

On Persistence in Mutual Fund Performance

MARK M. CARHART*

ABSTRACT

Using a sample free of survivor bias, I demonstrate that common factors in stock returns and investment expenses almost completely explain persistence in equity mutual funds' mean and risk-adjusted returns. Hendricks, Patel and Zeckhauser's (1993) "hot hands" result is mostly driven by the one-year momentum effect of Jegadeesh and Titman (1993), but individual funds do not earn higher returns from following the momentum strategy in stocks. The only significant persistence not explained is concentrated in strong underperformance by the worst-return mutual funds. The results do not support the existence of skilled or informed mutual fund portfolio managers.

PERSISTENCE IN MUTUAL FUND performance does not reflect superior stock-picking skill. Rather, common factors in stock returns and persistent differences in mutual fund expenses and transaction costs explain almost all of the predictability in mutual fund returns. Only the strong, persistent underperformance by the worst-return mutual funds remains anomalous.

Mutual fund persistence is well documented in the finance literature, but not well explained. Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995), and Wermers (1996) find evidence of persistence in mutual fund performance over short-term horizons of one to three years, and attribute the persistence to "hot hands" or common investment strategies. Grinblatt and Titman (1992), Elton, Gruber, Das, and Hlavka (1993), and Elton, Gruber, Das, and Blake (1996) document mutual fund return predictability over longer horizons of five to ten years, and attribute this to manager differential information or stock-picking talent. Contrary evidence comes from Jensen (1969), who does not find that good subsequent performance follows good past performance. Carhart (1992) shows that persistence in expense ratios drives much of the long-term persistence in mutual fund performance.

My analysis indicates that Jegadeesh and Titman's (1993) one-year momentum in stock returns accounts for Hendricks, Patel, and Zeckhauser's (1993) hot hands effect in mutual fund performance. However, funds that earn higher

* School of Business Administration, University of Southern California. I have benefited from helpful conversations with countless colleagues and participants at various workshops and seminars. I express particular thanks to Gene Fama, my dissertation committee chairman. I am also grateful to Gene Fama, the Oscar Mayer Fellowship, and the Dimensional Fund Advisors Fellowship for financial support. I thank Cliff Asness, Gene Fama, Ken French, and Russ Wermers for generously providing data. Finally, I thank Bill Crawford, Jr., Bill Crawford, Sr., and ICDI/Micropal for access to, and assistance with, their database.

one-year returns do so not because fund managers successfully follow momentum strategies, but because some mutual funds just happen by chance to hold relatively larger positions in last year's winning stocks. Hot-hands funds infrequently repeat their abnormal performance. This is in contrast to Wermers (1996), who suggests that it is the momentum strategies themselves that generate short-term persistence, and Grinblatt, Titman, and Wermers (1995), who find that funds following momentum strategies realize better performance before management fees and transaction expenses. While measuring whether funds follow the momentum strategy is imperfect in my sample, individual mutual funds that appear to follow the one-year momentum strategy earn significantly lower abnormal returns after expenses. Thus, I conclude that transaction costs consume the gains from following a momentum strategy in stocks.

I demonstrate that expenses have at least a one-for-one negative impact on fund performance, and that turnover also negatively impacts performance. By my estimates, trading reduces performance by approximately 0.95 percent of the trade's market value. Variation in costs *per* transaction across mutual funds also explains part of the persistence in performance. In addition, I find that fund performance and load fees are strongly and negatively related, probably due to higher total transaction costs for load funds. Holding expense ratios constant, load funds underperform no-load funds by approximately 80 basis points per year. (This figure ignores the load fees themselves.)

The joint-hypothesis problem of testing market efficiency conditional on the imposed equilibrium model of returns clouds what little evidence there is in this article to support the existence of mutual fund manager stock-picking skill. Funds with high past alphas demonstrate relatively higher alphas and expected returns in subsequent periods. However, these results are sensitive to model misspecification, since the same model is used to rank funds in both periods. In addition, these funds earn expected future alphas that are insignificantly different from zero. Thus, the best past-performance funds appear to earn back their expenses and transaction costs even though the majority underperform by approximately their investment costs.

This study expands the existing literature by controlling for survivor bias, and by documenting common-factor and cost-based explanations for mutual fund persistence. Section I discusses the database and its relation to other survivor-bias corrected data sets. Section II presents models of performance measurement and their resulting pricing error estimates on passively-managed benchmark equity portfolios. Section III documents and explains the one-year persistence in mutual fund returns, and Section IV further interprets the results. Section V examines and explains longer-term persistence, and Section VI concludes.

I. Data

My mutual fund database covers diversified equity funds monthly from January 1962 to December 1993. The data are free of survivor bias, since they

Table I
Mutual Fund Database Summary Statistics

The table reports time-series averages of annual cross-sectional averages from 1962 to 1993. TNA is total net assets, Flow is the percentage change in TNA adjusted for investment return and mutual fund mergers. Exp ratio is total annual management and administrative expenses divided by average TNA. Mturn is modified turnover and represents reported turnover plus 0.5 times Flow. Maximum load is the total of maximum front-end, rear-end, and deferred sales charges as a percentage of the investment. Live funds are those in operation at the end of the sample, December 31, 1993. Dead funds are those that discontinued operations prior to this date.

Time-Series Averages of Cross-Sectional Average Annual Attributes, 1962–1993									
Group	Total Number	Avg Number	Avg TNA (\$ millions)	Avg Flow (%/year)	Avg Exp Ratio (%/year)	Avg Mturn (%/year)	Percentage with Load	Avg Max Load	Avg Age (years)
All funds	1,892	509.1	\$217.8	3.4%	1.14%	77.3%	64.5%	7.33%	18.1
By fund category									
Aggressive growth	675	169.2	\$ 95.6	5.0%	1.55%	99.7%	58.2%	7.38%	12.3
Long term growth	618	168.5	\$221.4	5.5%	1.09%	79.5%	59.7%	7.38%	16.4
Growth & income	599	171.4	\$328.5	1.5%	0.91%	60.9%	70.0%	7.27%	23.8
By current status									
Live funds	1,310	352.3	\$268.7	4.3%	1.07%	76.2%	63.3%	7.29%	19.2
Dead funds	582	156.8	\$ 46.8	–1.2%	1.44%	83.1%	68.5%	7.44%	14.9

include all known equity funds over this period. I obtain data on surviving funds, and for funds that have disappeared since 1989, from Micropal/Investment Company Data, Inc. (ICDI). For all other nonsurviving funds, the data are collected from *FundScope Magazine*, *United Babson Reports*, *Wiesenberger Investment Companies*, the *Wall Street Journal*, and past printed reports from ICDI. See Carhart (1995a) for a more detailed description of database construction.

Table I reports summary statistics on the mutual fund data. My sample includes a total of 1,892 diversified equity funds and 16,109 fund years. The sample omits sector funds, international funds, and balanced funds. The remaining funds are almost equally divided among aggressive growth, long-term growth, and growth-and-income categories. In an average year, the sample includes 509 funds with average total net assets (TNA) of \$218 million and average expenses of 1.14 percent per year. In addition, funds trade 77.3 percent of the value of their assets (Mturn) in an average year. Since reported turnover is the *minimum* of purchases and sales over average TNA, I obtain Mturn by adding to reported turnover one-half of the percentage change in TNA adjusted for investment returns and mergers. Also, over the full sample, 64.5 percent of funds charge load fees, which average 7.33 percent.

By December 31, 1993, about one-third of the total funds in my sample had ceased operations, so a sizeable portion of the database is not observable in most commercially available mutual fund databases. Thus, survivor bias is an important issue in mutual fund research. (See Brown, Goetzmann, Ibbotson, and Ross (1992), Carhart (1995b), and Wermers (1996).) While my sample is, to my knowledge, the largest and most complete survivor-bias-free mutual fund database currently available, Grinblatt and Titman (1989), Malkiel (1995), Brown and Goetzmann (1995), and Wermers (1996) use similar databases to study mutual funds. Grinblatt and Titman (1989) and Wermers (1996) use quarterly "snapshots" of the mutual funds' underlying stock holdings since 1975 to estimate returns gross of transactions costs and expense ratios, whereas my data set uses only the net returns. Malkiel (1995) uses quarterly data from 1971 to 1991, obtained from Lipper Analytical Services. Although Malkiel studies diversified equity funds, his data set includes about 100 fewer funds each year than mine, raising the possibility of some selection bias in the Lipper data set. (We both exclude balanced, sector, and international funds.) Nonetheless, Malkiel's mean mutual fund return estimate from 1982 to 1990, 12.9 percent, is very close to the 13 percent that I find.

Brown and Goetzmann (1995) study a sample of mutual funds very similar to mine, but calculate their returns differently. Their sample is from the *Wiesenberger Investment Companies* annual volumes from 1976 to 1988. They calculate annual returns from the changes in net asset value per share (NAV), and income and capital gains distributions reported annually in *Wiesenberger*. As Brown and Goetzmann acknowledge, their data suffer from some selection bias, because the first years of new funds and last years of dead funds are missing. In addition, because funds voluntarily report this information to *Wiesenberger*, some funds may not report data in years of poor performance. Working in the opposite direction, Brown and Goetzmann calculate return as the sum of the percentage change in NAV (adjusted for capital gains distributions when available) and percentage income return. This procedure biases their return estimates downward somewhat, since it ignores dividend reinvestment. My data set mitigates these problems because I obtain monthly total returns from multiple sources and so have very few missing returns. In addition, I obtain from ICDI the reinvestment NAVs for capital gains and income distributions. Over the 1976 to 1988 period, Brown and Goetzmann report a mean annual return estimate of 14.5 percent, very close to the 14.3 percent in my data set. By these calculations, selection bias accounts for at least 20 basis points per year in Brown and Goetzmann's sample. It could be somewhat more, however, due to the downward bias in their return calculations.

