THE JOURNAL OF FINANCE • VOL. L, NO. 2 • JUNE 1995

Performance Persistence

STEPHEN J. BROWN and WILLIAM N. GOETZMANN*

ABSTRACT

We explore performance persistence in mutual funds using absolute and relative benchmarks. Our sample, largely free of survivorship bias, indicates that relative risk-adjusted performance of mutual funds persists; however, persistence is mostly due to funds that lag the S&P 500. A probit analysis indicates that poor performance increases the probability of disappearance. A year-by-year decomposition of the persistence effect demonstrates that the relative performance pattern depends upon the time period observed, and it is correlated across managers. Consequently, it is due to a common strategy that is not captured by standard stylistic categories or risk adjustment procedures.

A NUMBER OF EMPIRICAL studies demonstrate that the relative performance of equity mutual funds persists from period to period. Carlson (1970) finds evidence that funds with above-median returns over the preceding year typically repeat their superior performance. Elton and Gruber (1989) cite a 1971 Securities and Exchange Commission (SEC) study that indicates similar persistence in risk-adjusted mutual fund rankings. Lehmann and Modest (1987) report some evidence of persistent mutual fund alphas, and Grinblatt and Titman (1988, 1992) show that the effect is statistically significant. Goetzmann and Ibbotson (1994) conclude that the performance persistence phenomenon is present in raw and risk-adjusted returns to equity funds at observation intervals from one month to three years. In an in-depth study focused on growth funds, Hendricks, Patel, and Zeckhauser (1993) show that the performance persistence phenomenon appears robust to a variety of risk-adjustment measures. All of these studies lend strong support to the conventional wisdom that the track record of a fund manager contains information about future performance.

* Brown is from the Leonard Stern School of Business, New York University. Goetzmann is from the School of Management, Yale University. We thank Bent Christiansen, Edwin Elton, Mark Grinblatt, Martin Gruber, Gur Huberman, Roger Ibbotson, Philippe Jorion, Robert Korajczyk, Josef Lakonishok, Jayendu Patel, Nadav Peles, Stephen Ross, William Schwert, Erik Sirri, an anonymous referee, and editor René Stulz for helpful comments. We thank participants in the following conferences and workshops for their input: the 1993 Western Finance Association Meetings, the October 1993 NBER program on asset pricing, the 1994 AFA session on the behavior of institutional investors, and finance workshops at Columbia Business School, The University of Texas, the University of Illinois at Champagne, and the Yale School of Management. We also thank Nadav Peles for programming assistance and Stephen Livingston for data collection. We thank Ibbotson Associates and Morningstar, Inc., for providing data used in the analysis. All errors are the sole responsibility of the authors. Roger Ibbotson contributed substantially to the preliminary version of this article. This article was formerly circulated as "Attrition and Mutual Fund Performance."

In this article, we explore the phenomenon of performance persistence in equity mutual funds using a sample that contains defunct as well as surviving funds. The sample shows that poorly performing funds disappear more frequently from the mutual fund universe, suggesting that selection bias concerns can be relevant to mutual fund performance studies. We document performance persistence on this broad sample and show that it is robust to adjustments for risk. We find that much of the persistence is due to funds that repeatedly lag passive benchmarks. Though most previous studies have aggregated the results from different time periods in order to increase the power of tests designed to identify performance persistence, we find it instructive to break the analysis down on a year-by-year basis. This temporal disaggregation provides some clues regarding the source of the persistence phenomenon. While most years winners and losers repeat, occasionally the effect is dramatically reversed. These reversals suggest there are two possible reasons for persistence. First, persistence is correlated across managers. Consequently, it is likely due to a common strategy that is not captured by standard stylistic categories or risk adjustment procedures. Second, while losing funds have an increased probability of disappearance or merger, not all of them are eliminated. The market fails to fully discipline underperformers, and their presence in sample contributes to the pattern of relative persistence. The implications of our results for investors are that the persistence phenomenon is a useful indicator of which funds to avoid. However, evidence that the pattern can be used to beat absolute, risk-adjusted benchmarks remains weak. Future research should address the issues of cross-fund correlation and the persistence of poor performers. This article is organized as follows. Section I describes the database and the method of data collection. Section II considers the determinants of fund disappearance. Section III reports the results of the performance persistence tests. Section IV concludes.

I. Data

The Weisenberger Investment Companies Service reports information about virtually all publicly offered open-end mutual funds on an annual basis. Data were collected by hand from their *Mutual Funds Panorama*, a data section in Weisenberger (1977 through 1989), for the years 1976 through 1988, for all firms listed as common stock funds, or those specialty funds that invested in common stock (typically sector funds). For each fund, we record the name as it appeared that year, the year of origin, the fund objective, the net asset value at the end of the year, the net asset value per share at the beginning of the period, the twelve-month percentage change in net asset value per share adjusted for capital gains distributions, the income return, the capital gains distributions, and the expense ratio. We calculate the total return inclusive of capital appreciation, income, and capital gains distributions. In some cases, one or more of these data were not reported, and this prevents total return calculations.[1] We assign a unique number to each fund. Footnote at the end of the *Panorama* section indicate merged funds and name changes of funds. When one fund was merged into another, the acquired fund is deemed to have disappeared, while the acquiring fund is deemed to have continued in operation. Though Weisenberger seeks to collect information on all open-ended mutual funds, they may not have included small funds, or funds for which their data were incomplete. Since reporting to Weisenberger is at least in part discretionary, the data base may not be completely free of survivorship bias. In addition, by using only those funds for which an annual return may be calculated, we omit all funds that existed for less than one year. We therefore exclude from the sample the year that funds do poorly and merge or fail.[2] Despite these limitations, the Weisenberger database that we have assembled allows us to directly examine the process of fund attrition.

Table I reports the number of funds extant each year and breaks them into five categories by objective. The sample ranges from 372 funds in 1976 to 829 funds in 1988, with most of the growth occurring in the period 1983 to 1988. This breakdown by category shows the evolution of funds' styles through time, in a way that a backward-looking sample could not. Notice that in the early years of the sample, the Maximum Capital Gain category represented nearly 30 percent of the extant mutual funds. Presumably, these are funds that invested in high growth potential, or high price/earnings ratio (P/E) stocks. By the end of the sample, the popularity of this investment style had shrunk by half, when measured by number of funds. Table I also indicates the capitalization by category. The Weisenberger equity mutual fund universe grew from \$37.2 billion in 1976 to \$166.5 billion in 1988.

