This draft: August 9, 2001

MUTUAL FUND SURVIVORSHIP

Mark M. Carhart*, Jennifer N. Carpenter**, Anthony W. Lynch× and David K. Musto+

* Quantitative Strategies, Goldman Sachs Asset Management. email: mark.carhart@gs.com
** Stern School, New York University. email: jcarpen0@stern.nyu.edu
× Stern School, New York University. email: alynch@stern.nyu.edu
\+ Wharton School, University of Pennsylvania. email: musto@wharton.upenn.edu

Part of this research was conducted while Carhart was Assistant Professor at the Marshall School of Business, University of Southern California, and Senior Fellow, The Wharton Financial Institutions Center. We have benefitted from helpful comments from Stephen Brown, John Cochrane, Kent Daniel, Ned Elton, Gene Fama, Wayne Ferson, Will Goetzmann, Marty Gruber, Larry Harris, Bob Krail, Ananth Madhavan, Kevin J. Murphy, Russ Wermers, an anonymous referee, and the participants in seminars at the Western Finance Association annual meetings, Institutional Investor Institute Quantitative Round Table, University of California Riverside and the Institute for Quantitative Research in Finance. We also express appreciation to Bill Crawford and Lisa Simms at Micropal for assistance with their database. Carhart is grateful to Gene Fama, the Oscar Mayer Fellowship, and the Dimensional Fund Advisors Fellowship for financial support. Address correspondence and paper requests to Mark M. Carhart, Quantitative Strategies, Goldman Sachs Asset Management, 32 Old Slip, 24th Floor, New York, NY 10005. This paper was previously entitled, "Survivor Bias and Mutual Fund Performance."

MUTUAL FUND SURVIVORSHIP

ABSTRACT

This paper provides a comprehensive study of survivorship issues using the mutual fund data set of Carhart (1997). We demonstrate theoretically that when funds disappear primarily because of poor performance over several years, the average performance bias induced by survivor conditioning typically increases in the sample period length. This result is empirically relevant because Brown and Goetzmann (1995) find that funds disappear for exactly this reason. In the data, we find the annual bias increases from 0.07% for one-year samples to 1% for samples longer than fifteen years. We find empirically that survivor conditioning weakens evidence of performance persistence, as theory would suggest when survival depends on several years of past performance. Finally, we explain how survivor conditioning affects the relation between performance and fund characteristics and illustrate the effect empirically.

Fund disappearance, or attrition, affects almost every study of mutual funds, hedge funds, or pension funds. Many commercial data sets include only funds currently in operation. Test methodologies often require funds to survive a minimum time period to be included in the analysis. These forms of survivor-only conditioning can bias test results. This paper offers a theoretical and empirical analysis of the biases introduced by conditioning on survival. We study the effect of survivor conditioning on: (1) estimates of average performance, (2) tests of performance persistence, and (3) cross-sectional estimates of the relation between performance and fund attributes. In each case, the empirical results are consistent with the predictions of the theory.

Our database is virtually identical to the CRSP mutual fund data base, covering all known diversified equity mutual funds monthly from January 1962 to December 1995. Our paper is the first to use this data set to measure the effects of survivor conditioning on common mutual fund tests. Conditioning on survival can substantially alter the inferences from empirical tests, but the effects vary across the type of test, the form of survivor conditioning, and the sample period. Because survivor conditioning is relevant for many data sets and tests, the analysis in this paper has potential applications in other areas of financial economics.

To fix terminology, a single-period survival rule means that a fund with current period performance less than some threshold disappears at the end of the period, while a multi-period survival rule means that a fund disappears if its past *n*-period performance is less than some threshold. Some of the important theoretical insights about survivor biases pertain to a single-period rule (see, for example, Brown, Goetzmann, Ibbotson and Ross, 1992) . However, our theoretical work and that of others indicate that the effects of survivor conditioning depend critically on the nature of the survival rule in effect (see, for example, Brown, Goetzmann, Ibbotson and Ross, 1992 and Carpenter and Lynch, 1999). Evidence that lagged performance predicts survival, even in the presence of the most recent year's performance, suggest a multi-period survival rule for U.S. mutual funds (see Brown and Goetzmann, 1995).

Our paper contains several new results. We begin with the theoretical and empirical effects of survivor conditioning on estimates of average performance. We show that while a single-period survival criterion implies a constant survivor bias in estimates of average performance, a multi-period survival criterion typically causes survivor bias to increase in the sample length, though at an ever decreasing rate. Empirically, we find that the bias in annual performance is increasing in the sample length and is approximately 1% for subsets of our data longer than fifteen years. Theoretically, we explain why the bias in average performance need not always increase in the sample length, even with a multi-period survival rule.

We next examine empirically the impact of survivor conditioning on persistence tests, and find that the conditioning attenuates performance persistence relative to the full sample. This result could be anticipated theoretically given the multi-period nature of the industry's survival rule (see Brown, Goetzmann, Ibbotson and Ross, 1992, Grinblatt and Titman, 1989, and Carpenter and Lynch, 1999). However, our paper is the first to demonstrate the result empirically.[1]

Finally, we explain how survivor conditioning can affect the cross-sectional relations obtained between fund performance and fund characteristics. In particular, for the cross-sectional relation to be biased in a survivor-only sample, the fund characteristic in question must be related to the survivor bias in performance. We estimate the slope coefficient biases for commonly-used fund characteristics, find that the magnitude of these biases can be large, and show that their directions are consistent with intuition. We estimate the Heckman (1976, 1979) two-step correction for incidental truncation, and find that model misspecification may be a serious concern when attempting to use this procedure to control for the survivor biases.

Several recent papers have constructed mutual fund databases that attempt to control for survivor biases. Elton, Gruber and Blake (1996) follow the cohort of funds listed in Wiesenberger's 1977 volume from 1976 until 1993, constructing complete return histories up to the date of merger for funds with assets greater than \$15 million. Brown and Goetzmann (1995) use annual returns from 1977 to 1988 estimated from Wiesenberger's *Investment Companies*, while Malkiel (1995) uses quarterly returns from 1971 to 1991, obtained from Lipper Analytical Services.[2]

Section 1 describes methodology and the data. Section 2 considers the effects on average performance measures of requiring the sample funds to survive to the end of the sample period. Section 3 studies the effects of survivor conditioning on persistence measures, while section 4 examines the impact of survivor conditioning on cross-sectional regressions. Section 5 concludes.

