The Persistence of Mutual Fund Performance

Mark Grinblatt; Sheridan Titman

The Journal of Finance, Vol. 47, No. 5. (Dec., 1992), pp. 1977-1984.

Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199212%2947%3A5%3C1977%3ATPOMFP%3E2.0.CO%3B2-1

The Journal of Finance is currently published by American Finance Association.

THE JOURNAL OF FINANCE • VOL. XLVII, NO. 5 • DECEMBER 1992

The Persistence of Mutual Fund Performance

MARK GRINBLATT and SHERIDAN TITMAN*

ABSTRACT

This paper analyzes how mutual fund performance relates to past performance. These tests are based on a multiple portfolio benchmark that was formed on the basis of securities characteristics. We find evidence that differences in performance between funds persist over time and that this persistence is consistent with the ability of fund managers to earn abnormal returns.

THERE IS A LARGE and growing industry devoted to measuring mutual fund performance. This industry, as evidenced by newsletters, regular features in the financial press, and the existence of firms that professionally evaluate funds, is based on the idea that funds that do well (or poorly) in the past will continue to do so in the future.

While the efficient markets hypothesis has generated interest among academics in testing whether funds exhibit performance, there has been very little research devoted to testing for persistence in fund performance.[1] One reason for this may be that the traditional benchmarks used to evaluate performance are known to exhibit persistent biases. For example, the CAPM and APT-based benchmarks favor small capitalization and high dividend-yield stocks. Thus, small-firm funds and income-oriented funds may appear to persistently outperform other funds when traditional benchmarks are used.

In a recent paper,[2] we introduce a new benchmark that does not exhibit these biases and find evidence of significant differences in performance between funds. Our tests[3] reject the hypothesis that all funds with growth and aggressive-growth investment objectives have equal performance. This rejection holds for both the actual returns of the funds and for portfolio returns constructed from the funds' quarterly holdings that proxy for the funds' gross returns (before transaction costs and fees).

* Anderson Graduate School of Management, UCLA. Titman is also from The Hong Kong University of Science and Technology. We wish to thank James Brandon and Erik Sirri for research assistance on earlier versions of this research and the editors and referees, as well as Russ Wermers, for their comments.

[1] Exceptions include Jensen (1969), Beebower and Bergstrom (1977), and Lehmann and Modest (1987). All of these studies are subject to a benchmark bias and the latter has no test statistics.

[2] Grinblatt and Titman (1989a).

[3] See Tables 2 and 3 of Grinblatt and Titman (1989a).

The F-test used in this earlier study, which is based on a nonspecific alternative hyopthesis, does not tell us if past performance is a useful guide for selecting a fund. If we are interested in learning whether past performance provides useful information to an investor, more specific tests of the relation between past performance and future performance are needed. Such tests are provided in this paper. In contrast to the earlier paper, we use a larger sample of funds and focus only on the actual returns of the funds.

The organization of the paper is as follows. Section I describes our sample of funds, the sample of passive portfolios we use as a control, and the benchmark we use to evaluate performance. Section II describes our persistence tests and presents the empirical results, and Section III concludes the paper.

I. The Sample of Fund Returns, the Control Sample, and the Benchmark

Mutual fund data, consisting of monthly cash-distribution-adjusted returns and investment goals for 279 funds that existed from December 31, 1974 to December 31, 1984, were purchased from CDA Investment Technologies, Inc. To our best knowledge, this is the largest sample of mutual funds evaluated in the academic literature. The data was spot checked with data collected by hand and found to be accurate.

As with most mutual fund studies, the mutual fund return data are subject to survivorship bias. Since CDA's nonacademic clients are only interested in mutual funds that they can invest in, funds that went out of business prior to December 31, 1984 are excluded from the CDA data set. Grinblatt and Titman (1989a) conclude that the survivorship bias is negligible for a sample that includes surviving and nonsurviving funds over this time period. Moreover, the survivorship requirement biases our tests against finding persistence.[4]

To check the reliability of our persistence tests, a control sample of 109 passive portfolios is constructed from the CRSP daily returns and daily master files. Each portfolio equally weights a subset of the CRSP securities and is rebalanced monthly. Seventy-two of the portfolios, each containing