Table II reports the equally-weighted and value-weighted mean each year for the whole sample and for those funds that survived for the entire sample. The returns to the S&P 500 and the Vanguard Index Trust (an S&P Index fund) are reported for purposes of comparison. Note that the average returns for mutual funds differ significantly from year to year from the benchmarks. In 1979, for instance, mutual funds appear to substantially outperform the S&P 500. In most years of the 1980s they substantially underperform. The deviation of the aggregate mutual fund performance index from the S&P 500 may be due to differences in composition. S&P 500 returns lagged small stock returns for the years 1977 through 1983, and they exceeded small stock returns for the period 1984 through 1987. This is similar to the mutual fund

[1] The fund returns are calculated as follows:

$$Return_t = \frac{\Delta NAV_t}{NAV_{t-1}} + \frac{D_t}{NAV_{t-1}}$$

where ΔNAV_t is the change in net asset value per share adjusted to capital gains distributions as reported by Weisenberger, D_t is the dollar-denominated investment income per share at time t, and NAV_{t-1} is the net asset per share in the preceding period.

[2] The Elton *et al.* (1993) study takes great care to avoid this source of survivorship bias by computing returns for a buy-and-hold position in mutual funds from 1964 through 1984, accounting for merger terms of funds that are combined with other funds.

Table I

Number and Capitalization of Equity Mutual Funds by Category

Max Cap refers to managers seeking maximum capital gains. Growth refers to growth managers, Income refers to equity managers seeking income return. G&I refers to managers seeking a combination of the two. Other includes an Income and Growth category, presumably indicative of managers placing secondary emphasis on growth, funds with a major international component, and funds for which the objective is not identified. Quantities are expressed in millions of dollars.

	Number of Funds						Capitalization				
	Max Cap	Growth	Income	G&I	Other	Total	Max Cap	Growth	Income	G&I	Total
1976	106	158	7	90	11	372	4940	15574	50	15669	37218
1977	100	153	11	80	11	355	4532	13588	316	13416	32697
1978	102	148	11	80	9	350	4569	12941	441	12712	31348
1979	97	144	13	76	7	337	5719	14631	520	13644	35205
1980	99	138	15	72	7	331	8245	18797	683	15794	44346
1981	94	146	19	74	7	340	7505	17957	959	14252	41484
1982	93	157	22	78	7	357	10975	22331	1763	16440	52541
1983	89	188	26	81	11	395	16891	32563	3876	21380	76066
1984	97	215	33	94	9	448	17351	33332	4792	23181	80228
1985	93	289	41	91	10	524	23152	47089	9199	29580	110831
1986	97	356	59	107	8	627	28870	53500	14180	45426	144411
1987	104	404	68	131	1	708	30594	60680	14089	51606	156977
1988	109	472	81	165	2	829	32652	64297	14316	55193	166474

Table II

Annual Summary Statistics For Equity Mutual Funds

Extant in 1988 represents the sample of funds that were existing in 1988. EW refers to equally weighted mean, and VW refers to value-weighted mean. The value-weighted mean is calculated by the capitalization of the fund at the beginning of the period. Vanguard represents the Vanguard Market Index Trust. Means are calculated over the period 1977 through 1987.

	Whole Sample		Extant in 1988		Gone By 1988		Benchmarks	
Year	EW Mean	VW Mean	EW Mean	VW Mean	EW Mean	VW Mean	S&P 500	Van-guard
1977	0.015	−0.028	0.029	−0.025	−0.010	−0.040	−0.072	−0.078
1978	0.108	0.102	0.125	0.102	0.075	0.103	0.066	0.058
1979	0.292	0.261	0.306	0.260	0.261	0.264	0.184	0.178
1980	0.334	0.326	0.351	0.321	0.287	0.356	0.324	0.312
1981	−0.017	−0.040	−0.017	−0.036	−0.015	−0.066	−0.049	−0.052
1982	0.233	0.220	0.248	0.223	0.175	0.195	0.214	0.200
1983	0.214	0.212	0.224	0.218	0.159	0.141	0.225	0.212
1984	−0.024	−0.019	−0.022	−0.016	−0.039	−0.061	0.063	0.060
1985	0.267	0.270	0.273	0.272	0.215	0.252	0.322	0.308
1986	0.158	0.164	0.162	0.165	0.113	0.148	0.185	0.179
1987	0.022	0.028	0.024	0.030	−0.021	−0.078	0.052	0.051
1988	0.133	0.139	0.133	0.139	NA	NA	0.168	0.161
Mean	0.145	0.136	0.153	0.138	0.109	0.110	0.140	0.132

pattern. Mutual funds exceeded the S&P 500 for the period 1977 through 1982, and then lagged over the period 1983 through 1988.[3]

The effect of survivorship upon the estimated annual return to investment in mutual funds is not trivial. Table II divides the whole sample into two categories, those funds that had disappeared by 1988, and those funds that were extant in 1988. The equally-weighted average of defunct funds is below the equally-weighted average of the entire sample for every year since 1981. This is not true for the value-weighted indices, which track more closely. The implication is that most of the difference is due to the attrition of small funds that performed poorly and were shut down or merged into other funds. We find the difference between returns composed of the entire sample and returns composed of funds extant in 1988 to be 0.8 percent per year. When returns are scaled by capitalization, however, the margin is much lower: 0.2 percent per year.[4] This is not surprising, since we expect larger funds to have a higher probability of survival, and thus weigh more heavily in the mean calculation.

[3] Lakonishok, Shleifer, and Vishny (1992) find a similar result for pension funds and also attribute it to a small firm effect.

[4] These results are consistent with those reported in Grinblatt and Titman (1988) and Malkiel (1995).

II. Fund Disappearance

The Weisenberger database allows us to examine the factors contributing to fund disappearance. Mutual funds typically disappear as a result of being terminated or merged into other funds. As a result, fund disappearance is a management decision, which is presumably based upon fund profitability, and ultimately upon consumer demand. Studies of consumer behavior and the dollar flows into mutual funds indicate that investors select funds on the basis of past performance.[5] We also expect the age of the fund to influence consumer response to returns, since older funds have a longer track record from which to infer differential performance. Age and fund size may also interact with track record as a determinant of survival. A thorough description of the process that governs fund survival would require publicly unavailable revenue and expense data. Thus, we estimate a reduced form model that captures the major factors contributing to the decision to close or merge a fund.