II. Models of Performance Measurement

I employ two models of performance measurement: the Capital Asset Pricing Model (CAPM) described in Sharpe (1964) and Lintner (1965), and my (Carhart (1995)) 4-factor model. This section briefly describes these models, and

evaluates their performance estimates on quantitatively-managed portfolios of New York Stock Exchange (NYSE), American Stock Exchange (Amex), and Nasdaq stocks. For comparative purposes, this section also reports performance estimates from Fama and French's (1993) 3-factor model.¹

I construct my 4-factor model using Fama and French's (1993) 3-factor model plus an additional factor capturing Jegadeesh and Titman's (1993) one-year momentum anomaly.² The 4-factor model is consistent with a model of market equilibrium with four risk factors. Alternately, it may be interpreted as a performance attribution model, where the coefficients and premia on the factor-mimicking portfolios indicate the proportion of mean return attributable to four elementary strategies: high versus low beta stocks, large versus small market capitalization stocks, value versus growth stocks, and one-year return momentum versus contrarian stocks. I employ the model to "explain" returns, and leave risk interpretations to the reader.

I estimate performance relative to the CAPM, 3-factor, and 4-factor models as

$$r_{it} = \alpha_{iT} + \beta_{iT}\text{VWRF}_t + e_{it} \quad t = 1, 2, \dots, T \quad (1)$$

$$r_{it} = \alpha_{iT} + b_{iT}\text{RMRF}_t + s_{iT}\text{SMB}_t + h_{iT}\text{HML}_t + e_{it} \quad t = 1, 2, \dots, T \quad (2)$$

$$r_{it} = \alpha_{iT} + b_{iT}\text{RMRF}_t + s_{iT}\text{SMB}_t + h_{iT}\text{HML}_t + p_{iT}\text{PR1YR}_t + e_{it} \quad t = 1, 2, \dots, T \quad (3)$$

where r_{it} is the return on a portfolio in excess of the one-month T-bill return; VWRF is the excess return on the CRSP value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks; RMRF is the excess return on a value-weighted aggregate market proxy; and SMB, HML, and PR1YR are returns on value-weighted, zero-investment, factor-mimicking portfolios for size, book-to-market equity, and one-year momentum in stock returns.³

Summary statistics on the factor portfolios reported in Table II indicate that the 4-factor model can explain considerable variation in returns. First, note the relatively high variance of the SMB, HML, and PR1YR zero-investment portfolios and their low correlations with each other and the market proxies. This suggests the 4-factor model can explain sizeable time-series variation. Second,

¹ I find (Carhart (1995a)) that 3-factor performance estimates on mutual funds are more precise, but generally not economically different from the CAPM. Estimates from the 4-factor model frequently differ, however, due to significant loadings on the one-year momentum factor.

² This is motivated by the 3-factor model's inability to explain cross-sectional variation in momentum-sorted portfolio returns (Fama and French (1996)). Chan, Jegadeesh, and Lakonishok (1996) suggest that the momentum anomaly is a market inefficiency due to slow reaction to information. However, the effect is robust to time-periods (Jegadeesh and Titman (1993)) and countries (Asness, Liew, and Stevens (1996)).

³ SMB and HML are obtained from Gene Fama and Ken French. I construct PR1YR as the equal-weight average of firms with the highest 30 percent eleven-month returns lagged one month minus the equal-weight average of firms with the lowest 30 percent eleven-month returns lagged one month. The portfolios include all NYSE, Amex, and Nasdaq stocks and are re-formed monthly.

Table II
Performance Measurement Model Summary Statistics, July 1963 to December 1993

VWRF is the Center for Research in Security Prices (CRSP) value-weight stock index minus the one-month T-bill return. RMRF is the excess return on Fama and French's (1993) market proxy. SMB and HML are Fama and French's factor-mimicking portfolios for size and book-to-market equity. PR1YR is a factor-mimicking portfolio for one-year return momentum.

Factor Portfolio	Monthly Excess Return	Std Dev	<i>t</i> -stat for Mean = 0	Cross-Correlations				
				VWRF	RMRF	SMB	HML	PR1YR
VWRF	0.44	4.39	1.93	1.00				
RMRF	0.47	4.43	2.01	1.00	1.00			
SMB	0.29	2.89	1.89	0.35	0.32	1.00		
HML	0.46	2.59	3.42	-0.36	-0.37	0.10	1.00	
PR1YR	0.82	3.49	4.46	0.01	0.01	-0.29	-0.16	1.00

the high mean returns on SMB, HML, and PR1YR suggest that these three factors could account for much cross-sectional variation in the mean return on stock portfolios. In addition, the low cross-correlations imply that multicollinearity does not substantially affect the estimated 4-factor model loadings.

In tests not reported, I find that the 4-factor model substantially improves on the average pricing errors of the CAPM and the 3-factor model.⁴ I estimate pricing errors on 27 quantitatively-managed portfolios of stocks from Carhart, Krail, Stevens, and Welch (1996), where the portfolios are formed on the market value of equity, book-to-market equity and trailing eleven-month return lagged one month. Not surprisingly, the 3-factor model improves on the average pricing errors from the CAPM, since it includes both size and book-to-market equity factors. However, the 3-factor model errors are strongly negative for last year's loser stock portfolios and strongly positive for last year's winner stock portfolios. In contrast, the 4-factor model noticeably reduces the average pricing errors relative to both the CAPM and the 3-factor model. For comparative purposes, the mean absolute errors from the CAPM, 3-factor, and 4-factor models are 0.35 percent, 0.31 percent, and 0.14 percent per month, respectively. In addition, the 4-factor model eliminates almost all of the patterns in pricing errors, indicating that it well describes the cross-sectional variation in average stock returns.

III. Persistence in One-Year Return-Sorted Mutual Fund Portfolios

A. Common-Factor Explanations of One-Year Mutual Fund Persistence

In this section, I form portfolios of mutual funds on lagged one-year returns and estimate performance on the resulting portfolios, thus replicating the

⁴ These results are not included for sake of brevity, but are available from the author upon request.

methodology of Hendricks, Patel, and Zeckhauser (1993). On January 1 of each year, I form ten equal-weighted portfolios of mutual funds, using reported returns. Reported returns are net of all operating expenses (expense ratios) and security-level transaction costs, but do not include sales charges. I hold the portfolios for one year, then re-form them. This yields a time series of monthly returns on each decile portfolio from 1963 to 1993. Funds that disappear during the course of the year are included in the equal-weighted average until they disappear, then the portfolio weights are readjusted appropriately. For added detail, I subdivide the top and bottom portfolios into thirds.

The portfolios of mutual funds sorted on one-year past returns demonstrate strong variation in mean return, as shown in Table III. The post-formation monthly excess returns on the decile portfolios decrease nearly monotonically in portfolio rank, and indicate a sizeable annualized spread of approximately 8 percent. (This spread is 24 percent in the ranking year.) The subdivided extreme portfolios exhibit even larger return spreads. Portfolio 1A, which contains the top thirtieth of funds (14 funds on average), outperforms portfolio 10C, the bottom thirtieth of funds, by 1 percent per month. Cross-sectional variation in return is considerably larger among the previous year's worst performing funds than the previous year's best funds. The subportfolios of the top decile show a modest spread of 12 basis points per month (63 to 75), but the spread in the bottom decile is a substantial 50 basis points. Further, the bottom thirtieth of the previous year's funds seem to demonstrate anomalously poor returns. In the year after their bottom-decile ranking, these funds show high variance and still underperform T-bills by 25 basis points per month.

The CAPM does not explain the relative returns on these portfolios. The CAPM betas on the top and bottom deciles and subdeciles are virtually identical, so the CAPM alphas reproduce as much dispersion as simple returns. In addition, the performance estimates from the CAPM indicate sizeable positive abnormal returns of about 22 basis points per month (2.6 percent per year) for the previous year's top-decile funds, and even larger negative abnormal returns of about 45 basis points per month (5.4 percent per year) for the bottom-decile funds. If the CAPM correctly measures risk, both the best and worst mutual funds possess differential information, yet the worst funds appear to use this information perversely to reduce performance.

In contrast to the CAPM, the 4-factor model explains most of the spread and pattern in these portfolios, with sensitivities to the size (SMB) and momentum (PR1YR) factors accounting for most of the explanation. The top decile portfolios appear to hold more small stocks than the bottom deciles. More important, however, is the pronounced pattern in the funds' PR1YR coefficients. The returns on the top decile funds are strongly, positively correlated with the one-year momentum factor, while the returns in the bottom decile are strongly, negatively correlated with the factor. Of the 67-basis-point spread in mean monthly return between deciles 1 and 10, the momentum factor explains 31 basis points, or almost half. Further, of the 28-basis-point spread in monthly return not explained by the 4-factor model, the spread between the ninth and tenth deciles accounts for 20 basis points. Except for the relative underperfor-

Table III

Portfolios of Mutual Funds Formed on Lagged 1-Year Return

Mutual funds are sorted on January 1 each year from 1963 to 1993 into decile portfolios based on their previous calendar year's return. The portfolios are equally weighted monthly so the weights are readjusted whenever a fund disappears. Funds with the highest past one-year return comprise decile 1 and funds with the lowest comprise decile 10. Deciles 1 and 10 are further subdivided into thirds on the same measure. VWRF is the excess return on the CRSP value-weight market proxy. RMRF, SMB, and HML are Fama and French's (1993) market proxy and factor-mimicking portfolios for size and book-to-market equity. PR1YR is a factor-mimicking portfolio for one-year return momentum. Alpha is the intercept of the Model. The t-statistics are in parentheses.