[4] When the sample is split into two halves, we can characterize a fund's performance into four categories, (1) good in the first half, good in the second half, (2) good in the first half, bad in the second half, (3) bad in the first half, good in the second half, and (4) bad in the first half and bad in the second half. Cases (1) and (4) are cases of positive persistence, while cases (2) and (3) are indicative of negative persistence. If these cases are equally likely, one will find no persistence. Survivorship requirements, however, are most likely to eliminate funds in category (4). This would bias the remaining funds towards negative persistence. An alternative view, expressed by Brown, Boetzmann, Ibbotson, and Ross (1991), is that funds that do well in the first half are likely to be those that choose the riskiest strategies. This leads to a positive bias in measured persistence since the funds that do well in the past may be less likely to survive in the future since their riskier strategies make it more likely that they will do extremely poorly. A recent paper by Hendricks, Patel, and Zeckhauser (1991) examines persistence strategies on a sample without survival bias and concludes that the bias is small.

about $^1/_{12}$ of the CRSP securities, are formed on the basis of securities characteristics. These characteristics include firm size, dividend yield, coskewness with the monthly rebalanced equally weighted index, interest rate sensitivity, past returns over the previous three years, and beta computed against the equally weighted index. For each of the six characteristics, every CRSP security is grouped into one of twelve passive portfolios based on its ranking against other CRSP securities with respect to that characteristic. An additional 37 passive portfolios are formed on the basis of SIC industry groupings. Each of these is an equally weighted monthly rebalanced portfolio containing CRSP securities with the same "two-digit" SIC code as of the end of 1974. Only industries with at least 20 CRSP-listed firms are included.

The performance measure used in this study, an extension of the measure employed by Jensen (1968, 1969), is computed relative to the eight-portfolio benchmark, P8, used in Grinblatt and Titman (1989a). The basic idea underlying the formation of this benchmark is that various firm characteristics are correlated with their stocks' factor loadings. As a result, portfolios formed from stocks grouped by securities characteristics can be used as proxies for the factors. The P8 benchmark, constructed from groupings of the passive portfolios' returns described above, consists of four size-based portfolios, three dividend-yield-based portfolios, and the lowest past returns portfolio: The smallest $8\,^1/_3\%$ of firms comprise the first size-based portfolio; the average of the second- and third-smallest size portfolios (out of 12) comprise the second portfolio; the average of the fourth- through ninth-smallest size portfolios comprise the third portfolio; and the average of the three largest size portfolios comprise the fourth. The three portfolios formed from dividend-yield rankings are an equal weighting of the two lowest dividend-yield portfolios (out of 12), the fifth- and sixth-lowest dividend-yield portfolios, and the tenth and eleventh dividend-yield portfolios. To compute an abnormal return relative to the P8 benchmark, we use ordinary least squares to estimate the intercept in a time-series regression of the excess returns (above a one-month T-bill rate) of the fund or passive portfolio on the excess returns of the eight portfolios described above.[5]

II. Persistence Tests

The persistence of abnormal performance is analyzed with a three-step procedure. First, we split the ten-year sample of fund returns into two five-year subperiods. Second, we compute the abnormal returns of each fund for each five-year subperiod. Finally, we estimate the slope coefficient in a cross-sectional regression of abnormal returns computed from the last five years of data on abnormal returns computed from the first five years of data.

A significant positive t-statistic for the slope coefficient in this regression would reject the null hypothesis that past performance is unrelated to future

[5] For a theoretical discussion of the Jensen Measure and the inferences that can be drawn from it, see Grinblatt and Titman (1989b).

performance and support the alternative hypothesis that past performance is positively related to future performance. However, the standard method for computing the t-statistic for the slope coefficient in a cross-sectional regression results in a statistic that does not have a true t-distribution. This is because many funds have similar portfolios and hence have highly correlated residuals. To overcome this bias, we develop an alternative t-test that is derived from a time series procedure. The procedure for computing these "time-series t-statistics" is an extension of the technique introduced by Fama and MacBeth (1973) to overcome a similar problem in tests of the CAPM.

Define α_i to be the abnormal return of the ith fund computed from the first five years of returns in excess of the average abnormal return of all funds over these 60 months. By construction, the α_i's sum to zero, implying that a portfolio constructed with weights that are proportional to the α_i's has zero cost. Now consider the following weighted average of the (second-half) returns of the funds,

$$R_{pt} = \Sigma_i \alpha_i R_{it}/\text{var}(\alpha), \quad \text{where}$$

R_{it} = the return of the ith fund in month t, $t = 61, \ldots, 120$, and

var(α) = cross-sectional variance of the abnormal returns of the funds computed from the first five years times 279.