1. Methodology and Data

A. Aggregation method

Since a mutual fund sample is a panel data set, a method of aggregation across funds and time must be chosen. One approach is to pool all of the time-series and cross-section observations. Due to significant recent growth in the number of funds, this method skews results towards relations in the final few years of the sample. A second approach calculates statistics on the individual funds, then averages cross-sectionally. This approach gives the same weight to all funds irrespective of history length. Thus, dates at the end of the sample and funds with relatively few years of returns get more weight than in the first approach. A third approach calculates statistics cross-sectionally for each time period and then averages these estimates through time. This approach gives less weight to observations that occur on dates with many other observations, which seems reasonable given that the contemporaneous fund performances are likely to be correlated. We rely primarily on this third approach for aggregation.

B. Performance measurement

We employ two measures of performance. The first measure, group-adjusted performance, is the fund return minus the equal-weight average return on all funds with the same objective in that period. The different fund objectives are aggressive growth, growth and income, and long-term growth. When funds change objectives, they move to a new group.[3] The second performance measure is the time-series regression intercept, or alpha, from the 4-factor model of Carhart (1997). The 4-factor model uses Fama and French's (1993) 3 factors plus an additional factor capturing Jegadeesh and Titman's (1993) one-year momentum anomaly. The model is

$$r_i(t) = \alpha_i + b_i\, RMRF(t) + s_i\, SMB(t) + h_i\, HML(t) + p_i\, PR1YR(t) + e_i(t) \qquad \textbf{(1)}$$

where r_i is the return of asset *i* in excess of the one-month T-bill return, 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. We use the 4-factor model in an effort to adjust fund performance for well-known regularities in stock returns. It would also be interesting to assess performance using a conditional model like Ferson and Schadt (1996), but we leave this to future work.

C. *Survivor conditioning*

It is important to recognize that the survival criterion in effect for the industry interacts with the survivor conditioning in a sample to generate survivor biases in test statistics. In empirical work, the researcher is stuck with the market's survival criterion, but may have considerable control over the survivor conditioning imposed on the sample. Two forms of survivor conditioning are particularly important for mutual fund research. End-of-sample conditioning includes only the funds extant at the end of the sample period. Look-ahead conditioning requires funds to survive some minimum length of time after a reference date, known as the look-ahead period.[4] End-of-sample conditioning is usually a property of the data set, whereas look-ahead conditioning is typically imposed by the test methodology. An example of end-of-sample conditioning is Morningstar OnDisc, which reports performance since January 1976, but only for funds still existing at the end of the sample period. An example of look-ahead conditioning is the common performance persistence test methodology that regresses future n-period performance on a measure of past performance: the test conditions on survival for n periods beyond the evaluation date. In fact, some degree of look-ahead conditioning is inherent in any test of performance persistence. Since the imposition of a minimum survival period is often unavoidable, an important issue is how the resulting bias varies with the nature of the survival rule in effect for the industry.

D. Database

Our database covers all known diversified equity mutual funds monthly from January 1962 to December 1995, excluding sector funds, international funds, and balanced funds. We obtain data on surviving and nonsurviving funds from a variety of sources (see Carhart, 1997, for details). The sample includes a total of 2,071 diversified equity funds, 1,346 of them still operating as of December 31, 1995. We partition the sample into three primary investment objectives using Wiesenberger and ICDI classifications: aggressive growth, growth and income, and long-term growth.

The data set includes monthly returns and annual attributes. Return series do not include final partial-month returns on merged funds as in Elton, Gruber, and Blake (1996). Of the 725 nonsurviving funds, we obtain the date of merger, liquidation, or reorganization for 475. Within the sample of funds with known termination dates, the return series end within one week of the termination date for 330 funds. Of the remaining 145 funds, 32 do not include the final partial- or full-month return, 20 do not include the final two- to three-month return, 81 do not include the final four- to twelve-month return, and 12 funds are missing more than one year's returns. Of the 250 nonsurviving funds without exact termination dates, we do not observe any returns on 53 funds, often because they are too small to appear in any published sources.[5]

E. Summary statistics

The average annual fund attrition rate from 1962 to 1995 is 3.6%, with a standard deviation of 2.4%. Of the total 3.6%, 2.2% per year disappear due to merger and 1.0% disappear because of liquidation.[6] In the subsamples grouped by investment objective, aggressive growth funds perish at an annual rate of 4.5%, which is statistically significantly larger than the analogous rates of 2.9% for long-term growth funds and 3.3% for growth-and income funds. About 58% of all defunct funds disappear because of merger and 36% disappear due to liquidation. A further 2% vanish through other self-selected means, usually at the fund manager's request for removal, and the remainder depart for unknown reasons or are dropped from the sample by the database manager, not the fund itself.

Table 1 compares the performance of surviving and nonsurviving funds. Not surprisingly, nonsurviving funds exhibit considerably poorer performance than surviving funds. The cross-sectional average monthly performance estimates are the cross-sectional averages of the group-adjusted returns and 4-factor alphas of individual funds, estimated from the complete time series of their returns. By these measures, nonsurviving funds underperform survivors by about 4% per year. Liquidated funds exhibit the worst relative performance.

Table 1 also gives 4-factor model estimates for equal-weighted portfolios of funds. Nonsurviving funds remain in the equal-weighted average until they disappear.[7] The alphas for the portfolios are close to the average of the individual alphas of the funds in the portfolios, suggesting that the results are not sensitive to the method of aggregation. The performances of the portfolios of survivors and nonsurvivors are considerably different. Survivors achieve abnormal performance of -0.07% per month while nonsurvivors earn -0.33%. The difference between estimates of performance using survivors only and estimates using the complete sample is 0.08% per month. From the 4-factor loadings, we infer that relative to nonsurvivors, surviving funds have a less negative exposure to high book-to-market stocks, less positive exposure to small stocks, and similar exposures to the market and to the momentum factor.