Portfolio	Monthly		CAPM			4-Factor Model					
	Excess Return	Std Dev	Alpha	VWRF	Adj R-sq	Alpha	RMRF	SMB	HML	PR1YR	Adj R-Sq
1A	0.75%	5.45%	0.27% (2.06)	1.08 (35.94)	0.777	-0.11% (-1.11)	0.91 (37.67)	0.72 (19.95)	-0.07 (-1.65)	0.33 (11.53)	0.891
1B	0.67%	4.94%	0.22% (2.00)	1.00 (39.68)	0.809	-0.10% (-1.08)	0.86 (40.66)	0.59 (18.47)	-0.05 (-1.38)	0.27 (10.63)	0.898
1C	0.63%	4.95%	0.17% (1.70)	1.02 (44.65)	0.843	-0.15% (-1.92)	0.89 (49.76)	0.56 (20.86)	-0.05 (-1.61)	0.27 (12.69)	0.927
1 (high)	0.68%	5.04%	0.22% (2.10)	1.03 (43.11)	0.834	-0.12% (-1.60)	0.88 (50.54)	0.62 (23.67)	-0.05 (-1.86)	0.29 (13.88)	0.933
2	0.59%	4.72%	0.14% (1.75)	1.01 (57.00)	0.897	-0.10% (-1.78)	0.89 (66.47)	0.46 (22.95)	-0.05 (-2.25)	0.20 (12.43)	0.955
3	0.43%	4.56%	-0.01% (-0.08)	0.99 (70.96)	0.931	-0.18% (-3.65)	0.90 (76.80)	0.34 (18.99)	-0.07 (-3.69)	0.16 (11.52)	0.963
4	0.45%	4.41%	0.02% (0.33)	0.97 (85.70)	0.952	-0.12% (-2.81)	0.90 (90.03)	0.27 (18.18)	-0.05 (-3.12)	0.11 (9.40)	0.971
5	0.38%	4.35%	-0.05% (-1.10)	0.96 (93.93)	0.960	-0.14% (-3.31)	0.90 (89.65)	0.22 (14.42)	-0.05 (-3.27)	0.07 (6.18)	0.970
6	0.40%	4.36%	-0.02% (-0.46)	0.96 (91.94)	0.958	-0.12% (-2.82)	0.90 (86.16)	0.22 (14.02)	-0.04 (-2.37)	0.08 (6.01)	0.968
7	0.36%	4.30%	-0.06% (-1.39)	0.95 (92.90)	0.959	-0.14% (-3.09)	0.90 (85.73)	0.21 (13.17)	-0.03 (-1.62)	0.04 (2.89)	0.967
8	0.34%	4.48%	-0.10% (-1.86)	0.98 (85.14)	0.951	-0.13% (-2.52)	0.93 (75.44)	0.20 (10.74)	-0.06 (-3.16)	0.01 (0.84)	0.958
9	0.23%	4.60%	-0.21% (-3.24)	1.00 (67.91)	0.926	-0.20% (-3.11)	0.93 (60.44)	0.22 (9.69)	-0.10 (-3.80)	-0.02 (-1.17)	0.938
10 (low)	0.01%	4.90%	-0.45% (-4.58)	1.02 (46.09)	0.851	-0.40% (-4.33)	0.93 (42.23)	0.32 (9.69)	-0.08 (-2.23)	-0.09 (-3.50)	0.887
10A	0.25%	4.78%	-0.19% (-2.05)	1.00 (48.48)	0.864	-0.19% (-2.16)	0.91 (42.99)	0.33 (10.27)	-0.11 (-3.20)	-0.02 (-0.76)	0.891
10B	0.02%	4.92%	-0.42% (-3.84)	1.00 (40.67)	0.817	-0.37% (-3.45)	0.91 (35.52)	0.32 (8.24)	-0.09 (-2.16)	-0.09 (-2.99)	0.848
10C	-0.25%	5.44%	-0.74% (-5.06)	1.05 (32.16)	0.736	-0.64% (-4.49)	0.98 (28.82)	0.32 (6.29)	-0.04 (-0.73)	-0.17 (-4.09)	0.782
1-10 spread	0.67%	2.71%	0.67% (4.68)	0.01 (0.39)	-0.002	0.29% (2.13)	-0.05 (-1.52)	0.30 (6.30)	0.03 (0.53)	0.38 (10.07)	0.231
1A-10C spread	1.01%	3.87%	1.00% (4.90)	0.02 (0.42)	-0.002	0.53% (2.72)	-0.07 (-1.61)	0.40 (5.73)	-0.02 (0.32)	0.50 (8.98)	0.197
9-10 spread	0.22%	1.22%	0.23% (3.64)	-0.02 (-1.60)	0.004	0.20% (3.13)	-0.01 (-0.40)	-0.10 (-4.30)	-0.01 (-0.60)	0.07 (3.87)	0.118

mance by last year's worst performing funds, the 4-factor model accounts for almost all of the cross-sectional variation in expected return on portfolios of mutual funds sorted on lagged one-year return.

I also perform the Spearman nonparametric test on the rank ordering of performance measures. Here the null hypothesis is that the performance measures are randomly ordered. The Spearman test falls in the 5.7 percent fractile on the CAPM alphas and the 13.2 percent fractile on the 4-factor alphas. In both cases, random rank-ordering cannot be rejected. However, since the Spearman test treats the ordering of each decile portfolio equally, it lacks power against the hypothesis that predictability in performance is concentrated in the tails of the distribution of mutual fund returns.⁵

B. Characteristics of the Mutual Fund Portfolios

I now examine whether any of the remaining short-term persistence in mutual fund returns is related to heterogeneity in the average characteristics of the mutual funds in each decile portfolio. In each year, I calculate a cross-sectional average for each decile portfolio of fund age, total net assets (TNA), expense ratio, turnover (Mturn), and maximum load fees.

The average portfolio characteristics reported in Table IV indicate that expenses and turnover are related to performance. Decile 10 particularly stands out with higher than average expenses and turnover. The 70-basis-point difference in expense ratios between deciles 9 and 10 explains about six of the 20-basis-point spread between monthly 4-factor alphas on these portfolios. It does not appear that fund age, size, or load fees can explain the large spread in performance on these portfolios, since these characteristics are very similar for the top and bottom deciles.

Differences in portfolio turnover do not explain a sizeable portion of the remaining portfolio nine-ten spread in alphas. If funds pay 1 percent in costs per round-trip transaction, the difference in trading frequency between the ninth and tenth deciles accounts for only 0.5 basis points of the spread in 4-factor alphas. After accounting for expense ratios and turnover, tests on the difference between alphas on deciles 9 and 10 yield *t*-statistics of 2.69 relative to the CAPM, and 2.19 relative to the 4-factor model. Thus, expense ratios and

⁵ I also test the robustness of these findings to time period, performance measurement benchmark, and sorting procedure. In these tests, not reported but available from the author upon request, I divide the complete sample into three equal subperiods, with insignificant effects on the results. I also estimate performance on the mutual fund portfolios using the alternative performance measures of Ferson and Schadt (1996), Chen and Knez (1995), and Carhart, Krail, Stevens, and Welch (1996). Ferson and Schadt model time-variation in factor risk loadings as linear functions of instrumental variables. The Chen and Knez nonparametric method and the linear factor pricing kernel approach in Carhart et al. are cross-section estimators of the stochastic discount factor. Carhart et al. also permit time-variation in model parameters. The estimates from these methods do not change any inferences. As a final robustness check, I perform tests on the investment objective categories separately (aggressive growth, long-term growth, and growth-and-income) and find that the persistence evidence is virtually as strong in each objective category as the diversified equity universe as a whole.

Table IV

Characteristics of the Portfolios of Mutual Funds Formed on Lagged 1-Year Return

Mutual funds are sorted annually from 1963 to 1993 into equal-weight decile portfolios based on lagged one-year return. Funds with the highest past one-year return comprise decile 1, and funds with the lowest comprise decile 10. Deciles 1 and 10 are further subdivided into thirds on the same measure. The values in the table represent the time-series averages of annual cross-sectional averages of the funds in each portfolio. TNA is total net assets. Expense ratio is management, administrative, and 12b-1 expenses divided by average TNA. Mturn is modified turnover and represents reported turnover plus 0.5 times the percentage change in portfolio TNA adjusted for investment returns and mergers. Maximum load is the sum of maximum front-end, back-end, and deferred sales charges.

