction generates survivorship bias. Conversely, Harvey and Liu (2018) contend that, because of Fama and French's (2010) method of reducing survivorship bias, their analysis "lacks power to detect outperforming funds." Engaging in that debate and critiquing methods in depth is outside the scope of this study.

within the top 10% of fund performance. However, when I run an identical analysis using data from after KTWW's sample period ended (2003 to 2017), I find no evidence of alphas greater than would be expected based solely on luck. Likewise, if I use the combined sample (1975 to 2017), I find little evidence to suggest high alphas are not due to luck. Consistent with the actions of investors and research since KTWW's publication, the "sizable minority of managers [who] pick stocks well enough to more than cover their costs" appears to have substantially decreased in size—or perhaps disappeared entirely—during the last 15 years.[4]

My results indicate that KTWW's study does not reflect the modern mutual fund industry or the full testable history of the industry. Therefore, it is essential to use KTWW's study carefully in discussions related to active management. The time period dependence of KTWW's results must be acknowledged. Furthermore, my results suggest caution with respect to the utility of the large number of predictors of mutual fund performance that have been documented in the literature.[5] Such measures have been shown to have decreased predictive power out-of-sample (see Jones and Mo, 2018), and the lack of fund managers with outperformance not attributable to luck over the period 2003 to 2017 suggests such measures are searching for alpha that is increasingly difficult to find.

It is important to note what my results do not indicate. First, both KTWW's study and my study only evaluate domestic equity mutual funds and are accordingly unable to comment on the performance of other mutual fund investment styles. In 2017, domestic equity funds held 43% of all mutual fund assets—the largest category, but still less than half of all assets—so results from

[4] KTWW find the strongest evidence of outperformance that cannot be explained by luck alone among growth-oriented funds. If I limit my analysis to just that group of funds, my conclusions do not change.
[5] Among many others, see Amihud and Goyenko (2013); Cremers and Petajisto (2009); Doshi, Elkamhi, and Simutin (2015); Kacperczyk, Sialm, and Zheng (2005); and Kacperczyk, Sialm, and Zheng (2008) for predictors of mutual fund performance.

domestic equity funds cannot be taken to represent the entire mutual fund industry.[6] Huang, Lee, and Rennie (2018) perform a bootstrap procedure for domestic bond mutual funds over the period 1999 to 2016 and find that the top 50% of those funds pick bonds well enough to more than cover their costs. Consistent with Dyck, Lins, and Pomorski (2013) and Hoberg, Kumar, and Prabhala (2018), the value of active management could be greater for mutual funds (i) investing in markets less efficient than the domestic equity market or (ii) facing less competition from other mutual funds.

My results should also not necessarily be taken as evidence that mutual fund manager skill has decreased or disappeared. The Berk and Green (2004) model of the mutual fund industry suggests that, even if skilled managers exist, no mutual fund should have a persistent positive alpha because of the flow-performance relation and diseconomies of scale. Pastor, Stambaugh, and Taylor (2015) find that mutual fund managers are actually increasing in skill over time but show that increasing competition between fund managers has had an offsetting effect on performance. It is further possible that the results during the 2003 to 2017 period are not reflective of a downward trend, but rather indicative of particular market conditions that affect mutual fund performance.[7]

In summation, KTWW's finding that a significant number of mutual funds outperformed after costs over the period 1975 to 2002 does not hold over the subsequent 15 years (2003 to 2017) or in the most complete sample available today (1975 to 2017). When citing KTWW, it is important to note both the sensitivity of their results to time period and that their results are not reflective of the recent history of the mutual fund industry. Nonetheless, the fact that KTWW's

---

[6] This statistic is from Figure 2.2 in the 2018 Investment Company Factbook.

[7] Avramov and Wermers (2006); Kosowski (2011); Pastor, Stambaugh, and Taylor (2017); von Reibnitz (2017); and Dong, Feng, and Sadka (2018), among others, show evidence that mutual fund performance varies with market conditions.

results do not hold out-of-sample should not be used to draw overly broad conclusions about all actively managed mutual funds and their managers.