2. Survivor Bias Effects on Estimates of Average Performance

A. Theory

For convenience, we call the periods years, though they could be any length of time. An m-year survival rule causes funds at least m years old to disappear through liquidation or merger if the sum of their returns over the preceding m years falls below a threshold b. Returns could be any measure of performance. We assume that fund returns are cross-sectionally and intertemporally independent and identically distributed with mean μ.[8] Let $g \geq 0$ be the annual growth rate of the number of funds in the mutual fund industry.[9]

Let k be the length of the sample period of interest and let T be its ending date. By assumption, b, g, μ, and the variance of fund return are all independent of k. Consider the sample of all funds that have returns in the sample period and that are still in existence at time T, including new funds over the period. By construction, this sample imposes end-of-sample conditioning and includes funds with fewer than k years of performance. The estimate of average performance for an equal-weighted portfolio $\bar{\mu}_k^T$ is the time-series average of the yearly equal-weighted cross-sectional mean returns of these funds. We are interested in characterizing the behavior of $\bar{\mu}_k^T$ as a function of k.

Proposition 1*:* If a single-year survival rule causes fund disappearance (i.e., m=1), the annual end-of-sample bias in the average performance estimate is independent of the length of the sample period, k.

Proof: In any year of any sample period, the bias in the estimate of average performance of surviving funds is

$$E\left[R|R>b\right] - \mu \tag{2}$$

which is independent of k. ***Q.E.D.***

Now suppose a multi-year survival rule determines fund survival (i.e., m>1). Each of the funds that survive through time T survives performance cuts from the time it is m-years old until time T. Let C_t be the conditioning statement associated with the time t performance cut:

$$C_t \equiv [\ (\sum_{\tau=t-m+1}^{t} R_\tau) > b\ ], \tag{3}$$

and let x_i denote the survival probability after year t conditional on survival in previous years:

$$\begin{aligned} x_i &\equiv pr\ \{C_t \mid C_{t-1},\ldots, C_{t-i}\}, && i>0 \\ &\equiv pr\ \{C_t\}, && i=0 \\ &\equiv 1, && i<0. \end{aligned} \tag{4}$$

Let $\mu_{i,j}$ be the expected one-year return conditioned on having survived a set of j+1 consecutive performance cuts with the last cut occurring i-1 years after the return:

$$\mu_{i,j} \equiv E[R_{t+1-i} | C_{t-j}, \ldots, C_t]. \tag{5}$$

We define a cut whose return window includes the given return as a *direct* cut; otherwise, we have an *indirect* cut. If $\mu_{i,j}$ only involves indirect cuts, it must be equal to μ. With an m-year survival rule, there are m direct cuts for R_t which occur at times t through (t+m-1): C_t, ...,C_{t+m-1}. All other cuts are indirect with respect to R_t. For example, the time-(t+m-1) cut is direct with respect to R_t, because it is applied to the sum of R_t through R_{t+m-1}, which is a return window that overlaps with R_t. In contrast, the time-(t+m) cut is indirect with respect to R_t, because it is applied to the sum of R_{t+1} through R_{t+m}, which is a return window that does not overlap with R_t.

It might seem that indirect cuts should not affect a conditional mean return, but this is not the case. For example, even though R_t and C_{t-1} are independent, $E[R_t | C_{t-1}, C_t]$ is not equal to $E[R_t | C_t]$, because of the dependence between C_t and C_{t-1}. Nevertheless, imposing an additional direct cut tends to have a much greater effect on the conditional mean than imposing an additional cut that is indirect. Intuition suggests that the conditional mean of R_t is increasing in the number of direct cuts. This intuition implies a survivor-biased k-year sample mean $\bar{\mu}_k^T$ that is increasing in k when the survival rule uses more than one year of returns.

To illustrate why, let's consider a two-year survival rule. With m=2, the conditioning statement associated with the time-t performance cut becomes: $C_t \equiv [R_{t-1} + R_t > b]$. For simplicity, suppose that the expected 1-year return conditioned on having survived a set of performance cuts depends only on the number of direct cuts. Let μ_1 and μ_2 be the expected 1-year return conditional on one and two direct cuts respectively. Since intuition indicates that the conditional mean 1-year return is increasing in the number of direct cuts, we assume that $\mu < \mu_1 < \mu_2$.

With end-of-sample conditioning, the mean year-T return of 1-year-old funds is always the unconditional mean of μ since those funds are too young at the end of the sample to be subject to any cut. In earlier years, the mean return of the 1-year-old funds is the conditional mean return with one direct cut, μ_1, since a fund's first return can only be subjected to one cut. Similarly, the mean year-T return of all older funds is also μ_1, since a fund's last return in the sample can be subjected to at most one cut. In earlier years, the returns of funds older than 1 year at the time are subjected to two direct cuts, and so their mean return is μ_2.

Let μ_t^T be the cross-sectional mean year-t return of funds that survive through time T. The cross-sectional mean year-T return of funds that survive through time-T is a weighted average of the mean year-T return for 1-year-old funds, μ, and the mean year-T return for older funds, μ_1:

$$\mu_T^T = \hat{w}_{1,T}^T \, \mu + (1 - \hat{w}_{1,T}^T) \, \mu_1 \tag{6}$$

where $\hat{w}_{j,t}^T$ is the fraction of funds that are j years old at time t in the set of time-T survivors (after the time-T cut) that have a year-t return. For year $t < T$, the cross-sectional mean fund return of funds that survive through time T is a weighted average of the mean return for 1-year-old funds, μ_1, and the mean return for older funds, μ_2:

$$\mu_t^T = \hat{w}_{1,t}^T \, \mu_1 + (1 - \hat{w}_{1,t}^T) \, \mu_2. \tag{7}$$

Since the conditional one-year mean return is increasing in the number of direct cuts ($\mu < \mu_1 < \mu_2$), the cross-sectional mean fund return with end-of-sample conditioning is lower for the last year (T) than for the next to last year. Moreover, if the fraction of one-year-olds is constant over time (i.e., $\hat{w}_{1,t}^T \equiv \hat{w}_1$ for all t), the cross-sectional mean fund return with end-of-sample conditioning is the same for all years but the last. Consequently, the survivor-biased estimate of average performance across the k-year sample ending at T, $\bar{\mu}_k^T$, is increasing in k. The impact of the lower time-T cross-sectional mean on the time-series average becomes smaller as k increases.