Average Annual Portfolio Attributes					
Portfolio	Age (years)	TNA (\$ millions)	Expense Ratio	Mturn	Maximum Load
1A	11.7	110.0	1.38	116.2	3.93
1B	14.0	148.8	1.16	86.9	3.99
1C	16.5	127.4	1.11	75.8	4.62
1 (high)	14.1	128.7	1.22	92.9	4.18
2	16.6	190.8	1.08	75.3	4.97
3	17.3	194.3	1.10	76.3	4.72
4	17.6	183.7	1.11	67.2	4.82
5	18.3	185.9	1.09	68.4	4.71
6	17.5	199.1	1.15	65.8	4.33
7	18.3	169.7	1.14	62.2	4.50
8	17.5	149.3	1.13	65.3	4.76
9	15.8	145.6	1.22	75.1	4.59
10 (low)	13.6	77.1	1.92	81.4	4.38
10A	14.5	91.9	1.55	76.8	4.55
10B	14.4	87.4	1.71	76.7	4.57
10C	11.9	52.0	2.51	88.8	4.02

turnover alone cannot explain the anomalous negative abnormal performance by the worst-return decile of funds. This conclusion is even stronger when considering portfolio 10C, the bottom thirtieth of funds.

C. Characteristics of Individual Mutual Funds

Mutual fund managers claim that expenses and turnover do not reduce performance, since investors are paying for the quality of the manager's information, and because managers trade only to increase expected returns net of transactions costs. Thus, expenses and turnover should not have a direct negative effect on performance, as implied in the previous section, but rather a neutral or positive effect. I evaluate this claim by directly measuring the marginal effect of these and other variables on abnormal performance.

In each month, I estimate the cross-section regression:

$$\alpha_{it} = a_t + b_i x_{it} + \varepsilon_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T \quad (4)$$

where α_{it} is an individual fund performance estimate and x_{it} is a fund characteristic. As in Fama and MacBeth (1973), I estimate the cross-sectional relation each month, then average the coefficient estimates across the complete sample period. This yields 330 cross-sectional regressions which average 350 observations each for a combined sample of about 116,000 observations. To mitigate look-ahead bias, I estimate α_{it} as a one-month abnormal return from the 4-factor model, where the 4-factor model loadings are estimated over the prior three years:

$$\alpha_{it} \equiv R_{it} - R_{ft} - \hat{b}_{it-1} \text{RMRF}_t - \hat{s}_{it-1} \text{SMB}_t - \hat{h}_{it-1} \text{HML}_t + \hat{p}_{it-1} \text{PR1YR}_t. \quad (5)$$

I estimate one-month alphas each month on every fund, using a minimum of 30 observations, then estimate the cross-section relation of equation (4) using the Fama-MacBeth estimator.

The explanatory variables in equation (4) are expense ratio, turnover (Mturn), ln(TNA), and maximum load fees. Since I intend to explain performance, not predict it, I measure expense ratio and turnover contemporaneous with return. TNA is lagged one year to avoid spurious correlation (Granger and Newbold, 1974). Load fees are lagged one year to avoid the confounding possibility that funds change these fees in response to performance. I construct two additional explanatory variables from turnover to separate the effects of buy and sell trading. The latter two are

$$\text{Buy Turnover}_{it} = \text{Turnover}_{it} + \max(\text{Mflow}_{it}, 0)$$

and

$$\text{Sell Turnover}_{it} = \text{Turnover}_{it} - \min(\text{Mflow}_{it}, 0)$$

where Mflow_{it} measures the percentage change in TNA adjusted for investment returns and mergers. Because I find expense ratios are strongly related to the other variables, I estimate the cross-section regression for TNA, load, and the turnover measures using returns after adding back expense ratios.

The results in Table V indicate a strong relation between performance and size, expense ratios, turnover, and load fees. The resulting relation between performance and expense ratios and modified turnover suggest that mutual funds, on average, do not recoup their investment costs through higher returns. The -1.54 coefficient on expense ratio implies that for every 100-basis-point increase in expense ratios, annual abnormal return drops by about 154 basis points. The turnover coefficient of -0.95 suggests that for every 100-basis-point increase in turnover, annual abnormal return drops by about 95 basis points. We can interpret the turnover coefficient as a measure of the net costs of trading, since it reveals the marginal performance effect of a small change in turnover. Thus, the turnover estimate implies transactions costs of 95 basis points per round-trip transaction. When partitioned into *buy turnover*

Table V

Fama-MacBeth (1973) Cross-Sectional Regressions

Estimated univariate cross-sectional regressions for each month from July 1966 to December 1993 across all funds in the sample at that time. The dependent variable is the monthly residual from the 4-factor model, where the factor loadings are estimated on the prior 3 years of monthly returns. The independent variables are expense ratio, turnover, the natural log of TNA, maximum load fees, and measures of buy and sell turnover. Expense ratio is management, administrative, and 12b-1 expenses divided by average TNA. TNA is total net assets. Turnover represents reported turnover plus 0.5 times the percentage change in portfolio TNA adjusted for investment returns and mergers. Maximum load is the sum of maximum front-end, back-end and deferred sales charges. Buy turnover is reported turnover plus the maximum of 0 and the percentage change in TNA adjusted for investment returns and mergers. Sell turnover is reported turnover minus the minimum of 0 and the adjusted percentage change in TNA. Expense ratio and the turnover measures are divided by 12 and measured contemporaneous with the dependent variable. The reported estimates are time-series averages of monthly cross-sectional regression slope estimates as in Fama and MacBeth (1973). The *t*-statistics are on the time-series means of the coefficients. The regressions on TNA, maximum load, and the turnover measures use the residuals from reported returns after adding back expense ratios.

Independent Variables (Coefficients \times 100)	Estimate	<i>t</i> -statistic
Expense ratio (t)	-1.54	(-5.99)
Turnover (t) (Mturn)	-0.95	(-2.36)
ln TNA (t-1)	-0.05	(-0.66)
Maximum Load (t-1)	-0.11	(-3.55)
Buy turnover (t)	-0.43	(-1.16)
Sell turnover (t)	-1.26	(-3.00)

and *sell turnover*, the estimates imply a 21.5 basis point cost for (one-way) buy trades and a 63 basis point cost for sell trades.

TNA is insignificantly related to the cross-section of performance estimates, but maximum load fees are significantly negatively related to performance. The negative slope on load fees contradicts the oft-cited claim by load funds that their managers are more skilled and investment expenses lower than no-load funds. Although the coefficient appears small, it implies that annual abnormal returns are reduced by about 11 basis points for every 100 basis point increase in load fees. For a load fund with the average total sales charges of 7.3 percent, the reduction in annual return is 79 basis points. To test the sensitivity of this result to the poor-performing outliers, I repeat the analysis after removing the funds in the bottom two deciles. The results (not reported) are virtually unchanged. The underperformance of load funds is probably partially explained by higher total transactions costs, since load funds exhibit higher turnover than no-load funds (Carhart (1995a).)

D. Cross-Sectional Variation in Transaction Costs

Thus far, sensitivity to common factors and persistence in expense ratios explain most of the persistence in mutual fund performance. In addition, the cross-section tests indicate that turnover reduces performance for the average

fund. However, since turnover ratios on the worst-performing funds are only slightly higher than on the average fund, transaction costs can only explain the anomalous underperformance of the worst funds if these funds also have higher costs *per* transaction. This section evaluates whether estimates of costs per transaction explain any of the remaining abnormal performance not fully accounted for by the 4-factor model, expense ratios, and turnover.

I find that transaction costs describe most of the unexplained mutual fund performance. From the 4-factor model alphas, expense ratios, and turnover ratios, I assume market efficiency to infer the cost per transaction necessary to zero out the gross 4-factor alpha. The average fund's alpha of -0.15 percent, expense ratio of 1.14 percent, and turnover of 77.4 percent imply a cost of 85 basis points per round-trip transaction. The previously reported cross-section estimate of round-trip transactions cost is 95 basis points, with a standard error of 40 basis points. Thus, for the average fund, the implied transactions cost lies well within 0.25 standard errors of the estimated cost.

In addition to explaining performance on the average fund, transaction costs also explain much of the cross-sectional variation in return on the portfolios sorted on lagged one-year return. I sort the sample into quintiles to create subsamples large enough to yield reliable cross-section estimates. After repeating my calculations and cross-section estimates, I find that the implied transaction costs are very near to their cross-section estimates. Only in one quintile (quintile 2) is the estimated round-trip transactions cost more than two standard errors from implied. Although the quintile sort is coarse, cross-sectional variation in costs per transaction explains the return spread on these portfolios unrelated to the 4-factor model and expense ratios.

However, the estimated round-trip transaction costs in finer sorts of the bottom quintile are not large enough to explain its underperformance. In order to estimate transaction costs on the relatively small samples of the decile or subdecile portfolios, I pool the cross-section and time-series observations in the estimation. The estimated round-trip transaction costs on decile 10 undershoot the implied costs of 354 basis points by more than four standard errors. The implied costs on the three subportfolios of decile 10 suggest that portfolios 10B and 10C drive this unusually high implied transaction cost estimate. To fully explain 4-factor model abnormal performance, portfolio 10B requires a 356 -basis-point round-trip cost, and 10C requires a 582 -basis-point cost. At seven and 252 basis points, however, the cross-section estimates for these portfolios are considerably less than implied. While their pattern suggests that relative transaction costs play an important role in the cross-section of mutual fund performance, the magnitude of the cross-section estimates leaves unexplained much of the abnormal return in the worst-return mutual funds.