1. Methods

Because I carefully replicate KTWW's bootstrap procedure, I refer readers to their study for the full details. In short, their procedure compares the actual distribution of fund alphas to the distribution that would exist if all funds had a true alpha of zero. That second, purely hypothetical distribution is created by (i) constructing a time-series of returns attributable to systematic factors alone for each fund (i.e., the returns excluding alphas and residuals); (ii) adding a randomly drawn residual (with replacement) from a given fund's set to each of a given fund's systematic factor returns; and (iii) estimating the alpha for every fund using those constructed returns. Repeating steps (ii) and (iii) 1,000 times provides the information necessary to perform the comparison between the actual distribution and the true-zero-alpha distribution.

The main performance model used by KTWW to estimate the systematic exposures, alphas, and residuals is the Carhart (1997) model. Therefore, I use the same model in this study.

$$r_t = \alpha + \beta \cdot RMRF_t + s * SMB_t + h * HML_t + p * PR1YR_t + \varepsilon_t \tag{1}$$

r_t is the excess net return for a given fund in month t, $RMRF_t$ is the excess return on the value-weighted US equity market portfolio, and SMB_t, HML_t, and $PR1YR_t$ are zero-investment factors corresponding to size, value, and momentum. α is the alpha of the fund, and ε_t is the month t residual. KTWW also present some results using a conditional version of the Carhart (1997) model styled after Ferson and Schadt (1996). All of my conclusions are the same using that alternative model.[8]

[8] KTWW report that their results are robust to 15 different performance models. I note that if the Cremers, Petajisto, and Zitzewitz (2012) four-factor model—which corrects for biases in the Carhart (1997) model—is used, there is evidence of alpha that cannot be explained by luck alone in the 1975 to 2002 period (starting at the 80th percentile), but no evidence of such alpha in the 2003 to 2017 period.

KTWW use both alpha and the T-statistic associated with alpha as measures of performance. In my analysis, I only tabulate results using alpha. My general conclusions are the same using both measures, and I note in later discussion any instances where the alpha T-statistic results have meaningful deviations from the alpha results. Following KTWW, the alpha T-statistics are calculated from Newey and West (1987) standard errors robust to heteroskedasticity and autocorrelation.

2. Data

I follow KTWW and build a sample of domestic equity mutual funds using the CRSP Survivorship-Bias-Free US Mutual Fund database. I identify the investment style of a given fund using its CRSP objective code and its name. Any fund with a CRSP objective code of EDCM, EDCS, EDCI, EDYG, EDYB, or EDYI is included if its name does not contain certain terms (e.g., "international", "fund of funds", "real estate").[9] Funds with the CRSP objective code M (i.e., mixed) must clear the additional hurdle of holding at least 50% of their assets in domestic equities in the majority of years.[10] Index funds are removed from the sample by using a flag available in CRSP and by searching fund names for certain terms (e.g., "index"). As a whole, this process closely matches that used by KTWW.[11]

All of my analysis is conducted at the fund-level using monthly net returns. Individual shareclasses of a fund are identified using the WFICN variable available in the MFLINKS database. Fund-level returns are an asset-weighted average of the shareclass-level returns. The

[9] The full list of terms used in the fund name search is available upon request.
[10] This hurdle may seem somewhat low, but it is consistent with KTWW. My conclusions stay the same if I raise the hurdle (e.g., to 75% instead of 50%).
[11] CRSP objective codes were not available when KTWW performed their study. If, like KTWW, I form my sample using a combination of other companies' objective codes, my conclusions are unchanged. The exact set of objective code categories used by KTWW to identify a fund as domestic equity is uncertain.

only returns in the database that are discarded are those following a missing return, as the return following a missing return reflects the cumulative performance since the last nonmissing return.[12]

All three samples considered in this study (1975 to 2002, 2003 to 2017, and 1975 to 2017) are built using the same process described above. After imposing KTWW's requirement that each fund have at least 60 monthly returns in a given sample, my 1975 to 2002 sample has 1,825 unique funds.[13] An equal-weight portfolio of those funds has a four-factor alpha of −0.45% per year. In comparison, KTWW's 1975 to 2002 sample has 1,788 unique funds and an equal-weight portfolio alpha of −0.4% per year (see their Table 1). Hence, my replicated sample appears to be a close match for KTWW's original sample. In the extended sample, 2003 to 2017, there are 2,355 unique funds. The greater number of funds in the extended sample is consistent with documented industry growth.[14] The combined sample contains 3,151 unique funds.

3. Results

In this section, I first consider whether I can replicate KTWW's original results (1975 to 2002). I then consider whether those results hold out-of-sample (2003 to 2017) and in a combined sample (1975 to 2017).