While not generally true, the fraction of one-year olds in the sample of time-T survivors will be the same in all years if the probability of surviving a cut is the same irrespective of the number of cuts already survived (i.e., $x_i = x$ for all $i \geq 0$) and if the mutual fund industry is in steady-state growth. Letting $w_{j,t}$ be the fraction of funds with a time-t return that are age j at time t, we say that the mutual fund industry is in steady-state growth, given its growth rate g, if the age distribution is the same each year: i.e., $w_{j,t} = w_{j,\tau}$ for all j and any t and τ. To see why the assumptions of a steady state and a constant survival probability imply a constant fraction of one-year-olds, first note that with the industry in a steady state, the time-t age distribution conditional on survival through t is the same for all t. Second, note that with the probability of surviving a cut always the same, the time-t distribution of one-year-olds and older funds is the same conditional on survival through t or through any later date (i.e, $\hat{w}^{\tau}_{1,t} \equiv \hat{w}_{1,t}$ for all $\tau \geq t$). The reason is that both groups leave the sample at the exact same rate per year, (1-x), from time-(t+1) onward. Together, these results imply that the time-t age distribution conditional on survival through T must be the same for all t, as required.

The following proposition shows that this intuition generalizes to m-year survival rules with arbitrary $m \geq 2$.

Proposition 2*:* If an m-year survival rule causes fund disappearance, m>1, and

1) the conditional mean $\mu_{i,j}$ only depends on and is strictly increasing in its number of direct cuts: for all (i,j) pairs, $\mu_{i,j} = \mu_\tau$ where $\tau \in \{1,2,...m\}$ is the number of direct cuts involved in $\mu_{i,j}$; and, $\mu<\mu_1<\mu_2<...<\mu_m$;

2) the probability of surviving a cut is the same irrespective of the number of cuts already survived: $x_i = x$ for any $i \geq 0$;

3) the mutual fund industry is in a steady state: $w_{j,t} = w_{j,\tau}$ for all j and any t and τ;

then the end-of-sample bias in the average performance estimate $\overline{\mu}_k^T$ is increasing in the length of the sample period, *k*.

Proof: Appendix A.

Moving back in time from *T*, the cross-sectional mean increases for the first *m* years, at which point it reaches a steady-state value. The *m* means at the end of the sample can be expected to be lower since these returns are subjected to fewer direct cuts. As *k* increases, the greater weight on the steady-state means increases the sample average. This mechanism for delivering the result is a generalization of the *m*=2 case discussed above.

Since none of the three assumptions holds in general, it is possible to construct examples in which the sample mean is not increasing in the sample length.[10] Although direct cuts are generally expected to increase the conditional mean of R_t, indirect cuts can reduce this conditional mean. Roughly speaking, when direct cuts to R_t have already been applied, the lower part of the distribution of R_t has already been eliminated. Imposing incremental indirect cuts to R_t can then eliminate return paths that involve mainly good realizations of R_t, reducing its conditional mean.

Another complication is that funds of different ages may disappear at different rates, causing the weights of the different aged cohorts to change over time. Recalling that $x_i = pr\ \{C_t \mid C_{t-1},..., C_{t-i}\}$, it makes sense that x_i is changing as *i* goes from 0 to *m*-1 since each additional cut overlaps with C_t. However, x_i also varies as a function of *i* for *i*>*m*-1, because of the interaction of the cuts $C_{t-i},...,C_{t-m}$ with the cuts $C_{t-m+1},...,C_{t-1}$.

Finally, if the assumption of a steady state is relaxed, the cross-sectional mean may start declining in *k* for *k* sufficiently large, since the earliest years have only young funds whose early-year returns have few direct cuts. Thus, $\overline{\mu}_k^T$ may be hump-shaped as a function of *k*, rather than increasing.

More generally, the nature of the bias depends on the distribution of funds at the start of the sample. The construction of the sample is also important. Here we focus on the effects of end-of-sample conditioning for a sample that adds new funds as returns become available. Alternatively, the sample may follow a set of funds in existence at a point in time, as in Elton, Gruber and Blake (1996), and impose end-of-sample conditioning. The end-of-sample biases will be different for this sample because the cross-sectional distribution of funds will differ. For example, with fewer young funds, average fund volatility might be lower, leading to smaller survivor biases.

B. Calibration

Intuition developed in the previous subsection suggests that end-of-sample conditioning creates a bias in average performance that is increasing in the sample length. This effect follows from the fact that a return near the end of the survivor-only sample is conditioned on fewer direct cuts than other returns in the sample. However, the previous subsection also describes some counterintuitive effects of increasing the length of the sample period. To illustrate how the Proposition 2 effect typically dominates the counterintuitive effects in a more realistic setting, we generate a mutual fund history designed to match the U.S. mutual fund industry.

For each $m \in \{1, 2, 3, 4, 5, 10\}$ we simulate values for the conditional means $\mu_{i,j}$ and the survival rates x_j assuming that returns R_t are normally distributed with mean zero and standard deviation 5%. We set the growth rate in the industry equal to 5.5% to match the growth rate in the data. The choice of the critical return value *b* determines the average attrition rate for the sample. We choose this critical level in one of two ways. Panel A of Table 6 allows *b* to vary across *m* in such a way as to maintain a sample average attrition rate of 3.5%, the average annual attrition rate in the data. Panels B and C fix $b/\sqrt{m}$ at -9.06%, which makes the sample average attrition rate for the case $m = 1$ equal to 3.5%. For larger values of m, this choice of b leads to lower sample average attrition rates.

We make two assumptions about the starting composition of the industry. Panels A and B assume that all funds are m years old at time 1 which implies that the sample starts m-1 years after the industry. Panel C assumes that all funds are one year old at time 1, which implies that the sample and the industry start at the same time.

For a given subperiod length of k in the 34-year history, we compute average performance measures for survivor-only samples using the simulated conditional mean returns and attrition rates. In particular, when m is greater than 1, we do not impose conditions 1) or 2) of Proposition 2 but rather let the simulations determine the conditional means and survival probabilities. Finally, for each k, we average the performance estimate across all possible subperiods of length k. Table 2 reports this average (in percent) for k = 1, 2, 3, 4, 5, 10, and 30, and the change in this average (in basis points) for k going from 30 to 31, 31 to 32, 32 to 33 and 33 to 34.