For robustness, I employ a second method for inferring cross-sectional variation in transaction costs that exploits the time-series properties of the mutual fund portfolios. Since round-trip transactions costs should decrease in the trading liquidity of the underlying securities, mutual funds holding illiquid securities should be correlated with a factor-mimicking portfolio for trading liquidity. Assuming that the time-series properties of illiquid stocks differ from

liquid stocks, a portfolio long in illiquid stocks and short in liquid stocks should capture these patterns.

The liquidity factor-mimicking portfolio, VLMH, is the spread between returns on low- and high-trading-volume stocks orthogonalized to the 4-factor model.⁶ I find that the VLMH-loading estimates on mutual fund portfolios are strongly related to performance. The best one-year-return portfolios load significantly and negatively on VLMH, indicating relatively more liquid stocks. The worst portfolios load significantly and positively, indicating relatively more illiquid stocks. Since illiquid stocks are more costly to trade, the VLMH loadings suggest that the costs per transaction are higher for the lower-past-return portfolios. Although these results do not measure the incremental cost of trading illiquid stocks, they do suggest that higher transaction costs might explain the strong underperformance of the worst funds.

Overall, my results suggest that short-run mutual fund returns persist strongly, and that most of the persistence is explained by common-factor sensitivities, expenses, and transaction costs. The net gain in returns from buying the decile of past winners and selling the decile of losers is 8 percent per year. I explain 4.6 percent with size, book-to-market and one-year momentum in stock returns; 0.7 percent with expense ratios; and 1 percent with transaction costs. However, of the 5.4 percent spread between deciles 1 and 9, the 4-factor model explains 4.4 percent and expense ratios and transaction costs explain 0.9 percent, leaving only 0.1 percent annual spread unexplained. Underperforming by twice its expense ratio and estimated transaction costs, the performance on the lowest decile is still anomalous after these explanations. Thus, the cross-section of average mutual fund returns not explained by the variables is almost entirely concentrated in the spread between the bottom two past-returns sorted decile portfolios.

IV. Interpreting the Performance on Past-Winner Mutual Funds

Previous sections demonstrated strong patterns in 4-factor model coefficients on portfolios of mutual funds sorted on one-year return. This finding suggests sorting funds on one-year return groups with similar time-series properties, at least over the period while they are ranked in a particular decile. There are at least two possible explanations for this groupwise commonality. First, the funds in each portfolio might be relatively stable with consistent strategies through time. Second, the funds in each portfolio might be unstable through time, but the funds in a particular decile might hold similar securities while they are in that portfolio. The implications of these two explanations differ drastically, since the former suggests that managers follow consistent strategies that determine their expected returns, whereas the latter is consistent with managers choosing securities randomly but holding them for one to two years.

⁶ Details on the construction of the VLMH portfolio, and the specific estimates discussed in this section, are available from the author upon request.

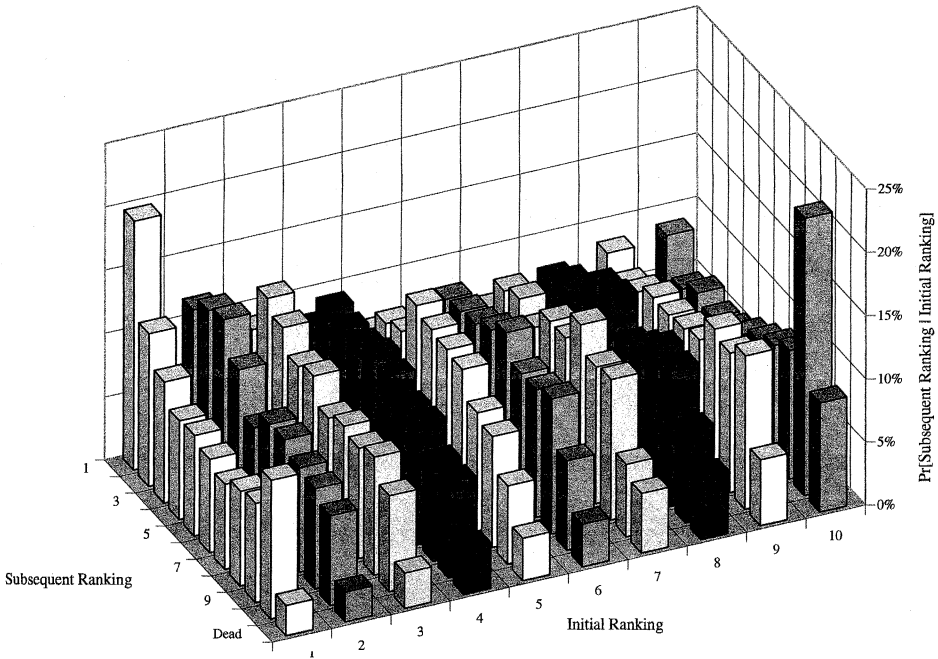


Figure 1. Contingency table of initial and subsequent one-year performance rankings. In each calendar year from 1962 to 1992, funds are ranked into decile portfolios based on one-year gross return. These initial decile rankings are paired with the fund's subsequent one-year gross return ranking. Funds that do not survive the complete subsequent year are placed in a separate category for dead funds. The bars in cell (j, i) represent the conditional probability of achieving a subsequent ranking of decile j (or dying) given an initial ranking of decile i . I estimate gross returns by adding back expense ratios to reported returns.

A. Consistency in Ranking

I test the consistency in fund ranking by constructing a contingency table of initial and subsequent one-year mutual fund rankings. I use simple returns gross of expense ratios to remove the predictable expense element in reported returns. The contingency table is displayed in Figure 1. The bars for initial rank i and subsequent rank j represent $\Pr(\text{rank } j \text{ next year} \mid \text{rank } i \text{ last year})$.

From the figure, it is apparent that winners are somewhat more likely to remain winners, and losers are more likely to either remain losers or perish. However, the funds in the top decile differ substantially each year, with more than 80 percent annual turnover in composition. In addition, last year's winners frequently become next year's losers and vice versa, which is consistent with gambling behavior by mutual funds. Further, the probability of disappearing from the database decreases monotonically in the previous-year's return. Thus, while the ranks of a few of the top and many of the bottom funds persist, the year-to-year rankings on most funds appear largely random.

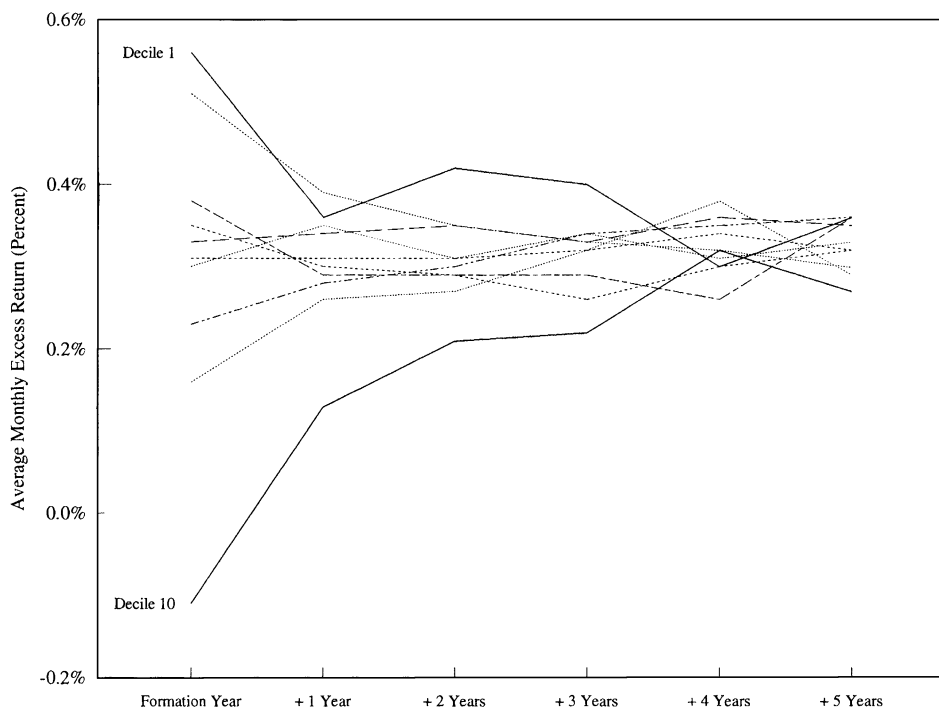


Figure 2. Post-formation returns on portfolios of mutual funds sorted on lagged one-year return. In each calendar year from 1962 to 1987, funds are ranked into equal-weight decile portfolios based on one-year return. The lines in the graph represent the excess returns on the decile portfolios in the year subsequent to initial ranking (the “formation” year) and in each of the next five years after formation. Funds with the highest one-year return comprise decile 1 and funds with the lowest comprise decile 10. The portfolios are equally weighted each month, so the weights are readjusted whenever a fund disappears from the sample.

B. Returns on the Portfolios of Mutual Funds after Ranking

The large number of top-decile funds that revert to lower ranks suggests that the relatively high returns on the funds in this portfolio are short-lived. Figure 2 presents the average returns of the funds in each decile portfolio in each of the five years after their original formation. From the figure, it is clear that one-year performance persistence is mostly eliminated after one year. Except for the persistent underperformance by the worst funds, mean returns and abnormal performance across deciles do not differ statistically significantly after one year.