The first specification models disappearance as a function of relative return (i.e., fund return in the year less average fund return that year), relative size (fund size less average fund size that year), expense ratio, and age (expressed in years since fund inception). The results are reported in Table III. All variables are significant at the 95 percent level, and only one, the expense ratio, increases the probability of disappearance. The signs on the relative return coefficient are consistent with previous research into customer response to investment performance. The probit shows that poor performers not only lose customers, but also have a higher probability of disappearing. The negative coefficient on relative size indicates that the bigger the fund, the less likely it is to disappear. Of course, it is difficult to separate this effect from past performance, since a good track record attracts customers. The positive coefficient on expense indicates that funds with higher expense ratios also have a higher probability of disappearance. This is consistent with the existence of some fixed operating costs for funds—if all costs were variable, companies would have little incentive to shut even the small funds down. It is also interesting to note the negative coefficient on age. Younger funds clearly have a higher probability of disappearance.

In the second model we include additional variables that might explain fund disappearance. These are the lagged relative return, relative new money, and relative new money lagged.[6] New money is included because we conjecture that the decision to close the fund depends upon customer response to returns, rather than returns themselves. In addition, we consider other interaction terms among variables. For instance, it seems plausible that large funds might be less susceptible to closure as a result of poor performance or shorter histories. In addition, older funds with poor returns might be less likely to close than younger ones. To account for these effects, we specify interaction terms between relative return and age, relative return and rela-

[5] See Patel, Hendricks, and Zeckhauser (1990), Kane, Santini, and Aber (1991), Ippolito (1992), Lakonishok, Shleifer, and Vishny (1992), and Sirri and Tufano (1992), for examples.

[6] New money is calculated as: $NM_t = [NAV_t - (1 + r_t)NAV_{t-1}]/NAV_{t-1}$.

Table III

Probit Model of Fund Disappearance

The probit models the odds of fund disappearance as a function of the specified variables. A negative coefficient implies that a higher value for the variable decreases the chance of fund disappearance. Coefficients for the probit model are estimated via maximum likelihood. T-statistics are in parentheses. Relative return $(t-1)$ is the $t-1$ period total return less the average across funds that year. Relative new money $(t-1)$ is the percentage increase in fund net asset value due to share purchases, less the average across funds that year. Relative size is the net asset value less average across funds that year. Expense ratio is expressed $\times 100$ and Age is in years since inception of the fund. Interactions are represented by colons.

	Model 1	t-Statistic	Model 2	t-Statistic	Model 3	t-Statistic	Model 4	t-Statistic
Intercept	−1.642	(29.00)	−1.838	(20.05)	−1.782	(16.42)	−1.551	(21.69)
Relative return $(t-1)$	−1.402	(5.60)	−1.100	(2.02)			−1.245	(2.85)
Relatie return $(t-2)$			−0.691	(2.09)			−0.510	(1.91)
Relative return $(t-3)$							−0.594	(2.26)
Relative new money $(t-1)$			−0.003	(1.21)	0.0234	(2.07)		
Relative new money $(t-2)$			−0.004	(0.67)	0.0052	(0.91)		
Relative new money $(t-3)$					−0.0057	(0.74)		
Relative size $(t-1)$	−0.117	(3.47)	−0.070	(1.06)	−0.191	(2.18)	−0.150	(2.48)
Expense ratio $(t-1)$	0.047	(2.47)	0.086	(3.11)	0.102	(3.06)	0.046	(2.04)
Age $(t-1)$	−0.005	(2.08)	−0.003	(0.71)	−0.005	(1.17)	−0.009	(2.97)
Relative Return: Age			0	(0.001)			0.009	(0.33)
Relative return: Relative size			0.414	(0.32)			−0.264	(0.72)
Age: Relative size			−0.001	(0.456)			0.002	(0.84)
Relative NM: Age					−0.001	(1.69)		
Relative NM: Relative size					−0.001	(0.50)		
Age: Relative size					0.001	(0.73)		
Observations	5580		3981		3255		4923	

tive size, and age and relative size. The results of this specification are somewhat surprising. Relative returns and lagged relative returns are both significant predictors, but new money is not. The sign on new money indicates that past positive inflows decrease the probability of closure, although the effect is weak. The third specification shows that eliminating returns and replacing them with lagged values of new money makes the first lag of new money a significant predictor of disappearance, but not earlier lags. This may be due to the fact that fund managers care about total growth, rather than new money, or it may be due to our inability to adequately model the past and future interactions between returns and new money. The coefficients on the interaction terms among the other variables all proved to be insignificantly different from zero.

The fourth model includes longer lags for returns. The rationale for considering longer lags is the possibility that the extended track record contributes to the closure decision. When we include three years of past relative returns in the model, we find all three years to be significant or near-significant predictors of fund disappearance. The coefficients on past relative returns lagged one and two years are of the same sign as the relative returns in the first model, and a bit less that half of the magnitude. In other words, poor performance in earlier years is not as powerful a factor in fund disappearance, but certainly is an important one.

The results of the probit analysis suggest that past performance over several years is a major determinant of fund disappearance. Surprisingly, net fund growth, at least as we have defined it, contributes only marginally to prediction of fund disappearance. Other variables clearly play a role in predicting fund closure. Size and age are negatively related to fund disappearance, and expense ratio is positively related to fund disappearance.

III. Repeat Performers

Following Brown *et al.* (1992) and Goetzmann and Ibbotson (1994), we track the evolution of the mutual fund universe using a nonparametric methodology based upon contingency tables. Table IV reports the frequency counts for each year. The table identifies a fund as a winner in the current year if it is above or equal to the median of all funds with returns reported that year. The same criterion is used to identify it as a winner or loser for the following period. Thus, Winner-Winner (WW) for 1976 is the count of the winners in 1976 that were also winners in 1977. The same principle defines the other categories. Winner-Gone indicates the number of winners that disappeared from the sample in the following period. No Data indicates that Weisenberger listed the fund for the following period, but was unable to collect data necessary to calculate a total return. New Fund indicates the number of new funds that appear in that year. Cross-Product Ratio reports the odds ratio of the number of repeat performers to the number of those that do not repeat; that is, $(WW*LL)/(WL*LW)$. The null hypothesis that performance in the first period is unrelated to performance in the second period corresponds to an odds ratio of one. In large samples with independent observations, the standard error of the natural log of the odds ratio is well approximated.[7]