Consistent with Proposition 1, all three panels show that the survivorship bias in the average performance is constant across k for $m = 1$. Turning to the cases with $m>1$, the first two panels of Table 2, which have only m-year-olds at time 1, show that the bias uniformly increases in sample period length k for $m > 1$. In contrast, Panel C only has one-year-olds at time 1, and the intuition described earlier causes the sample average as a function of k to start declining for k close to 34. However, even in this extreme case in which the sample starts with all one-year-olds, the largest decline for a one-year increase in k is only 0.04 basis points. Thus, we conclude that the counterintuitive effects of increasing the sample period length discussed in the previous subsection are not likely to play an important role in realistic settings like the U.S. mutual fund industry.

C. Evidence

We now examine the relation between average performance bias and sample period length in the data. Brown and Goetzmann (1995) find that past annual performance out to at least 3 lags affects fund survival, though in a more complicated fashion than the multi-period rule described in subsection 3.A. above. The calibration results suggest that the Proposition 2 effect is likely to outweigh the counterintuitive effects in realistic settings. Nevertheless, we would like to assess whether the calibration results still apply qualitatively to the U.S. mutual fund industry, despite the additional complexity of the empirical survival rule. We would also like to measure the magnitude of the bias in average performance, as a function of the sample length. We find that the empirical relation between the bias and the sample length is strong and positive, as the calibration results suggest.

We consider all the possible samples with end-of-sample conditioning that might be assembled from our database over the 1962 to 1995 period. For example, a researcher might assemble a five-year sample in 1972 or a ten-year sample in 1985. For each sample period length k, we consider all the possible (usually overlapping) annual return samples, and estimate the bias in average annual group-adjusted return induced by including only survivors. We report the average end-of-sample bias across all possible *k*-year samples. We also calculate correlation-adjusted standard errors assuming that the survivor bias in annual sample equal-weighted return is independent and identically distributed.[11]

Table 3 shows that the survivor bias increases in the sample length. For a survivor-biased sample of only one year, the bias in average return is only 0.07%, whereas the bias is 0.37% per year for survivor-biased samples of five years. For samples greater than fifteen years, the hypothesis that survivor bias is 1% per year is not rejected. So consistent with the calibration results in the previous subsection, the bias is an increasing concave function of sample length that is virtually flat at sufficiently long sample lengths. Figure 1 plots the survivor bias as a function of the sample length. For time periods of fifteen years or longer, 1% is probably a good approximation of the bias in mean annual performance estimates introduced by end-of-sample conditioning.

3. Survivor and Look-Ahead Bias Effects on Estimates of Persistence in Performance

A. Theory

Persistence is defined as a positive relation between performance in an initial ranking period and a subsequent evaluation period. Brown, Goetzmann, Ibbotson and Ross (1992) show that if mutual fund returns are independently distributed with the same mean but differing variances, and if a single-period survival rule causes fund disappearance, then tests on surviving samples show spurious persistence. Conditional on making the ranking-period cut, higher volatility funds have higher means. In a sample of survivors, the same high volatility funds tend to win in the subsequent evaluation period. Brown, et al. (1992) also demonstrate a spurious reversal effect. In the absence of cross-sectional dispersion in volatility and in the presence of a multi-period survival rule, survivorship bias causes spurious reversals instead of persistence in performance. A multi-period survival rule removes loser-losers in greater proportion than winner-losers, loser-winners, or winner-winners, leaving the sample more heavily weighted toward reversers. Grinblatt and Titman (1992) make a similar argument. Carpenter and Lynch (1999) study these effects when both cross-sectional dispersion in fund volatility and a multi-period survival rule are present. They find that although the spurious persistence effect stemming from cross-sectional dispersion in volatility is always at work, the reversal effect tends to dominate when the multiple-period survival rule is in force. Which of the various effects dominates in the data depends on the nature of real survival rules and the degree of cross-sectional dispersion in volatility.

B. Evidence

This section studies the effect of end-of-period and look-ahead bias on the persistence tests of Hendricks, Patel and Zeckhauser (1993) and Carhart (1997) in our sample of U.S. mutual funds. Annually, we form ten equal-weighted portfolios of mutual funds sorted on either lagged return or lagged 4-factor alpha. We hold the portfolios for one year, then re-form them. This yields a time-series of monthly returns on each portfolio from 1962 to 1995 less the initial performance estimation period. The performance measures are one-year return, five-year return, and three-year estimates of alpha from the 4-factor model. Funds disappearing during the ranking period are not used to determine the performance deciles, but if a fund disappears during the evaluation period, its returns are included in the decile performance averages right up until the time the fund disappears. At that point, its decile portfolio is re-weighted equally across the remaining funds.

Panel A of Table 4 reports tests of persistence in fund returns and 4-factor alphas for three different samples. The "full" sample includes all returns on disappearing funds in our database. Consistent with Carhart (1997), the full sample portfolios demonstrate strong persistence in mean return, most of which is explained by the 4-factor model. The end-of-sample-biased portfolios show less persistence. Spreads in mean return and 4-factor model performance shrink considerably relative to the full sample, and the statistical significance diminishes as well.[12] The look-ahead biased sample requires that funds survive a look-ahead period after portfolio formation that is equal in length to the ranking period. That is, the lagged one-year results include only funds surviving a full year after sorting on the previous year's return, and the lagged five-year sample requires survival for an additional five years after sorting.[13] Using the look-ahead biased sample changes the inference relative to the full sample only for the five-year returns-sorted portfolios, the longest look-ahead period.

Finally, we undertake Hendricks, Patel and Zeckhauser's (1996) test for spurious persistence due to survivorship. The HPZ J-shape t-statistic is the t-statistic on the linear term of a quadratic regression of the evaluation period portfolio rank on the ranking period portfolio rank. Under the hypothesis that performance persists spuriously due to survivorship, the HPZ J-shape t-statistic should be reliably negative. However, Carpenter and Lynch (1999) present simulation evidence that the HPZ J-shape t-statistic is rarely reliably positive unless performance is truly persistent. We find that the HPZ J-shape t-statistics are all positive and often significant in our survivor-biased samples.

Myers (2001) finds that persistence in pension fund performance is attributable to differences in fund returns across fund styles rather than within a style. To investigate this possibility for mutual funds, Panel B of Table 4 repeats the tests of Panel A using group-adjusted returns instead of raw returns to measure evaluation-period performance. We include 4-factor alphas for group adjusted-returns to capture some of the elements of fund style that grouping by fund objective might miss.[14] In general, the evaluation period's decile spreads for both the group-adjusted returns and their 4-factor alphas are of magnitudes similar to those for the raw returns and their 4-factor alphas, respectively. This suggests that the persistence in raw returns and their four-factor alphas reflects more than just differences in fund style.