Furthermore, the returns on the top and bottom decile funds are not nearly so strongly related to the one-year momentum effect in stock returns outside of the ranking and formation years. In the year before ranking, funds that will comprise decile 1 have a PR1YR loading of 0.18, and funds that will comprise decile 10 have a PR1YR loading indistinguishable from zero. In the year after portfolio formation, decile 1 funds have a PR1YR loading of only 0.14, and decile 10 funds have a PR1YR loading of 0.04. These coefficients contrast

sharply with the top- and bottom-decile PR1YR loadings in the portfolio formation year of 0.29 and -0.09 . (See Table III.)

C. Portfolios Sorted on PR1YR Loadings

The results from the previous two sections suggest that most top-ranked mutual funds do not maintain their high relative returns. However, funds that follow a momentum strategy in stocks might consistently earn above-average returns, even if their 4-factor model performance is not abnormal. To test whether momentum managers earn consistently higher returns, I sort mutual funds into portfolios on their 4-factor model PR1YR loadings and find that one-year momentum funds do not earn substantially higher returns than contrarian funds.⁷ Relative to the 4-factor model, in fact, one-year momentum funds underperform one-year contrarian funds. Momentum funds also have high turnover and expense ratios, suggesting that most of the gains from following the one-year momentum strategy are consumed by higher expenses and transaction costs. This result contrasts with Wermers (1996), who finds that momentum funds outperform on a gross performance basis.

My results suggest that Jegadeesh and Titman's (1993) spread in mean return among last-year's winning and losing stocks is not an investable strategy at the individual security level. My results also suggest that there is a simple explanation for the strong pattern in PR1YR loadings on portfolios sorted on lagged one-year returns: These mutual funds don't follow the momentum strategy, but are funds that accidentally end up holding last year's winners. Since the returns on these stocks are above average in the ensuing year, if these funds simply hold their winning stocks, they will enjoy higher one-year expected returns and incur no additional transaction costs for this portfolio. With so many mutual funds, it does not seem unlikely that some funds will be holding many of last year's winning stocks simply by chance.

D. Evidence that PR1YR and VLMH Loadings Capture Momentum and Trading Volume

Daniel and Titman (1997) suggest that firms' actual size and book-to-market equity contain more explanatory power for mean returns than do time-series estimates of factor loadings. Their results suggest that generalizations about the securities held or strategies followed by mutual funds based on time-series factor loadings might be misleading. Fama and French (1993) find that SMB and HML loadings are related to the average market capitalization and book-to-market equity on their test portfolios. Thus, I examine the information content of PR1YR and VLMH loadings by comparing the factor loadings with direct measures of momentum and trading liquidity. If the factor loadings capture the liquidity and momentum of these quantitatively-managed portfolios, they should strongly correlate with direct measures of liquidity and momentum.

⁷ This material is available from the author upon request.

To test this hypothesis, I construct two sets of 25 value-weighted stock portfolios by sorting all NYSE, AMEX, and Nasdaq stocks first on size, and then on one-year momentum or trading volume. The patterns in VLMH loadings on the size-trading volume portfolios support my previous generalizations about the relative liquidity of stocks held by mutual funds.⁸ Within each size quintile, the VLMH loadings decrease in the dollar volume of trading. Since VLMH is long in low-trading-volume stocks and short in high-volume stocks, I expect this inverse relation between trading volume and VLMH loading. Further, VLMH is constructed orthogonally to the size factor, so the VLMH loading does not reveal the magnitude of trading volume, only the magnitude of trading relative to firm size. After subtracting the average trading volume for each size quintile, the correlation between trading volume and VLMH coefficients is 0.74.

I also find that PR1YR loadings are informative on the momentum of stocks in each portfolio. On the size-momentum portfolios, the PR1YR loadings are monotonic in momentum within every size quintile, and the overall correlation between momentum and factor loadings is 0.95. Thus, covariance with the PR1YR factor appears to be a relatively good indication of the momentum of the underlying stocks in a portfolio.

V. Longer-Term Persistence in Mutual Fund Portfolios

A. Two- to Five-Year-Return Sorted Portfolios

Contrary to the suggestions of Hendricks, Patel, and Zeckhauser (1993), mutual fund manager stock-picking skill is not required to explain the one-year persistence in mutual fund returns. However, if manager skill exists, a one-year return is probably a noisy measure. To reduce the noise in past-performance rankings, I form portfolios of mutual funds on lagged two- to five-year returns. I then repeat my earlier analyses to examine how much cross-sectional variation in mean return can be explained by the 4-factor model, expense ratios, and transaction costs. Figure 3 summarizes these and the results from the one-year past-return sorted portfolios.

Using longer intervals of past returns does not reveal more information about expected future mutual fund return or 4-factor performance. While the 4-factor model explains more than half the spread in return on the one-year-return portfolios, it explains a smaller fraction of return spread in the two- to four-year portfolios, and none of the spread in the five-year portfolios. It turns out that the 4-factor model explains less of the return spread because of a less-pronounced pattern in PR1YR loadings and a more pronounced pattern in HML loadings. Past-winner mutual funds load negatively on HML and positively on PR1YR, while past-loser mutual funds do not load significantly on either factor. Expense ratios explain a similar return spread across sorting intervals, approximately 1 percent per year. Estimates of total transaction costs from turnover and cross-section estimates of costs per transaction ex-

⁸ This material is available from the author upon request.

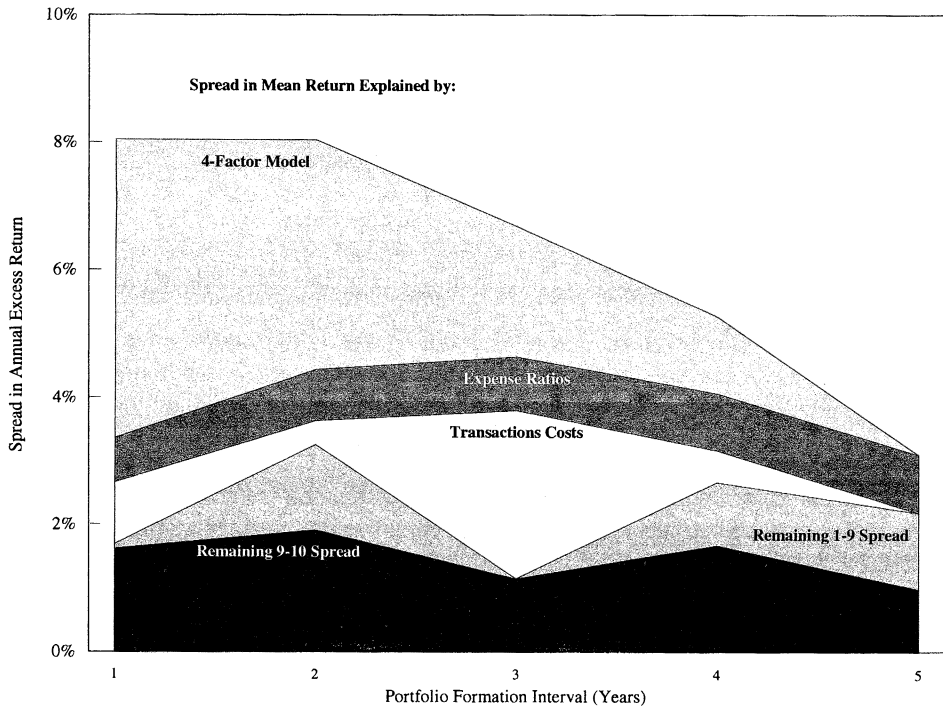


Figure 3. Summary of explanations for persistence in mutual fund performance. On January 1 of each year, funds are ranked into equal-weight decile portfolios based on returns over the prior one-, two-, three-, four-, and five-year periods. Funds with the highest return comprise decile 1, and funds with the lowest comprise decile 10. The height of the graph represents the annual spread in mean return between deciles 1 and 10 for the portfolios formed on one- to five-year returns. The top shaded region represents the spread in annual return that is explained by the 4-factor model, where the 4-factor model captures common variation in return associated with size, book-to-market equity, and one-year return momentum. The second region from the top represents the difference in the average expense ratios of deciles 1 and 10. The third region from the top represents the difference in estimates of total transaction costs for deciles 1 and 10. Total transaction costs are modified turnover times the cross-section estimates of roundtrip transaction costs. Of the remaining spread in annual return after the 4-factor model, expense ratios and transaction costs explanations, the fourth region from the top represents the portion of the unexplained spread attributable to the difference between returns on deciles 1 and 9. The bottom region represents the unexplained spread attributable to the difference between returns on deciles 9 and 10.

plain between zero and 2.6 percent of the spread in annual return. Of the spread in annual return remaining after the 4-factor model, expense ratios, and transaction costs, approximately two-thirds is attributable to the spread between the ninth and tenth decile portfolios. This amounts to approximately 1.5 percent.⁹

⁹ The samples are not held constant across sorting intervals. The sample of one-year past-return portfolios averages 411 mutual funds per year, whereas the sample of five-year portfolios averages only 306 funds per year.