Brown *et al.* (1992) show that fund attrition and cross-fund dependencies tend to bias the cross-product ratio test toward rejection.[8] The degree of this bias in the cross-product ratio test is dependent both upon the correlation structure of the mutual fund universe and upon the attrition rate. To address this problem, we bootstrap the distribution of the odds ratio conditional upon an actual correlation matrix of mutual fund returns, and upon attrition rates for winners and losers. In effect, we replicate the effect of fund attrition, cross-sectional dependencies, and heteroskedasticity upon the test statistics.[9]

In Table IV we report the test statistic for the odds ratio test, as well as its bootstrap probability value, conditional upon the sample correlation matrix of fund returns and the observed attrition rates. Under both the biased and the corrected hypothesis tests, we find that seven years (eight years for the bias-corrected distributions) of the sample indicate significant positive persistence, and two years indicate significant negative persistence. The bootstrapped probability values generally agree with the theoretical distribution of the test statistics and alternative procedures designed to address survivorship.[10]

[7] This is given as:

$$\sigma_{\log(odds\ ratio)} = \sqrt{\frac{1}{W,W} + \frac{1}{W,L} + \frac{1}{L,W} + \frac{1}{L,L}}$$

See Christensen (1990) p. 40, for instance.

[8] Patel and Zeckhauser (1992) also investigate the distribution of the performance persistence test statistics under the alternative of a performance threshold. They observe that there are some interesting consequences to dividing the samples up into octiles each period, rather than into halves. Their simulations yield an interesting result. Performance may *reverse* for the poorest performing group of managers who survive both periods.

[9] To simulate the distribution of the log odds ratio, we used a de-meaned two-year sample of monthly mutual fund returns (obtained from Morningstar, Inc.) over the period 1987 to 1988, from which we selected a sample without replacement, of size corresponding to the sum of *WW*, *LW*, *WL*, and *LL* for the given year. We simulated return series through randomization over the dimension of time, and eliminated the appropriate number of funds in the appropriate category. For losing funds, we assume that the poorest performers would be eliminated. This represents a conservative approach, since it maximizes any potential bias of the sort reported in Brown *et al.* (1992). We calculate the odds ratio for the 2 × 2 table of winners and losers, and the likelihood ratio statistic for the contingency table including funds that disappeared. This is performed 100 times to generate simulated distributions that correspond to the null hyopthesis of no performance persistence, conditional upon a typical correlation structure, fund variances, and actual attrition rates. The simulated distributions are used for comparison to the actual statistics. Note that this procedure relies upon a variance-covariance matrix that is singular, since there are more securities than time periods. Thus, we are unable to completely address the issue of the sampling error of the statistic.

[10] For comparison to the simulated distributions of cross-product ratios, we applied standard limited dependent variable procedures to the problem of estimating year by year cross-sectional relationships between returns, conditional upon survival. Using the inverse Mills ratio estimated from the probit regression helps explain some differential in annual performance, but does little to affect the general performance persistence results.

Table IV

Frequency of Repeat Performers and Related Categories: Entire Sample

Winner-winner indicates the number of above median funds in the year that were also above median funds in the following year. Loser-Winner, Winner-Loser, and Loser-Loser are defined similarly. Winner-Gone and Loser-Gone indicate the number of funds that were above median and disappeared, and those that were below median and disappeared. No Data indicates a lack of return data for that year. New Fund indicates the number of new funds that appeared in that year. The cross-product ratio, also referred to as the odds-ratio, is calculated as: (Winner-Winner * Loser-Loser)/(Loser-Winner * Winner-Loser). The Z-statistic is the log odds ratio divided by its standard error, and is asymptotically normally distributed, under the assumption of independence of the observations. Bootstrap p-value refers to the probability value taken from the numerical simulation of the odds ratio as described in the text. It explicitly incorporates cross-dependencies in the observations, and conditions upon fund attrition counts each period.

Year	Total	Winner-Winner	Loser-Winner	Winner-Loser	Loser-Loser	Winner-Gone	Loser-Gone	No Data	New Fund	Cross-Product Ratio	Z-Statistic	Bootstrap p-Value
1976	372	106	62	64	104	15	19	2	NA	2.78	4.53	0.00
1977	355	111	56	52	97	11	20	8	16	3.70	5.50	0.00
1978	350	114	48	52	102	6	21	7	23	4.66	6.35	0.00
1979	337	115	41	46	106	5	18	6	14	6.46	7.36	0.00
1980	331	50	95	101	49	9	15	12	12	0.25	−5.53	1.00
1981	340	90	55	61	93	7	10	24	29	2.49	3.85	0.00
1982	357	89	67	74	86	7	16	18	27	1.54	1.92	0.01
1983	395	92	79	81	83	6	16	38	57	1.19	0.81	0.23
1984	448	102	99	87	101	13	2	44	76	1.20	0.88	0.20
1985	524	139	83	90	136	2	12	62	93	2.53	4.78	0.00
1986	627	162	103	104	156	12	19	71	117	2.36	4.80	0.00
1987	708	134	182	181	124	11	19	57	118	0.50	−4.20	1.00
Total	5144	1304	970	993	1237	104	187	349	582	1.67	27.78	

The disaggregation reveals some interesting features of the persistence phenomenon. We find evidence of significant persistence seven or eight out of twelve years. While persistence is more common, it is important to emphasize that reversal also occurs. One of the years that indicated a significant reversal pattern was 1987. Winning funds in 1987 tended to be losing funds in 1988. Malkiel (1995) finds reversals in two of the years following our sample period, which suggests that the probability of reversal is high and confirms that the strongest evidence for repeat performance is over the late 1970s and early 1980s.

The reversals also indicate that persistence is correlated across managers. This is important because it tells us that persistence is probably *not* due to individual managers selecting stocks that are overlooked or ignored by other managers. Whatever the cause of winning, it is evidently a group phenomenon. While this correlation in persistence is consistent with recently identified herding behavior among equity fund managers (see Grinblatt, Titman, and Wermers (1994)), it is also consistent with correlated dynamic portfolio strategies, such as portfolio insurance (see Connor and Korajczyk (1991)).