To summarize, both the end-of-sample and look-ahead conditioning reduce the degree of persistence regardless of whether the performance measure is group-adjusted or not. The downward bias in the persistence measures induced by survivor conditioning is consistent with the theory, given the multi-period nature of the survival rule documented by Brown and Goetzmann (1995). Since fund performance exhibits persistence in all three samples, our results provide further evidence that the performance of U.S. mutual funds is truly persistent.

4. Effects of Survivor Bias on Cross-Section Tests

A. Theory and Evidence

Survivor-only conditioning can affect estimates of the cross-sectional relations between fund performance and fund characteristics, but only when the fund characteristics in question are related to the survivor bias in performance. The direction and magnitude of the characteristic's impact on the performance bias determine the direction and magnitude of the slope coefficient bias in the cross-sectional regression. If the survivor bias in performance is positively related to the fund characteristic, the characteristic's slope coefficient in the cross-sectional regression also possesses positive bias in the survivor-only sample. Conversely, if the performance bias is negatively related to the characteristic, the slope coefficient is downward biased. The biases in cross-sectional regressions introduced by survivor conditioning are an example of sample selection, or incidental truncation, which has been the subject of an enormous recent literature, both theoretical and applied.[15]

To illustrate the problem, we run pooled time-series, cross-sectional regressions of annual group-adjusted returns (as defined in section 1) on five explanatory variables: net expense ratio, relative turnover, lagged relative total net assets (TNA), lagged maximum load fees, and lagged annual group-adjusted returns. The fund's relative TNA at the end of *t* is the fund's TNA divided by the average TNA of all other funds on that date. The fund's relative turnover is its modified turnover over the average modified turnover of all funds that year.[16] The fund's maximum load fee in year *t* is the sum of maximum front-end, back-end and deferred sales charges in that year.

We run two sets of regressions. In the first set, there is one simple regression per explanatory variable, whereas in the second set, the five explanatory variables are taken together in a multiple regression. Within each set, we run two regressions. The first regression uses the "full" sample and is a Seemingly Unrelated Regressions model (SUR). The full sample uses all available returns prior to a fund's disappearance. A fund's year-y+1 return is deemed available if we observe its TNA, expense ratio and sales loads in year *y*. Since a fund does not typically have a full-year return in its year of disappearance, it is not possible to run an OLS regression with full-year return as the dependent variable and still include all available monthly returns. Instead, the SUR model consists of twelve regressions with the twelve separate monthly returns in fund-year y+1 as the dependent variables. For each independent variable, we sum the coefficients across the twelve regressions to get a full sample slope coefficient for annual return as the dependent variable.

By contrast, the second regression uses the "5-year look-ahead" sample and uses the fund's annual y+1 group-adjusted return as the dependent variable. The 5-year look-ahead sample includes only the available year-y+1 returns of funds that did not disappear in years y+1 through y+5. The second regression is OLS.

The slope coefficients from the simple regressions on each explanatory variable are in the first two columns of Table 5 while the slope coefficients from the multiple regression are in the fifth and sixth columns. T-statistics for significant differences from zero are below the coefficients.[17]

The full sample results indicate a coefficient on the contemporaneous expense ratio which is insignificantly different from -1 in both the simple and multiple regressions. This result indicates a 1:1 tradeoff between performance and expenses (i.e., more expensive funds do not perform better, they are just more expensive). Consistent with the persistence results in the previous section, the coefficient on lagged group-adjusted return is significantly positive in both regressions for the full sample. Most other coefficients are either insignificant or not robust across the two specifications. For example, the positive and significant simple relation between lagged fund size and group-adjusted returns disappears in the multiple regression.

With regard to the effects of survivor conditioning, Table 5 shows that the regression coefficients from the 5-year look-ahead sample are very different from those obtained using the full sample. For example, the simple regression slope on relative TNA goes from being positive and significant in the full sample to negative and insignificant in the 5-year look-ahead sample. The multiple regression slope on expense ratio declines from -1.150 to -0.440 going from the full to the 5-year look-ahead sample. While this slope is significantly different from 0 but not -1 in the full sample, the converse is true in the survivor-only sample.

For a given fund characteristic, the direction and magnitude of the bias in the slope coefficient is determined by the correlation of the characteristic with the survivor bias in performance. The results of a probit analysis described in Appendix B can be used to infer the correlation. From Table 6, we see that the probability of disappearance during years $y+1$ through $y+5$ is negatively related to fund size at the end of y, holding other fund characteristics fixed. This suggests that the survivor bias in performance is likely to be decreasing in fund size. This implies a negative survivor bias in the slope coefficient on fund size, which is exactly what Table 5 shows.

A similar argument explains the positive survivor bias in the coefficient on net expense ratio.

B. Heckit Correction

Heckman (1976, 1979) develops a correction for incidental truncation that has been used in a variety of applications. It has become a standard technique to adjust for sample selection. We would like to assess its ability to correct for the effects of survivor conditioning in the cross-sectional regressions. To this end, we run a third regression which uses the "5-year look-ahead" sample and uses the fund's annual y+1 group-adjusted return as the dependent variable. This regression implements the two-stage "Heckit" procedure by adding, as an additional regressor, the inverse Mills ratio, commonly denoted λ. The inverse Mills ratio is obtained from a probit model that predicts disappearance in years y+1 through y+5 using variables available at the end of year y. Details of the probit are contained in Appendix B. Since we have the full sample, we can assess the ability of the Heckit estimator to correct for survivor biases in the sample with 5-year look-ahead conditioning.

The cross-sectional regression results using the Heckit procedure are also reported in Table 5. The Heckit procedure generally moves both the point estimates and significance levels toward those for the full sample in the multiple regression, but not in the simple regressions. An explanation for this finding requires an examination of the loadings on λ. Conditioning on fund characteristics available at the end of year y, we expect a fund's return in year y+1 to be positively correlated with the fund's probability of survival in years y+1 to y+5. Consequently, theory tells us that the coefficient on λ is positive. But while the loading on λ is positive and significant in the multiple regression, it is negative in all but one of the simple regressions, the exception being the one with lagged group-adjusted return as the independent variable.