These results differ somewhat from Grinblatt and Titman (1992), who study persistence in five-year mutual fund returns and find slightly stronger evidence of persistence with a similar methodology. However, Grinblatt and Titman condition on five-year subsequent survival, and their sample period includes the very high attrition period of 1975 to 1978 (see Carhart (1995b)). Further, Grinblatt and Titman's (1989) P-8 benchmark does not capture the one-year momentum effect in stock returns. They construct the P-8 model to explain variation in return associated with firm size, dividend yield, three-year past returns, interest-rate sensitivity, co-skewness, and beta. As evidence that the omission of a momentum factor is significant, the intercept from the regression of $PR1YR$ on the P-8 benchmark over Grinblatt and Titman's sample period yields a statistically significant intercept of 0.46 percent per month, with an r -squared of only 0.6. Finally, Grinblatt and Titman do not attempt to account for differences in performance attributable to expenses or transaction costs.

B. Three-Year, 4-Factor, Alpha-Sorted Portfolios

Since I evaluate performance relative to the 4-factor model, sorting mutual funds on alphas from the same model should measure stock-picking talent more accurately. However, using the same asset pricing model to sort *and* estimate performance will also pick up the model bias that appears between ranking and formation periods. For example, if the factor-mimicking portfolios impose risk premia that are too high or too low, funds with consistent 4-factor model loadings will show persistent 4-factor model performance. A similar problem exists if there is an omitted factor in the model. Because of the joint-hypothesis problem, I cannot directly test model bias. Therefore, I interpret the results from these tests with caution.

Table VI reports statistics on decile portfolios formed on lagged three-year alpha estimates from the 4-factor model. Sorting on 4-factor alphas does not achieve as large a spread in mean return as one-year past return (0.43 percent per month versus 0.67 percent), but it does identify funds with larger positive and negative abnormal performance relative to the 4-factor model. The spread in 4-factor alphas is 0.45 percent, substantially greater than the 0.28 percent for portfolios sorted on one-year simple return in Table III.

The Spearman test for rank independence (not reported) fails to reject with a p -value of 7.2 percent, but the top and bottom past-performance deciles are clearly separated from the average-performing midranked funds. The top decile achieves a positive 4-factor model alpha that is eight basis points per month and almost two standard errors above the second-ranked portfolio. Likewise, the bottom past-performance decile underperforms the ninth by 24 basis points per month, a difference of more than three standard errors. Patterns in 4-factor model loadings are not as pronounced, with both the top and bottom past-performance decile funds concentrating in small, growth, and momentum stocks. As in the one-year past-return sorted portfolios, the CAPM beta estimates of the alpha-sorted portfolios are very similar to one another

Table VI

Portfolios of Mutual Funds Formed on 3-Year Past 4-Factor Model Alphas

Mutual funds are sorted on January 1 each year from 1966 to 1993 into equal-weight decile portfolios based on their 4-factor model alphas estimated over the prior 3 years. I require a minimum of 30 return observations for this estimate. Funds with the highest alpha estimates comprise decile 1 and funds with the lowest comprise decile 10. The 4-factor model consists of the RMRF, SMB, HML, and PRIYR factor-mimicking portfolios. RMRF, SMB, and HML are Fama and French's (1993) market proxy and factor-mimicking portfolios for size and book-to-market equity. PRIYR is a factor-mimicking portfolio for one-year return momentum. Expense ratio and turnover are time-series averages of annual cross-sectional averages of the funds in each portfolio. Expense ratio is management, administrative, and 12b-1 expenses divided by average total net assets (TNA). Turnover represents reported turnover plus 0.5 times the percentage change in portfolio TNA adjusted for investment returns and mergers. Roundtrip transactions costs are estimated monthly, using cross-sectional regressions on turnover across all funds in each of 5 quintile sorts on lagged 4-factor alpha. The dependent variable in these regressions is the monthly residual from the 4-factor model, where the factor loadings are estimated on the prior 3 years of gross monthly returns after adding back expense ratios. Alpha is the 4-factor model intercept estimate, and alpha- t is the t -statistic on this estimate. Adjusted alpha is the 4-factor alpha plus 1/12 of expense ratio and 1/12 of turnover times roundtrip transaction costs.

Portfolio	Excess Return	Standard Deviation	4-Factor Model Ordinary Least Squares (OLS) Estimates						Exp Ratio	Turn (Mturn)	Roundtrip Transaction Costs	Adjusted Alpha
			Alpha	Alpha- t	RMRF	SMB	HML	PRIYR				
1 (high)	0.62%	5.07%	0.02%	(0.41)	0.93	0.48	-0.14	0.14	1.13	91.1	-0.17%	0.10%
2	0.47%	4.60%	-0.06%	(-1.37)	0.90	0.32	-0.10	0.10	1.00	69.5	-0.17%	0.01%
3	0.49%	4.49%	-0.03%	(-0.81)	0.90	0.25	-0.06	0.09	0.94	63.7	0.95%	0.10%
4	0.43%	4.43%	-0.05%	(-1.46)	0.91	0.20	-0.04	0.06	0.95	58.6	0.95%	0.08%
5	0.39%	4.45%	-0.13%	(-3.34)	0.90	0.25	-0.03	0.09	0.97	57.4	0.84%	-0.01%
6	0.40%	4.40%	-0.11%	(-2.76)	0.90	0.20	-0.03	0.08	0.98	61.1	0.84%	0.01%
7	0.38%	4.46%	-0.17%	(-3.97)	0.90	0.26	-0.01	0.10	1.10	68.7	1.03%	-0.02%
8	0.40%	4.54%	-0.16%	(-3.29)	0.90	0.29	-0.03	0.11	1.11	64.4	1.03%	-0.01%
9	0.37%	4.68%	-0.19%	(-3.19)	0.89	0.38	-0.06	0.10	1.26	72.4	1.24%	-0.01%
10 (low)	0.19%	5.10%	-0.43%	(-5.89)	0.93	0.49	-0.06	0.11	1.76	96.1	1.24%	-0.18%
1-10 spread	0.43%	1.33%	0.45%	(5.95)	0.00	-0.01	-0.08	0.03	-0.63	-5.0	-1.41%	0.29%
9-10 spread	0.18%	1.07%	0.24%	(3.12)	-0.04	-0.11	0.00	-0.01	-0.50	-23.7	NA	0.17%

(not reported), so the CAPM does not explain the cross-sectional variation in return either. Using longer-term estimates or appraisal ratios (α_i/σ_{ie}), as suggested by Brown et al. (1992), does not substantially affect the results.

While the 4-factor model explains none of the spread in return on past alpha-sorted portfolios, expenses and transaction costs explain about 2 percent of the spread. The expense ratio on the lowest-ranked portfolio exceeds the expense ratio on the highest-ranked fund by 0.63 percent per year. Further, estimates of round-trip transaction costs of the two extreme deciles differ by 1.41 percent. Since the lowest-ranked portfolio trades slightly more frequently, the net difference in total transaction cost estimates is 1.35 percent per year. Thus, of the 5 percent annual spread in mean return between the highest and lowest past alpha-ranked portfolios, the 4-factor model explains nothing, and expenses and transaction costs explain slightly less than one-half.

Underperformance by decile 10 funds relative to decile 9 is still quite pronounced and statistically significant in these portfolios. Decile 10 underperforms decile 9 by 18 basis points per month in mean return, and by 24 basis points per month in 4-factor performance. Differences in expense ratios of 0.5 percent account for only four basis points of the nine-ten spread. Differences in turnover of 24 percent and estimated transaction costs of 1.24 percent explain only another 2.5 basis points of the spread. Even after considering the higher expense ratios and turnover for decile 10, the spread in 4-factor alphas between deciles 9 and 10 is a statistically significant 18 basis points.

Unlike the highest one-year past-return mutual funds, the returns on high past-alpha mutual funds remain above average long after fund ranking. Figure 4 displays the mean monthly excess returns on the funds in each decile portfolio in the first five years after funds are ranked in past-alpha deciles. Although the mean returns on the lowest nine past-performance deciles converge after two years, the highest decile maintains a persistently high mean return a full five years after the portfolio is initially formed. Apparently, a relatively high 4-factor model alpha is a reasonably good indicator of the relative long-term expected return on a mutual fund. However, the 4-factor model alpha on this portfolio over the five-year post-ranking period (not reported) averages only three basis points per month and is not reliably different from zero. This suggests that these funds aren't providing returns substantially beyond those predicted by the 4-factor model. Thus, high-alpha funds also have high sensitivities to the factors in the 4-factor model.

If alpha measures portfolio manager skill, mutual funds should maintain their 4-factor alpha ranking in subsequent, nonoverlapping periods. A contingency table of fund ranks (not reported) finds that relatively few funds stay in their initial decile ranking. Only funds in the top and bottom deciles maintain their rankings more frequently than expected. Funds initially in decile 1 have a 17 percent probability of remaining in that decile, and funds in decile 10 have a 46 percent chance of remaining in decile 10 or disappearing from the sample altogether. Given the high five-year expected return on the highest decile funds versus the second-highest decile, it is surprising that so few funds are able to maintain their top ranking.