A. *Risk Adjustment*

One possible explanation for the secular trend in performance persistence is that systematic risk differs across managers. The annual frequency of the Weisenberger data makes it difficult to use traditional risk adjustments. To address this problem, we use the Morningstar monthly database for the period 1976 through 1988 for a subset of the funds in the Weisenberger sample. By merging these two datasets, we obtain fund characteristics as well as monthly return data for a substantial subset of the mutual fund universe. We use this merged database to model fund betas and residual errors as linear functions of other mutual fund characteristics. We specify a traditional single index model, as well as the Elton *et al.* (1993) three-index model, and report estimates for the coefficients and residual errors in Table V.[11] Beta and residual risk measures differ significantly according to fund classification, prior year size and expense ratios, and period since inception of the fund. The R^2 indicates that the model performs well, and rankings of three index beta measures by fund characteristic correspond to those reported by Elton *et al.* (1993) on the basis of annual data. We then use the estimated coefficients from this model to extrapolate beta and residual risk measures on the basis of characteristics of *all* funds in the Weisenberger sample.

[11] In the table, the Growth fund classification represents the base case. The linear model is described in the notes to Table V. It was estimated via OLS and GLS, the latter being used to correct for heteroskedasticity. Results for the two models were nearly indistinguishable, and we report the GLS estimates. The estimation of forecasting of mutual fund betas and factor sensitivities as a function of observable factors is a topic of on-going research. See, for instance, Ferson and Schadt (1995).

Table V

Regression of Beta and Residual Risk on Fund Characteristics

Results in this table are obtained by regressing excess monthly returns realized on 521 mutual funds with at least 12 months of data reported by Morningstar for the period January 1976 through December 1988 and for which Weisenberger reports prior year end net asset value, expense ratios, and fund descriptors. Columns under the Single Index Model and Three-Index Model represent the coefficients β_{ki}, $k = 0, \ldots, 10$, $k = 1, \ldots, I$ estimated from the regression

$$R_{jt} - R_f = \alpha_0 + \sum_k \alpha_k \times f_{kt} + \sum_{i=1}^{I} (R_i - R_f) \left[\beta_{0i} + \sum_k \beta_{ki} \times f_{kt} \right] + e_{jt}$$

In the single index model case ($I = 1$), the single index is the total return on the S&P 500 Index. In the three-index model case ($I = 3$) the first index is the total return on the S&P 500 Index, the second is the Ibbotson Small Firm total return orthogonal to the S&P 500 return, and the third is a government bond return orthogonal to the first two and composed of 80 percent intermediate term bonds and 20 percent long-term bonds (see Elton *et al.* (1992)). R_f is the total return on U.S. Treasury Bills with one month to maturity. The variables f_{kt} represent fund descriptors given in the left column. Variables 2 through 4 are dummy variables indicating style categories 1 if the fund belongs to the corresponding category, 0 otherwise (fund descriptors change significantly through time). This information and the remaining variables are taken from the previous year-end Weisenberger data. The case case is that of Growth funds, while *G&I, MCG,* and *INC* represent the style dummy variables 2 through 4. The regression results are obtained by weighted least squares, where the weights are proportional to the estimate of residual standard deviation for each fund. The Residual Risk column refers to regressing the log of absolute errors from the three-index equation for each fund on the corresponding fund descriptors. N signifies number, and DW signifies Durbin-Watson statistic.

	Single Model		Three-Index Models						Residual Risk	
	S&P Index Beta		S&P Index Beta		Small Firm Index Beta		Government Index Beta		Log of Absolute Error	
	Coefficient	t-Value	Coefficient	t-Value	Coefficient	t-Value	Coefficient	t-Value	Coefficient	t-Value
Constant	0.8311	124.06	0.8483	134.21	0.3346	35.93	−0.0027	−0.17	−4.2656	−229.54
Growth & Income	0.0833	13.96	0.0457	8.11	−0.3114	−37.91	0.0222	1.55	−0.4516	−18.46
Maximal Capital Gain	0.0642	4.72	0.0336	2.65	0.0797	4.51	0.0433	1.48	0.0994	3.65
Income	−0.1954	−12.49	−0.2073	−13.95	−0.1747	−8.14	0.1584	4.19	−0.2169	−4.46
Log net asset value	0.0355	13.10	0.0435	17.04	−0.0095	−2.68	−0.0191	−2.90	−0.1126	−12.49
Expense ratio	−0.0042	−3.58	−0.0020	−1.77	0.0097	6.11	0.0108	3.98	−0.0039	−1.21
Time since inception	0.0016	5.62	0.0007	2.63	−0.0035	−8.87	0.0025	3.78	−0.0063	−9.46
Time ∗ *G&I*	−0.0037	−12.36	−0.0028	−9.91	0.0046	11.10	−0.0002	−0.23	0.0040	4.47
Time ∗ *MCG*	0.0027	3.83	0.0046	6.98	0.0055	5.57	−0.0020	−1.30	−0.0014	−1.00
Time ∗ *INC*	0.0034	2.67	0.0043	3.58	0.0037	2.03	0.0008	0.27	−0.0210	−4.90
Diagnostics	$R^2 = 0.90$ $N = 42822$ DW = 1.987		$R^2 = 0.92$ $N = 42822$ DW- = 2.003						$R^2 = 0.035$ $N = 42822$ DW = 1.621	

The persistence tests for risk-adjusted returns are represented in Table VI. Risk adjustment does not appear to affect the pattern of persistence. In part, this may be due to the fact that systematic risk differences across managers as estimated by the model are not great. Depending upon which risk adjustment measures are used, we find that from five to seven years show evidence of significant persistence, and, as with the raw returns, 1980 to 1981 and 1987 to 1988 show evidence of a significant reversal in the pattern.

Brown *et al.* (1992) suggest using an appraisal ratio (the alpha measured in standard deviation units) to measure persistence because, in the presence of survivorship, ex post superior returns and alphas appear positively related to idiosyncratic risk. Scaling by this risk reduces the bias. The table reports tests using the single index and multiple index appraisal ratio. Neither appears to reduce the evidence for persistence. When single index appraisal ratios are used as opposed to Capital Asset Pricing Model (CAPM) alphas, the performance persistence pattern changes marginally, but the changes are not statistically significant.

An alternative way of adjusting for risk is to identify funds according to their style category. Is the repeat performance in the sample driven by the styles effect, or by individual fund deviations from the style average? We address this question by examining the performance persistence pattern of deviations from the average style return. The last panel in Table VI shows that the pattern of performance persistence is little affected by subtracting off the style benchmark. Clearly, if the Weisenberger style codes are meaningful, then the performance persistence in the sample is not driven by picking the winning management style each year. The fund reversals are due to correlation across fund managers; however, convential stylistic classifications fail to control for this correlation.