Now the Heckit λ is negatively related to the probit's predicted survival probability. Since the probit results in Table 6 indicate that lagged group-adjusted return is an important predictor of 5-year survival, it follows that Heckit λ is negatively related to lagged performance. From Table 4, lagged performance is strongly positively related to current performance. Consequently, the negative relation between lagged performance and λ may cause cross-sectional regressions that omit lagged performance to load negatively on λ. Thus, the simple regressions that do not include lagged performance suffer from an omitted variable problem. Our results suggest that such misspecification may be a serious concern when attempting to use the Heckit procedure to control for survivor biases in cross-sectional regressions.

5. Summary and Conclusions

Evidence suggests that funds disappear following poor multi-year performance. Using Carhart's (1997) sample of U.S. mutual funds, we demonstrate both analytically and empirically that this survival rule typically causes the bias in estimates of average annual performance to increase in the sample length. At the same time, our results provide a warning that the nature of the biases imparted can be complicated. In our sample, the bias is economically small at 0.07% for one-year samples, but a significantly larger 1% for samples longer than fifteen years. In tests of mutual fund performance persistence, conditioning on survival weakens the evidence of persistence. This result is again consistent with the evidence of a multi-period survival rule in effect for the U.S. mutual fund industry.

We also explain how the relation between performance and fund characteristics can be affected by the use of a survivor-only sample and show that the magnitudes of the biases in the slope coefficients are large for fund size, expenses, turnover and load fees in our sample. Many areas of finance run cross-sectional regressions with performance as the independent variable. The use of a survivor-only sample may seriously bias such regressions. For example, researchers often relate cross-country differences in equity-market performance to cross-country differences in equity-market characteristics. Our analysis suggests that data unavailability for failed equity markets can have important ramifications for such comparisons, particularly if the characteristics in question are related to survival.

Researchers forced to use survivor-only samples need to consider carefully the likely impact of using such samples on the test statistics of interest. It would seem that finance researchers are often in this position. For example, Goetzmann and Jorion (1999) document how equity market disappearance is conditioned upon a downward drift in performance over time, which suggests that survivor biases are likely to be a problem for empirical studies using international data. Our work suggests that both the nature of the survival rule and the sample period length are likely to be important when attempting to characterize survivorship biases.

APPENDIX A

Proof of Proposition 2:

First, notice that only funds that are at least τ years old at T have a return at time T+1-τ. Therefore,

$$\mu_{T+1-\tau}^{T} = \frac{\sum_{j=\tau}^{J} \hat{w}_{j,T}^{T}\ \mu_{\tau,j-m}}{\sum_{j=\tau}^{J} \hat{w}_{j,T}^{T}} \qquad \textbf{(8)}$$

where J is the age of the oldest funds alive at T. As τ increases, increasingly younger cohorts are omitted from the summation.

Start with a k-period sample ending at time T. Its survivor-biased mean is $\bar{\mu}_k^T$. To see the effect of lengthening the sample period, we could either add a year to the end of the sample period and compare $\bar{\mu}_k^T$ to $\bar{\mu}_{k+1}^{T+1}$, or else add a year to the beginning of the sample period and compare $\bar{\mu}_k^T$ to $\bar{\mu}_{k+1}^T$. With the fund industry in a steady-state, these are equivalent and so for expositional convenience, we consider the case of adding a year to the beginning of the sample period.

Assumptions 2) and 3) allow us to write $\hat{w}_{j,T}^{T}$ in the following way:

$$\begin{aligned} \hat{w}_{j,T}^{T} &= \hat{w}_{1,T}^{T}\left(\frac{1}{1+g}\right)^{j-1}, && j = 1,\ 2,\ldots,m-1, \\ &= \hat{w}_{i,T}^{T}\left(\frac{x}{1+g}\right)^{j-i} && j,\ i \geq m. \end{aligned} \qquad \textbf{(9)}$$

Substituting this expression into **(8)** and exploiting assumption 1) gives the following expressions for the cross-sectional survivorship-biased mean for time T+1-τ:

$$\mu_{T+1-\tau}^{T} \equiv \frac{\hat{w}_{1,T}^{T}\left(\sum_{j=\tau}^{m-1}\frac{1}{(1+g)^{j-1}}\ \mu + \sum_{j=m}^{m-1+\tau}\frac{x^{j-(m-1)}}{(1+g)^{j-1}}\ \mu_{j-(m-1)} + \sum_{j=m+\tau}^{\infty}\frac{x^{j-(m-1)}}{(1+g)^{j-1}}\ \mu_{\tau}\right)}{\hat{w}_{1,T}^{T}\left(\sum_{j=\tau}^{m-1}\frac{1}{(1+g)^{j-1}} + \sum_{j=m}^{m-1+\tau}\frac{x^{j-(m-1)}}{(1+g)^{j-1}} + \sum_{j=m+\tau}^{\infty}\frac{x^{j-(m-1)}}{(1+g)^{j-1}}\right)} \qquad \textbf{(10)}$$

for $\tau = 1, 2, ..., m$-1; and,

$$\mu^T_{T+1-\tau} = \frac{\hat{w}^{\,T}_{\tau-1,T}\left(\sum_{j=1}^{m} \left(\frac{x}{1+g} \right)^j \mu_j + \sum_{j=m+1}^{\infty} \left(\frac{x}{1+g} \right)^j \mu_m \right)}{\hat{w}^{\,T}_{\tau-1,T}\left(\sum_{j=1}^{\infty} \left(\frac{x}{1+g} \right)^j \right)} \qquad \textbf{(11)}$$

for $\tau > m$-1. Under the assumption that the μ_i are increasing in i, it follows from **(10)** that $\mu^T_{T+1-\tau}$ is increasing in τ for $\tau = 1, 2, ..., m$. Moreover, equation **(11)** shows that $\mu^T_{T+1-\tau}$ is constant for $\tau \geq m$. Thus, $\overline{\mu}^T_k$ must be increasing in k for all k.

Q.E.D.

APPENDIX B

Details of the Heckit Probit.