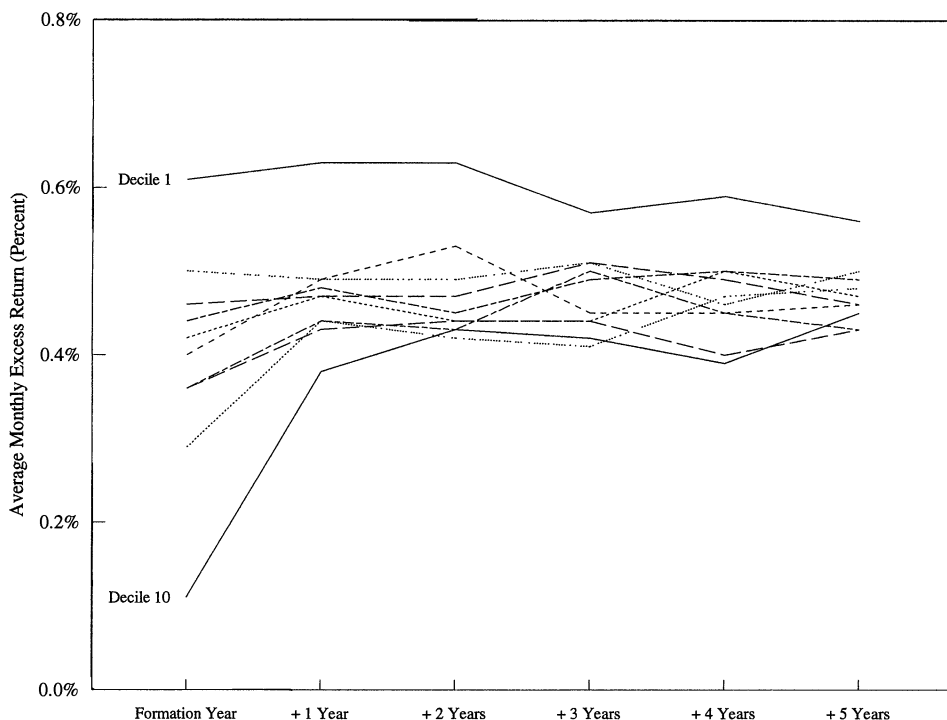


Figure 4. Post-formation returns on portfolios of mutual funds sorted on lagged three-year estimates of 4-factor alpha. In each calendar year from 1962 to 1987, funds are ranked into equal-weight decile portfolios based on three-year estimates of 4-factor alpha. The lines in the graph represent the excess returns on the decile portfolios in the year subsequent to initial ranking (the "formation" year) and each of the next five years after formation. Funds with the highest 4-factor alpha comprise decile 1, and funds with the lowest comprise decile 10. The portfolios are equally weighted each month, so the weights are readjusted whenever a fund disappears from the sample.

Apparently, neither expense ratios nor turnover completely explain the persistent spread and pattern in 4-factor abnormal returns on mutual funds. About 0.6 percent of the 5 percent annual spread in net alphas can be explained by expense ratios; variation in transaction costs accounts for another 1.4 percent. The most striking result is the size of the spread captured by the strong underperformance in the lowest-ranked funds, even after adjustments for expenses and transaction costs.

VI. Conclusion

This article does much to explain short-term persistence in equity mutual fund returns with common factors in stock returns and investment costs. Buying last year's top-decile mutual funds and selling last year's bottom-decile funds yields a return of 8 percent per year. Of this spread, differences in the market value and momentum of stocks held explain 4.6 percent, differences in

expense ratios explain 0.7 percent, and differences in transaction costs explain 1 percent. Sorting mutual funds on longer horizons of past returns yields smaller spreads in mean returns, all but about 1 percent of which are attributable to common factors, expense ratios, and transaction costs. Further, the spread in mean return unexplained by common factors and investment costs is concentrated in strong underperformance by the bottom decile relative to the remaining sample. Of the spread in annual return remaining after the 4-factor model, expense ratios, and transaction costs, approximately two-thirds is attributable to the spread between the ninth- and tenth-decile portfolios.

I also find that expense ratios, portfolio turnover, and load fees are significantly and negatively related to performance. Expense ratios appear to reduce performance a little more than one-for-one. Turnover reduces performance about 95 basis points for every buy-and-sell transaction. Differences in costs per transaction account for some of the spread in the best- and worst-performing mutual funds. Surprisingly, load funds substantially underperform no-load funds. After controlling for the correlation between expenses and loads, and removing the worst-performing quintile of funds, the average load fund underperforms the average no-load fund by approximately 80 basis points per year.

This article offers only very slight evidence consistent with skilled or informed mutual fund managers. Mutual funds with high 4-factor alphas demonstrate above-average alphas and expected returns in subsequent periods. However, these results are not robust to model misspecification, since the same model is used to estimate performance in both periods. In addition, the higher expected performance for high-alpha funds is only relative, since these funds do not earn significantly positive expected future alphas. The evidence is consistent with the top mutual funds earning back their investment expenses with higher gross returns.

Overall, the evidence is consistent with market efficiency, interpretations of the size, book-to-market, and momentum factors notwithstanding. Although the top-decile mutual funds earn back their investment costs, most funds underperform by about the magnitude of their investment expenses. The bottom-decile funds, however, underperform by about twice their reported investment costs. Apparently, these results are not confined to mutual funds: Christopherson, Ferson, and Glassman (1995) reach qualitatively similar conclusions about pension fund performance. However, the severe underperformance by the bottom-decile mutual funds may not have practical significance, since they are always the smallest of the funds, averaging only \$50 to \$80 million in assets, and because the availability of these funds for short positions is doubtful.¹⁰

Buying last year's winners is an implementable strategy for capturing Jegadeesh and Titman's (1993) one-year momentum effect in stock returns virtually without transaction costs, since the actual trading costs are shifted to

¹⁰ Jack White & Co. permits short selling on about 100 funds of the 4,000 no-load funds in their network.

the long-term holders of mutual funds. However, the current mutual fund practice of selling shares at NAV cannot be a long-run equilibrium after this strategy is widely followed: Equilibrium requires mutual funds to charge transaction fees to incoming and outgoing investors to compensate for their perturbing effects on performance. This practice is already becoming common among many funds that hold illiquid stocks such as the Vanguard Small Capitalization Index Fund and Dimensional Fund Advisors Emerging Markets Index Fund.

The evidence of this article suggests three important rules-of-thumb for wealth-maximizing mutual fund investors: (1) Avoid funds with persistently poor performance; (2) funds with high returns last year have higher-than-average expected returns next year, but not in years thereafter; and (3), the investment costs of expense ratios, transaction costs, and load fees all have a direct, negative impact on performance. While the popular press will no doubt continue to glamorize the best-performing mutual fund managers, the mundane explanations of strategy and investment costs account for almost all of the important predictability in mutual fund returns.

REFERENCES

- Asness, Clifford S., 1994, Variables that explain stock returns, Unpublished Ph.D. dissertation, Graduate School of Business, University of Chicago, Chicago, Ill.
- Asness, Clifford S., John M. Liew, and Ross L. Stevens, 1996, Parallels between the cross-sectional predictability of stock and country returns, Working paper, Goldman Sachs.
- Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679–698.
- Brown, Stephen J., William Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.
- Carhart, Mark M., 1992, Persistence in mutual fund performance re-examined, Working paper, Graduate School of Business, University of Chicago, Chicago, Ill.
- Carhart, Mark M., 1995a, Survivor bias and persistence in mutual fund performance, Unpublished Ph.D. dissertation, Graduate School of Business, University of Chicago, Chicago, Ill.
- Carhart, Mark M., 1995b, Survivor bias and mutual fund performance, Working paper, School of Business Administration, University of Southern California, Los Angeles, Cal.
- Carhart, Mark M., Robert J. Krail, Ross L. Stevens, and Kelly D. Welch, 1996, Testing the conditional CAPM, Working paper, Graduate School of Business, University of Chicago, Chicago, Ill.
- Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, Forthcoming, *Journal of Finance*.
- Christopherson, Jon A., Wayne E. Ferson, and Debra A. Glassman, 1995, Conditioning manager alphas on economic information: Another look at the persistence of performance, Working paper, University of Washington School of Business Administration, Seattle, Wash.
- Chen, Zhiwu, and Peter J. Knez, 1995, Portfolio performance measurement: Theory and applications, *Review of Financial Studies* 9, 511–555.
- Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross-sectional variation in stock returns, *Journal of Finance* 52, 1–33.
- Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Christopher R. Blake, 1996, The persistence of risk-adjusted mutual fund performance, *Journal of Business* 69, 133–157.
- Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Matt Hlavka, 1993, Efficiency with costly information: A re-interpretation of evidence from managed portfolios, *Review of Financial Studies* 6, 1–21.

- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–53.
- Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.
- Fama, Eugene F., and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.
- Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–462.
- Goetzmann, William N., and Roger G. Ibbotson, 1994, Do winners repeat? Patterns in mutual fund performance, *Journal of Portfolio Management* 20, 9–18.
- Granger, C. and P. Newbold, 1974, Spurious regressions in econometrics, *Journal of Econometrics* 2, 111–120.
- Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62, 393–416.
- Grinblatt, Mark, and Sheridan Titman, 1992, The persistence of mutual fund performance, *Journal of Finance* 42, 1977–1984.
- Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088–1105.
- Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1993, Hot hands in mutual funds: Short-run persistence of performance, 1974–88, *Journal of Finance* 48, 93–130.
- Jegadeesh, Narasimham, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.
- Jensen, Michael C., 1969, Risk, the pricing of capital assets, and evaluation of investment portfolios, *Journal of Business* 42, 167–247.
- Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.
- Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549–572.
- Sharpe, William F., 1964, Capital Asset Prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–42.
- Wermers, Russ, 1996, Momentum investment strategies of mutual funds, performance persistence, and survivorship bias, Working paper, Graduate School of Business and Administration, University of Colorado at Boulder, Boulder, Col.