B. Absolute Benchmarks

In this section, we consider the effect of redefining winner as a fund that exceeds an absolute, rather than a relative benchmark. The simplest benchmark is the one that would have been most familiar to industry participants over the period of our study: the S&P 500. Figure 1 shows the effect of redefining a winner as a mutual fund that beat the total return of the S&P 500 in a given year. Notice that the *absolute* repeat-winner and repeat-loser pattern follows its own trend through time that closely matches the relative success of mutual funds reported in Table I. Over the second half of the sample, repeat-losers dramatically dominate. When results are aggregated across years, most of the persistence phenomenon is due to repeat-losers rather than to repeat-winners.

Table VI reports the result of using a risk-adjusted absolute benchmark. When a winner is defined as a fund with a positive alpha, the aggregate persistence pattern is only slightly affected. However, on a desegregated basis, we find that much of the effect is due to only a few statistically significant years in the sample period. Five of the years show significant

Table VI

Performance Persistence Patterns

This table reports cross-product ratios and Z-statistics for year-by-year performance persistence applied to a number of measures. The first is Raw Returns, calculated on an annual basis, assuming dividend reinvestment. The second is the Capital Asset Pricing Model (CAPM) alpha, estimated according to the model in Table V. The third is the Traynor-Black Appraisal Ratio, which is calculated as the CAPM alpha scaled by the residual standard deviation. The fourth is the Three-Index Appraisal Ratio, which calculated as the Three-Index alpha, described in Table V, scaled by the residual standard deviation. The fifth is the raw return minus the return for the fund style. The sixth defines winning funds as those with positive CAPM alphas. The seventh defines winners as funds with positive alphas, and it includes funds with a two-year performance in the lowest octile. Cross-product ratios and Z-statistics are defined as in Table IV. Asterisks indicate that the probability value of the bootstrapped statistic exceeds 0.95, while the Z-statistic does not.

	Raw Returns		CAPM Alpha		Treynor-Black Appraisal Ratio		Three-Index Appraisal Ratio		Style Adjusted		Positive Alpha		Positive Alpha Excluding Poor Performers	
Year	CPR	Z	CPR	Z	CPR	Z	CPR	Z	CPR	Z	CPR	Z	CPR	Z
1976	2.78	4.53	1.13	0.55	1.33	1.31	3.69	5.68	4.22	6.19	1.13	0.53	0.15	−1.58
1977	3.70	5.50	3.40	5.19	2.88	4.53	4.82	6.47	4.03	5.83	3.13	4.49	0.10	−2.15
1978	4.66	6.35	4.39	6.15	4.66	6.35	4.39	6.15	3.22	4.98	3.74	4.08	0.92	−0.22
1979	6.46	7.36	5.35	6.75	6.06	7.16	4.73	6.33	4.89	6.44	7.69	5.68	2.89	2.86
1980	0.25	−5.53	0.15	−7.32	0.16	−7.02	0.25	−5.65	0.26	−5.43	0.09	−7.90	0.25	−3.77
1981	2.49	3.85	1.44	1.56	1.16	0.64	2.17	3.28	2.23	3.40	1.07	0.28	1.06	0.16
1982	1.54	1.92*	1.50	1.80*	2.11	3.25	1.29	1.13	1.11	0.45	1.76	2.30	1.07	0.16
1983	1.19	0.81	1.31	1.25	0.92	−0.37	1.34	1.34	1.52	1.89*	1.13	0.40	0.21	−4.01
1984	1.20	0.88	1.44	1.78*	0.97	−0.14	1.05	0.25	1.38	1.58	1.34	0.97	0.50	−2.63
1985	2.53	4.78	4.74	7.65	4.56	7.48	3.00	5.60	2.78	5.24	4.59	7.26	1.82	2.49
1986	2.36	4.80	2.44	4.98	1.86	3.52	2.12	4.46	2.60	5.32	2.28	4.27	1.60	2.12
1987	0.50	−4.20	0.50	−4.21	0.41	−5.39	0.41	−5.380	0.42	−5.23	0.30	−6.60	0.41	−2.24

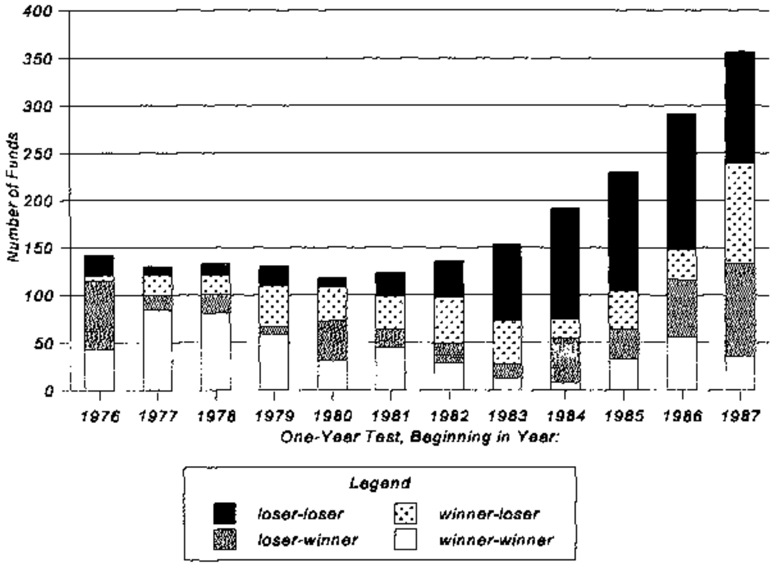

Figure 1. Frequency of Repeat Losers and Winners. The figure shows the effect of defining a winner as a mutual fund that beat the total return to the S&P 500 in a given year. The bars indicate the number of winning and losing funds each year that were winners or losers in the following year. Winner-Winner indicates a fund whose return exceeded the S&P 500 year.

positive persistence, two of the years show significant negative persistence, and five are ambiguous. While not reported in the table, we find that adding back expenses to returns also makes little difference in the overall persistence results for growth funds—consistent losers are not simply those with higher fees. The net effect of redefining winner in absolute terms is to reduce the *reliability* of the persistence effect for mutual funds. It matters little whether more sophisticated measure of absolute performance, such as multiple-factor appraisal ratios or positive alphas, are used.