Since the weight on fund i is $\alpha_i/\text{var}(\alpha)$, we can regard $\tilde{R}_{pt}$ as the return (in the second five years) of a zero cost portfolio of the funds. The intercept from a multiple regression of R_{pt} on the excess returns of the eight portfolios in the P8 benchmark is the abnormal return of this portfolio (or, alternatively, the same weighted average of the abnormal returns of the individual funds). This intercept is algebraically identical to the least squares slope coefficient from the cross-sectional persistence regression. However, in contrast to the t-statistics estimated with the cross-sectional persistence regressions, the t-statistics generated with the time-series regression will not be biased under the null hypothesis that residuals are i.i.d. normal (an assumption that is better approximated with the time-series regression than with the cross-sectional regression).[6]

Panel A of Table I presents the intercept from this time-series regression along with the corresponding t-statistic. This persistence statistic indicates that mutual funds in the second five-year period are expected to realize a 0.28% greater abnormal return in the second five years for every 1% abnormal return achieved in the first five years. This coefficient is highly significant (at almost the 1% level in a two-tailed test).

As a control, we also performed the same procedure on the sample of 109 passive portfolios described earlier. The coefficient here, also reported in Panel A, is of a larger magnitude than the coefficient for the 279 mutual funds, but is statistically insignificant. To reconcile the larger coefficient

[6] For most of our results, the time-series t-statistic is about half the size of the biased t-statistic derived directly from the cross-sectional regression.

Table I

Regression Tests of the Persistence of Performance

Panel A: Slope Coefficient and *t*-Statistics[a] for Regressions of Abnormal Returns of 279 Mutual Funds and 109 Passive Portfolios in the Last 60 Months on Abnormal Returns from the First 60 Months

Sample	Cross-Sectional R^2	Slope Coefficient	*t*-Statistic
Mutual finds	0.06	0.281	2.64*
Passive portfolios	0.15	0.395	1.59

Panel B: Slope Coefficient and *t*-Statistics[a] for Regressions of the Abnormal Returns of 279 Mutual Funds and 109 Passive Test Portfolios in 60 Randomly Selected Months on the Abnormal Returns of the Portfolios from the Remaining Sixty Months

Sample	Cross-Sectional R^2	Slope Coefficient	*t*-Statistic
Mutual funds	0.13	0.420	4.59**
Passive portfolios	0.00	−0.053	−0.29

Panel C: Abnormal Returns and *t*-Statistics[a] in the Last Sixty Months (or Random Sixty Months) of Equally Weighted Portfolios of Mutual Funds Consisting of Either the Top or Bottom 10% of the Funds Based on Performance from the Other Half of the Sample

Ranking Period	Best Performers Abnormal Returns[b]	Worst Performers Abnormal Returns[b]	*t*-Statistic for Difference
First 60 months	0.0000 (0.02)	−0.0029 (−2.10)*	2.10*
Random 60 months	0.0046 (2.05)*	−0.0028 (−2.31)*	4.32**

[a] *t*-Statistic is unbiased *t*-statistic calculated with the time-series technique.
[b] *t*-Statistics are in parentheses.
* Significant at 0.05 level (two-tailed test).
** Significant at 0.01 level (two-tailed test).

estimated for the sample of passive portfolios with its statistical insignificance, we have to recognize that these passive portfolios were constructed to be very different while many of the mutual funds are very similar. As a result, the returns of the costless portfolio of funds used to calculate the time-series *t*-statistics have a much lower standard deviation than the analogous returns constructed from the passive portfolios.

Although the coefficient estimate in the regression for the passive portfolios in Panel A is not reliably different from zero, its magnitude makes us hesitant to conclude that the slope coefficient for the funds is entirely due to persistence of managerial talent. Moreover, recent work by Jegadeesh and Titman (1991) provides evidence of significant persistence in the long-run abnormal returns of individual stocks. Panel B of Table I addresses whether this is the source of persistence for the mutual funds. To obtain the coefficients in Panel B, we perform the same regressions as in Panel A, but randomly sort the 120 months of returns into two 60-month samples, separately sorting the Januaries so that each subsample would include the same number of Januaries (this allows us to avoid possible biases due to the January effect).

The results in Panel B from the randomly partitioned months are even more striking than those from the chronologically partitioned months in Panel A. First, there is no evidence of persistence in the passive portfolios. Moreover, persistence of the mutual funds is much stronger with a random partitioning of the monthly returns than with a chronological partitioning. The coefficient estimate, 0.42, is much larger and the t-statistic, 4.59, is highly significant. This indicates that persistence in individual securities returns is unlikely to be the source of the persistence observed for either the mutual funds or the passive funds in Panel A.