3.1. Replicating Kosowski, Timmermann, Wermers, and White (2006)

An essential part of my study is evaluating whether I can replicate KTWW's results during KTWW's sample period (1975 to 2002). If my in-sample results do not reasonably match those presented by KTWW, then my out-of-sample results will be difficult to interpret. Specifically,

---

[12] If I discard returns that occur before a fund is at least 2 years old and has reached at least $20 million in assets (remaining thereafter regardless of future assets) to control for Evans (2010) incubation bias, then I find no evidence of outperformance that cannot be explained by luck, regardless of the sample. Without such a filter, at least some incubated returns are included in the analysis irrespective of the minimum return history required.

[13] KTWW analyze the potential survivorship bias created by this requirement and state that the "bias has almost no impact on [their] bootstrap results." If I change the requirement to 12 months to lessen any potential bias, my conclusions are not affected.

[14] Table 5 in the 2018 Investment Company Factbook shows that there were only 430 domestic equity funds in 1984 (the first year the data was available), compared to 3,201 in 2017.

without an in-sample match, it will be unclear whether a difference between KTWW's results and my out-of-sample results represents an (i) actual change or (ii) a deviation in procedure. Therefore, to make robust conclusions, it is vital that I be able to successfully replicate KTWW's results.

Table 1 shows different percentiles in the distribution of actual fund alphas. For each percentile, a *p*-value is also reported. For the below median percentiles, the *p*-values are associated with a null hypothesis of not underperforming (based on the true-zero-alpha distribution). For the above median percentiles, the *p*-values are associated with a null hypothesis of not outperforming. As an example, if the 95th percentile shows an alpha of 1.00% per month and a *p*-value of 0.01, then the conclusion is that the outperformance of the fund at that point in the distribution cannot be explained by luck alone (alternatively, there is only a 1% chance that alpha is a fluke). The results presented in KTWW's study (see their Table 2) are presented side-by-side with my replication. KTWW only reported monthly alphas to one decimal place, whereas my results report them to two.

[Table 1 about here]

While not exact, my replication closely matches KTWW's results and leads to the same conclusions. Both KTWW and my replication indicate lower alpha than would be expected based on luck alone within the bottom 50% of funds. At the 90th percentile and above, both sets of results identify better performance than would be expected based on luck alone. Consequently, KTWW and I both agree that "a sizable minority of managers pick stocks well enough to more than cover their costs" during the 1975 to 2002 period.[15] Given this match quality, any changes in results I

[15] Using the alpha T-statistic instead of alpha itself, I find outperformance starting at about the 97th percentile, whereas KTWW find such performance starting at about the 95th percentile. Thus, using that alternative measure, KTWW and I both agree that a sizable minority exists, but slightly disagree on the precise number of funds in the group.

find between the 1975 to 2002 period and the 2003 to 2017 period should be due to the results actually changing, not a deviation in procedure.

3.2. Extending Kosowski, Timmermann, Wermers, and White (2006)

I next consider whether KTWW's results hold out-of-sample (2003 to 2017) and in a combined sample (1975 to 2017).[16] The out-of-sample test evaluates whether KTWW's results have continued to hold over the 15 years since KTWW's sample ended. Put another way, does a group of mutual funds that can pick stocks well enough to more than cover costs still exist? The combined sample test evaluates whether KTWW's results still hold in general. That is, in the largest sample available today, does that same group of mutual funds exist?

To answer those questions, Table 2 repeats the analysis in Table 1 using the 2003 to 2017 sample and the 1975 to 2017 sample. As shown, there is no evidence in either sample that any fund performs better than would be expected based on luck alone. Even at the 99th percentile of each distribution, the *p*-value associated with a null hypothesis of not outperforming is 1.00.[17] Therefore, the clear conclusions based on these results are (i) that there no longer exists "a sizable minority of managers [who] pick stocks well enough to more than cover their costs" and (ii) that the sizable minority does not exist in the largest available sample.[18]

[Table 2 about here]

[16] If I split the out-of-sample period into three equal subperiods (2003 to 2007, 2008 to 2012, and 2013 to 2017), I find evidence of alpha that cannot be explained by luck alone starting at the 70th percentile in the earliest subperiod, while results for the latter two subperiods are similar to those reported for the full out-of-sample period.
[17] Using just the funds in the 1975 to 2002 sample, but performing the analysis using all returns from 1975 to 2017, I find similar results to those presented.
[18] If the tests in this section are repeated using the alpha T-statistic as the measure of performance, I still find no sizable minority in the 2003 to 2017 period. Using the 1975 to 2017 period, I find some limited evidence of outperformance that cannot be explained by luck alone at the 99th percentile (*p*-value of 0.08).