C. Investment Implications

Can the persistence effect be used to earn excess risk-adjusted returns? In other words, can it be used to beat the market? Insight into this question can be gained by considering gradations of performance finer than the binary Winner-Loser classification we have used to this point. In Table VII we report average realized second year returns and alphas to a portfolio strategy where we invest an equal amount in funds that fall in each octile by total return in the first year. At the aggregate level, the results correspond to those reported in the earlier studies: top-octile performers do well, and bottom octile performers do poorly. It is interesting to note that these results are not sensitive to the choice of benchmark: we show results for the CAPM alpha computed using total returns on the S & P 500 Index, and for an equally weighted average of mutual funds in the sample (assuming, in this case, that the betas

Table VII

Rank Portfolios: Summary Statistics and Individual Year Results

Eight rank portfolios, equally weighted and reconstituted each year, are formed based on total return performance ranks for the year in the left column, with total returns on the portfolios computed for the following year. The results correspond to results reported in Hendricks, Patel, and Zeckhauser (1993), Table III, for a four-quarter evaluation period. Mean excess returns are measured in excess of 30-Day Treasury Bill returns; the standard deviations are also reported. Capital Asset Pricing Model (CAPM) Betas and Treynor-Black appraisal ratios are calculated using coefficients and residuals errors estimated using a prediction model based upon a monthly database of mutual fund returns. The methodology is described in the text and related notes. Alphas refer to Jensen's $-\alpha$ measure computed using returns *subsequent* to the portfolio formation year given in the leftmost column. EWMF Alpha refers to returns in excess of an equally weighted average of returns on funds in our sample. It corresponds to Jensen's-α where we assume fund betas are unity. The poor performer exclusion omits from the analysis funds whose two-year returns fall below the lowest octile of two-year performance for all funds in the sample. t-Values are in parentheses.

Panel A: Summary Statistics for Rank Portfolios

	1 (Worst)	2	3	4	5	6	7	8 (Best)	Best-Worst	Best-Worst Excluding Poor Performers
Mean Excess return	1.48	5.23	4.41	5.51	6.48	6.53	7.22	10.17	8.70	0.55
Standard Deviation	9.84	12.78	11.21	12.15	13.15	14.88	14.88	17.48	7.63	5.34
CAPM Beta	0.98	1	1.01	1.01	1.02	1.02	1.02	1.02	0.04	0.02
CAPM Alpha	−3.98	−0.30	−1.14	−0.01	1.04	0.99	1.65	4.64	8.62	0.66
	(−1.69)	(−0.17)	(−0.76)	(−0.01)	(0.59)	(0.51)	(0.75)	(1.46)	(2.29)	(0.17)
EWMF Alpha	−4.40	−0.65	−1.47	−0.37	0.60	0.66	1.35	4.29	8.70	0.55
	(−2.33)	(−0.47)	(−1.83)	(−0.69)	(0.99)	(0.92)	(1.39)	(1.73)	(2.15)	(0.14)

Table VII—*Continued.*

Panel B: CAPM Alphas on Rank Portfolios by Year

Year	1 (Worst)	2	3	4	5	6	7	8 (Best)	Treynor-Black Appraisal Ratio: Cross-Product	Treynor-Black Appraisal Ratio: Z-Statistic
1976	6.59	3.53	7.21	4.39	8.73	9.69	9.39	17.21	1.33	1.31
1977	0.68	−0.21	2.43	3.39	3.57	7.14	6.40	13.77	2.88	4.53
1978	−0.81	9.36	7.24	11.62	14.58	12.53	17.96	17.41	4.66	6.35
1979	−13.53	−2.62	−6.32	−1.67	−0.37	5.55	6.85	18.86	6.06	7.16
1980	10.72	5.77	3.77	5.81	4.37	−0.79	−0.36	−3.57	0.16	−7.02
1981	−0.60	−1.73	−2.06	0.58	1.98	5.10	5.81	2.75	1.16	0.64
1982	−4.56	7.78	−3.75	−3.43	−3.50	−3.35	−0.56	−1.27	2.11	3.25
1983	−15.11	−4.70	−7.35	−6.02	−7.19	−8.40	−7.17	−8.48	0.92	−0.37
1984	−13.65	−4.76	−5.53	−4.47	−1.98	−6.49	−6.09	−3.65	0.97	−1.14
1985	−5.66	−9.57	−5.37	−3.83	−2.35	−3.86	−2.92	7.16	4.56	7.48
1986	−10.48	−6.94	−3.94	−4.18	−5.20	−1.97	−3.34	9.62	1.86	3.52
1987	−1.34	0.52	0.01	−2.34	−0.11	−3.27	−6.14	−14.14	0.41	−5.39

of rank portfolios are all unity). Similar results are found for the three index benchmark, as well as using the S & P 500 total return as a benchmark.

We also report the effect of a simulated strategy of buying winners and shorting losers, indicated as the Best-Worst column in the top panel. We find that, if such a strategy were feasible, its benefits depend considerably upon the poor returns of those funds in the lowest octile. What happens if we eliminate these "bottom feeders" from the sample? When the persistent losers are eliminated, where persistent loser is defined as being in the lowest octile of *two-year* returns, the mean excess return to the strategy is positive, but insignificant.[12] The effect upon earlier tests of excluding bottom feeders is reported in the final two columns of Table VI. The significance is practically eliminated when the persistent losers are eliminated from the sample, and winners are defined as those with positive alphas or positive appraisal ratios.

Of course, investors are concerned with risk as well as return. In the top panel of Table VII, the average betas for the lowest and highest octiles are practically the same, but the annual standard deviation of returns to the octile portfolios differ considerably. Chasing winners is clearly a volatile strategy. While one might argue that investors are unconcerned with total risk because it is diversifiable, the strong correlation of winning funds noted in the earlier section suggests that diversification is not consistent with picking a portfolio of winners. Differences in risk are clearly evident in the annual decomposition of octile portfolio returns. While performance is correlated across all eight groups, the track record of the top octile is the most variable. It is interesting to note that, because of the high relative volatility of top octile funds, when they fail, they fail dramatically. For instance, the top octile performers in 1980 ended up in the bottom octile in 1981. Again in 1987, the top octile funds ended up in the bottom octile in 1988.