The results in Panels A and B for the mutual funds could also have been generated by persistent differences in fees and transaction costs across funds. To test this, Panel C examines the average abnormal returns of the 10% best- and worst-performing funds in a five-year period outside of the ranking period. The same chronological sorting and random sorting is used. If persistence is entirely due to fees and transaction costs, the best-performing funds from the ranking period would exhibit negative abnormal returns of approximately the same magnitude as their transaction costs in the test period. However, significantly positive abnormal returns for the funds ranked best would imply that there was persistence in abilities. The t-test in the difference column of Panel C also tests whether the abnormal returns of the 10% best-performing funds equals the abnormal performance of the 10% worst-performing funds in the five-year test period.

Panel C indicates that when the sample is split in half chronologically, the worst-performing funds demonstrate persistently poor performance; both the best- and worst-performing funds demonstrate persistence when the fund is split in half randomly. Funds in the lowest decile from 1975–1979 and in the random half of the sample used for ranking have abnormal performance of about -3.5 percent per year in the other half of the sample.[7]

In Panel C, the portfolio of funds with abnormal returns in the top decile for the randomly selected ranking months achieves an abnormal return of about 5.5 percent per year in the remaining sixty months. However, this result appears to be due to chance. Out of four additional random partitions of the 120 months, none resulted in significant out-of-sample abnormal

[7] This finding appears to be robust to changes in the random numbers generated by the computer. Four additional randomly generated sample splits were also examined. Three of the four had negative returns for the worst-performing funds of approximately the same magnitude as that for the first random split. These resulted in significant differences between the returns of the best- and worst-performing funds in the second half of the sample. In addition, the persistence regressions for the 109 passive portfolios are all insignificant with the P8 benchmark. The t-statistic for the difference between the 10% best and worst performers with the chronological sample split used in Panel C was 1.53. The comparable t-statistics for the five random splits were respectively: -0.74, -1.05, 0.12, 0.23, and -0.34.

returns for the best-performing funds from the ranking period. However, all of the out-of-sample abnormal returns are positive.[8]

III. Conclusion

The results presented in this paper indicate that there is positive persistence in mutual fund performance. The persistence cannot be explained by inefficiencies in the benchmark that are related to firm size, dividend yield, past returns, skewness, interest rate sensitivity, or CAPM beta. These findings are consistent with there being persistent differences in fees and transaction costs across funds, although the results in Panel C suggest that this is not the sole explanation for our results. Irrespective of the source or sources of the persistence that we find in this paper, we can assert that the past performance of a fund provides useful information for investors who are considering an investment in mutual funds.

One issue that we do not address in this paper is how to optimally weight information about past performance in selecting a mutual fund. A recent paper by Hendricks, Patel, and Zeckhauser (1991) argues that only the most recent past performance provides information about future performance. They present evidence of higher levels of abnormal performance for strategies that buy mutual funds based on their performance measured over the past 2 to 8 quarters.

[8] The results for the other four random partitions are as follows:

Ranking Period	Best Performers Abnormal Returns	Worst Performers Jensen Measure[b]	t-Statistic for Difference[a]
Random2			
60 months	0.0020 (0.91)	−0.0031 (−2.54)*	2.30*
Random3			
60 months	0.0018 (0.88)	−0.0037 (−3.21)**	3.26**
Random4			
60 months	0.0022 (1.19)	−0.0031 (−1.97)	2.80**
Random5			
60 months	0.0009 (0.48)	−0.0003 (−0.17)	0.42

(See Table I for explanation of symbols)

REFERENCES

Beebower, Gilbert L., and Gary L. Bergstrom, 1977, A performance analysis of pension and profit-sharing portfolios: 1966–1975, *Financial Analysts Journal* 33, 31–42.

Brown, Stephen, Will Goetzmann, Robert Ibbotson, and Stephen Ross,1991, Survivorship bias in performance studies, Working paper, Yale University.

Fama, Eugene, and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Grinblatt, Mark, and Sheridan Titman, 1989a, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business*, 62, 393–416.

———, 1989b, Portfolio performance evaluation: Old issues and new insights, *Review of Financial Studies*, 2, 393–421.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1991, Hot hands in mutual funds: Short-run persistence of performance, 1974–88, Working paper, Harvard University.

Jegadeesh, Narasimhan, and Sheridan Titman, 1991, Return to buying winners and selling losers: Implications for stock market efficiency, Working paper, UCLA.

Jensen, Michael, 1968, The Performance of mutual funds in the period 1945–1964, *Journal of Finance*, 23, 389–416.

———, 1969, Risk, the pricing of capital assets, and the evaluation of investment portfolios, *Journal of Business*, 42, 167–247.

Lehmann, Bruce, and David Modest, 1987, Mutual fund performance evaluations: A comparison of benchmarks and benchmark comparisons, *Journal of Finance*, 42, 233–265.