4. Conclusions

Fama and French (2010) remark that the performance of mutual funds from 1975 to 2002 is "irrelevant for today's investors." Confirming that idea, my out-of-sample extension of KTWW's results shows that while many mutual funds outperformed after costs from 1975 to 2002, such performance has been exceedingly rare or nonexistent during the subsequent 15 years. Furthermore, while it is notable that, during certain periods, some outperforming mutual funds existed, the full historical record available today shows little evidence of outperformance that cannot be attributed to luck. In short, KTWW's results have not continued stably beyond publication and, in some respects, are obsolete as of 2018. As discussed previously, the failure of KTWW's results to hold out-of-sample should not be used to draw overly broad conclusions (e.g., 'all fund managers are unskilled' or 'active management does not create value'); however, when citing KTWW, it is essential to understand and note that their results are time period dependent and not reflective of the modern state of the mutual fund industry.

References

Amihud, Yakov, and Ruslan Goyenko. 2013. Mutual fund's R^2 as predictor of performance. *Review of Financial Studies* 26, 667-694.

Avramov, Doron, and Russ Wermers. 2006. Investing in mutual funds when returns are predictable. *Journal of Financial Economics* 81, 339-377.

Barras, Laurent, Olivier Scaillet, and Russ Wermers. 2010. False discoveries in mutual fund performance: Measuring luck in estimated alphas. *Journal of Finance* 65, 179-216.

Berk, Jonathan, and Richard Green. 2004. Mutual fund flows and performance in rational markets. *Journal of Political Economy* 112, 1269-1295.

Bernstein, Peter. 1998. Where, oh where are the .400 hitters of yesteryear? *Financial Analysts Journal* 54, 6-14.

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

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2008. Liquidity and market efficiency. *Journal of Financial Economics* 87, 249-268.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2011. Recent trends in trading activity and market quality. *Journal of Financial Economics* 101, 243-263.

Conrad, Jennifer, Sunil Wahal, and Jin Xiang. 2015. High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics* 116, 271-291.

Cremers, Martijn, and Antti Petajisto. 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22, 3329-3365.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz. 2012. Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review* 2, 1-48.

Dong, Xi, Shu Feng, and Ronnie Sadka. 2018. Liquidity risk and mutual fund performance. *Management Science*, forthcoming.

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin. 2015. Managerial activeness and mutual fund performance. *Review of Asset Pricing Studies* 5, 156-184.

Dyck, Alexander, Karl Lins, and Lukasz Pomorski. 2013. Does active management pay? New international evidence. *Review of Asset Pricing Studies* 3, 200-228.

Evans, Richard. 2010. Mutual fund incubation. *Journal of Finance* 65, 1581-1611.

Fama, Eugene, and Kenneth French. 2010. Luck versus skill in the cross-section of mutual fund performance. *Journal of Finance* 65, 1915-1947.

Ferson, Wayne, and Rudi Schadt. 1996. Measuring fund strategy and performance in changing economic conditions. *Journal of Finance* 51, 425-461.

Harvey, Campbell, and Yan Liu. 2018. False (and missed) discoveries in financial economics. Working paper.

Hoberg, Gerard. Nitin Kumar, and Nagpurnanand Prabhala. 2018. Mutual fund competition, managerial skill, and alpha persistence. *Review of Financial Studies* 5, 1896-1929.

Huang, Lifa, Wayne Lee, and Craig Rennie. 2018. Selection and timing skill in bond mutual fund returns: Evidence from bootstrap simulations. Working paper.

Jones, Christopher, and Haitao Mo. 2018. Out-of-sample performance of mutual fund predictors. Working paper.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2005. On the industry concentration of actively managed equity mutual funds. *Journal of Finance* 60, 1983-2011.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2008. Unobserved actions of mutual funds. *Review of Financial Studies* 21, 2379-2416.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. 2006. Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance* 61, 2551-2595.