Regardless of the risk and return characteristics of chasing winners, the penalties to holding funds in the lowest octile are unambiguous. Preceding year performance appears to be an *excellent* predictor of negative alphas. Positive alphas were obtained in only three of the twelve years in the sample for the lowest octile. How can it be so easy to use simple historical information to identify a dominated asset in an efficient financial market? The answer seems to be the inability of investors to short most losing mutual funds. Investors can respond to poor performance by reclaiming shares, but not by arbitrage. The table focuses upon performance measures, but is silent

[12] Excluding those managers who perform poorly in aggregate over the two-year period does of course lead to an upward bias in excess returns; however, there is no reason to expect that it will affect *relative* returns. We verify this conclusion in two ways. First, we perform a simulation, reported in earlier versions of this article, in which we eliminate the lowest octile of *two-year* performers. We find that the two-year performance cut failed to induce differential performance across the octiles. An imposition of a 10 percent cut on the simulated sample based upon *one-year* returns does increase the lowest octile returns as expected. Second, we eliminate losers based upon a two-year period *preceding* the final year for which the evaluation is made. This alternative definition of bottom feeder only slightly increases the returns to the Best-Worst strategy.

about the money invested in the lowest octile funds. Evidence form Goetzmann and Peles (1993) suggests that only two to three percent of mutual fund investors hold shares in this bottom octile.

IV. Conclusion

Our study of a relatively survivorship-bias-free data set of equity mutual funds allows us to examine mutual fund performance with a database that largely controls for survivorship bias. We report basic summary measures for the Weisenberger equity fund universe. These directly show the magnitude of survival bias in mutual fund samples observed ex post. They also document the evolution of fads in the mutual fund industry. An analysis of the factors contributing to the disappearance of funds shows that a poor track record is the strongest predictor of attrition. Size, age, and the fund's expense ratio are also important. Attrition is negatively related to the relative growth of the fund assets due to new share purchases, but not strongly so.

The primary focus of the article is upon the issue of performance persistence. Our study takes a different tack from earlier researchers who identified the existence of repeat winners. By desegregating the persistence tests on an annual basis, we find that the phenomenon is strongly dependent upon the time period of study. This result is validated by Malkiel's (1995) analysis that extends the period of observation. Perhaps more significant than the uncertainty about repeat performance is the correlation across managers that is implied by the periods when the pattern is reversed. This suggests that future investigation of the persistence effect should concentrate upon a search for *common* management strategies. Recent candidates for such strategies include dynamic rebalancing of the type proposed by Connor and Korajczyk (1991), trend-chasing, identified by Grinblatt, Titman, and Wermers (1993), and common conditioning upon macroeconomic variables, suggested by Ferson and Schadt (1995).

Using methods designed to control for the survivorship bias identified by Brown *et al.* (1992), and a substantially larger database than previously assembled, we find clear evidence of *relative* performance persistence. Investors can use historical information to beat the pack. Evidence that historical information can be used to earn returns in excess of ex ante benchmarks, such as the S & P 500, positive appraisal ratios, and positive alphas is weaker and depends upon the time period of analysis.

An analysis of the risk and return characteristics of chasing the winners suggests that, while it is a positive alpha strategy, it also has a high level of total risk. Because of the correlation across winning funds, this total risk is not diversifiable, and thus it matters to risk-averse investors. Indeed the correlation of winning strategies suggests the possibility that winning funds are loading up on a macroeconomic factor, unassociated with the major components of equity returns, that may be priced. It is clear from these results that the nature of mutual fund persistence is more complicated than previous researchers, including the current authors, have understood. These issues are fertile ground for future inquiry regarding the basis for correlated active strategies among fund managers and motives for fund mergers.

REFERENCES

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.

Carlson, Robert S., 1970, Aggregate performance of mutual funds, *Journal of Financial and Quantitative Analysis* 5, 1–32.

Christensen, Ronald, 1990, *Log-Linear Models* (Springer-Verlag, New York).

Connor, Gregory, and Robert Korajczyk, 1991, The attributes, behavior, and performance of U.S. mutual funds, *Review of Quantitative Finance and Accounting* 1, 5–26.

Elton, Edwin J., and Martin J. Gruber, 1989, *Modern Portfolio Theory and Investment Management* (John Wiley and Sons, New York).

Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Matthew Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence for managed portfolios, *Review of Financial Studies* 6, 1–22.

Ferson, Wayne, and Rudi Schadt, 1995, Measuring fund strategy and performance in changing economic conditions. Working paper, University of Washington School of Business Administration.

Goetzmann, William N., and Roger G. Ibbotson, 1994, Do winners repeat? Patterns in mutual fund performance, *Journal of Portfolio Management* 20, 9–17.

Goetzmann, William N., and Nadav Peles, 1993, Cognitive dissonance and mutual fund investors, Working paper, Columbia Business School.

Grinblatt, Mark, and Sheridan Titman 1988, The evaluation of mutual fund performance: An analysis of monthly returns, Working paper, The John E. Anderson Graduate School of Management at UCLA.

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

Grinblatt, Mark, Sheridan Titman, and Russell Wermers, 1994, Momentum strategies, portfolio performance and herding, Working Paper, Johnson School of Management, UCLA.

Hendricks, Darryl, Jayendu Patel, and Richard Zeckhauser, 1993, Hot hands in mutual funds: The persistence of performance, 1974–1988, *Journal of Finance* 48, 93–130.

Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45–70.

Kane, Alex, Danilo Santini, and John W. Aber, 1991, Lessons from the growth history of mutual funds, Working paper, Boston University.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1992, The structure and performance of the money management industry, *Brookings Papers on Economic Activity: Microeconomics 339–391.*

Lehmann, Bruce N., and David Modest, 1987, Mutual fund performance evaluation: A comparison of benchmarks and a benchmark of comparisons, *Journal of Finance* 21, 233–265.

Malkiel, Burton, 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549–572.

Patel, Jayendu, and Richard J. Zeckhauser, 1992, Survivorship and the 'u' shaped pattern of response, *Review of Economics and Statistics*, Forthcoming.

Patel, Jayendu, Richard J. Zeckhauser, and Daryl Hendricks, 1990, Investment flows and performance: Evidence from mutual funds, cross border investments and new issues, in Richard Satl, Robert Levitch, and Rama Ramachandran, Eds.: *Japan, Europe and the International Financial Markets: Analytical and Empirical Perspectives* (Cambridge University Press, New York).

Sirri, Erik, and Peter Tufano, 1992, The demand for mutual fund services by individual investors, Working paper, Harvard Business School.

Weisenberger Inc., 1977 through 1989, *Investment Companies* (Weisenberger Investment Companies, New York).