Kosowski, Robert. 2011. Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. *Quarterly Journal of Finance* 1, 607-664.

Newey, Whitney, and Kenneth West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.

Pastor, Lubos, Robert Stambaugh, and Lucian Taylor. 2015. Scale and skill in active management. *Journal of Financial Economics* 116, 23-45.

Pastor, Lubos, Robert Stambaugh, and Lucian Taylor. 2017. Do funds make more when they trade more? *Journal of Finance* 72, 1483-1528.

von Reibnitz, Anna. 2017. When opportunity knocks: Cross-sectional return dispersion and active fund performance. *Critical Finance Review* 6, 303-356.

Table 1: Replicating Kosowski, Timmermann, Wermers, and White (2006)

Description: This table shows key results from Table 2 of Kosowski, Timmermann, Wermers, and White (2006) and my replication of those results. The columns with the heading 'Original' provide KTWW's results, and the columns with the heading 'Replication' provide my replication. The columns labeled 'Alpha' show various percentiles in the distribution of mutual funds' four-factor alphas. The reported values are in percent per month. The columns labeled '*p*-value' show the cross-sectionally bootstrapped *p*-value associated with a null hypothesis of not underperforming (outperforming) for below (above) median percentiles. The period of analysis is 1975 to 2002. KTWW's sample contains 1,788 mutual funds, while my sample contains 1,825.

Interpretation: During the 1975 to 2002 period, mutual funds in the top 10% of the distribution of alpha show outperformance that cannot be explained by luck alone.

	Original		Replication	
Percentile	Alpha	*p*-value	Alpha	*p*-value
1%	-0.8%	0.02	-0.82%	<0.01
3%	-0.6%	<0.01	-0.55%	<0.01
5%	-0.5%	<0.01	-0.47%	<0.01
10%	-0.4%	<0.01	-0.35%	<0.01
20%	-0.2%	<0.01	-0.24%	<0.01
30%	-0.2%	<0.01	-0.17%	<0.01
40%	-0.1%	<0.01	-0.11%	<0.01
Median	-0.1%		-0.07%	
60%	0.0%	1.00	-0.02%	1.00
70%	0.1%	1.00	0.04%	1.00
80%	0.1%	0.99	0.12%	1.00
90%	0.3%	0.02	0.26%	0.02
95%	0.4%	<0.01	0.39%	0.03
97%	0.6%	<0.01	0.53%	<0.01
99%	1.0%	<0.01	0.88%	<0.01

Table 2: Extending Kosowski, Timmermann, Wermers, and White (2006)

Description: This table shows results from repeating the analysis from Table 2 of Kosowski, Timmermann, Wermers, and White (2006) using two different sample periods, 2003 to 2017 and 1975 to 2017. The columns with the heading '2003 to 2017' provide the former, and the columns with the heading '1975 to 2017' provide the latter. The columns labeled 'Alpha' show various percentiles in the distribution of mutual funds' four-factor alphas. The reported values are in percent per month. The columns labeled '*p*-value' show the cross-sectionally bootstrapped *p*-value associated with a null hypothesis of not underperforming (outperforming) for below (above) median percentiles. The 2003 to 2017 sample contains 2,355 mutual funds, while the 1975 to 2017 sample contains 3,151.

Interpretation: During the 2003 to 2017 and 1975 to 2017 periods, there is no outperformance by mutual funds that cannot be explained by luck alone.

	2003 to 2017		1975 to 2017	
Percentile	Alpha	*p*-value	Alpha	*p*-value
1%	-0.61%	<0.01	-0.74%	<0.01
3%	-0.43%	<0.01	-0.52%	<0.01
5%	-0.36%	<0.01	-0.43%	<0.01
10%	-0.29%	<0.01	-0.31%	<0.01
20%	-0.20%	<0.01	-0.21%	<0.01
30%	-0.15%	<0.01	-0.15%	<0.01
40%	-0.12%	<0.01	-0.10%	<0.01
Median	-0.09%		-0.07%	
60%	-0.06%	1.00	-0.03%	1.00
70%	-0.02%	1.00	0.01%	1.00
80%	0.01%	1.00	0.06%	1.00
90%	0.08%	1.00	0.14%	1.00
95%	0.13%	1.00	0.21%	1.00
97%	0.18%	1.00	0.25%	1.00
99%	0.26%	1.00	0.37%	1.00