For each year *y*, we collect all funds alive at the end of that year and set SURV equal to zero if the fund disappears in years *y*+1 through *y*+5. Otherwise, SURV is set equal to one. To predict the value of *SURV*, we use variables that describe relative fund size, management pricing, past performance, new money flow, and factor loadings. We include the relative TNA of the fund at the end of *y* since the fixed costs of running a fund suggest a smaller probability of survival as the relative TNA of the fund declines. We also include the fund's group-adjusted return for years y-4 to y to capture the effects of past performance and the same lags of the fund's relative net flow to capture the effects of relative net new investment. The fund's relative net flow in year t is the fund's net flow minus the average net flow for other funds that year. To avoid throwing out funds that disappear within five years of inception, we set both of these to zero for lag L if one of them does not exist for this lag, and set the indicator variable *MISS* (*y*-*L*) to one. Otherwise, we set *MISS* (*y*-*L*) to zero. Management pricing is represented by the fund's net expense ratio in year *y*, and the fund's maximum load fee. Coefficients from the four-factor regression of equation **(1)** are also included as predictive variables since Table 1 suggests that four-factor loadings differ across survivors and nonsurvivors, at least on the SMB and HML factors. The regression is run over the five years ending with the end of *y*-1, and requires at least 30 observations. If there are not 30 observations then the coefficients are set to zero, and the variable *CMISS* (*y*-1), which is otherwise zero, is set to one. The probit results are reported in Table 6.

REFERENCES

Brown, Stephen J, and William N. Goetzmann, 1995, "Performance Persistence," *Journal of Finance*, 50, 679-698.

Brown, Stephen J, and William N. Goetzmann, 1997, "Mutual Fund Styles," *Journal of Financial Economics*,

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.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, 1995, "Survival," *Journal of Finance*, 50, 853-873.

Carhart, Mark M., 1995, "Survivor Bias and Persistence in Mutual Fund Performance," Unpublished Ph.D. dissertation, Graduate School of Business, University of Chicago, Chicago, IL.

Carhart, Mark M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57-82.

Carpenter, Jennifer N., and Anthony W. Lynch, 1999, "Survivorship Bias and Attrition Effects in Measures of Performance Persistence," *Journal of Financial Economics*, 54, 337-374.

Coggin, Daniel and Charles Trzcinka, 2000, "Performance Measurement of Institutional Money Managers", Working paper, Indiana University.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 1996, "Survivorship Bias and Mutual Fund Performance," *Review of Financial Studies*, 9, 1097-1120.

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 James MacBeth, 1973, "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 607-36.

Ferson, Wayne E., and Rudi W. Schadt, 1996, "Measuring Fund Strategy and Performance in Changing Economic Conditions," *Journal of Finance*, 51, 425-462.

FundScope Monthly Investment Company Magazine, March 1962 to February 1976, Los Angeles.

Goetzmann, William N. and Philippe Jorion, 1999, "Global Stock Markets in the Twentieth Century," *Journal of Finance*, 54, 953-980.

Greene, William, 2000, "Econometric Analysis," 4th edition, Prentice Hall.

Grinblatt, Mark, and Sheridan Titman, 1992, "The Persistence of Mutual Fund Performance," *Journal of Finance*, 47, 1977-1984.

Heckman, J., 1976, "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement*, 5, 475-92.

Heckman, J., 1979, "Sample Bias as a Specification Error," *Econometrica*, 47, 153-62.

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.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1996, "The J-Shape of Performance Persistence Given Survivorship Bias," *The Review of Economics and Statistics*.

Investment Dealers' Digest, Spring 1993, *The Investment Dealer's Digest Mutual Fund Directory*.

Ippolito, Richard A., 1989, "Efficiency with Costly Information: A Study of Mutual Fund Performance, 1965-1984," *Quarterly Journal of Economics*, 104, 1-23.

Jegadeesh, Narasimham, and Sheridan Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65-91.

Malkiel, Burton G., 1995, "Returns from Investing in Equity Mutual Funds 1971-1991," *Journal of Finance*, 50, 549-572.

Myers, David, 2001, "Survivorship and Predictability in Pension Fund Returns," Working paper, Lehigh University.

Prabhala, N.R., 1997, "Conditional Methods in Event Studies and an Equilibrium Justification for Standard Event-study Procedures," *Review of Financial Studies,* 10, 1-38.

Sirri, Erik R., and Peter Tufano, 1998, "The Demand for Mutual Fund Services by Individual Investors," *Journal of Finance*, 53, 1589-1622.

United and Babson Mutual Fund Selector, semi-monthly, January 1976 to December 1983, Wellesley, MA.

Wiesenberger, Arthur & Co., 1962-1993, *Investment Companies*, Warren, Gorham & Lamont, Inc., Boston, MA.

ENDNOTES

1. Myers (2001) finds this empirically for pension funds.

2. Myers (1999) and Coggin and Trzcinka (2000) examine survivor biases associated with U.S. pension funds. Myers (1999) finds that end-of-sample conditioning reduces persistence measures.

3. Brown and Goetzmann (1997) document that some funds game their stated objectives to improve their relative performance, so we reconstruct the annual series of stated objectives to remove short-term objective "flips." In our data set, the change in benchmark increases prior-year's group-adjusted performance an average of only 0.61% (t-statistic of 1.63), considerably less than the 9.8% reported by Brown and Goetzmann.

4. End-of-sample conditioning can be thought of as look-ahead conditioning with longer look-ahead periods for earlier reference dates.

5. Since mergers and liquidations need shareholder approval, these reorganizations require at least several months to complete. Thus, missing final returns probably do not differ substantially from the prior observed returns on these funds. The evidence from Elton, Gruber and Blake's (1996) sample supports this conclusion: Marty Gruber, in a personal communication, indicates that the final partial-month return on merged funds does not significantly differ from the average nonsurvivor's return. Of greater concern is the 250 funds without exact termination dates, particularly the 53 without any return data. Since these 53 are likely non-survivors, the lack of any return data imparts a survivorship bias to the measures obtained for the full sample. As a consequence, comparisons of the full sample to the survivor-only sample are likely to understate the effects of survivor-only conditioning for the U.S. mutual

fund industry.

6. By contrast, Elton, Gruber, and Blake (1996) find an attrition rate of only 2.3% in their sample. However, Elton et al. (1996) study only a single cohort of funds, so each year's sample requires funds to have survived some time in the past.

7. We obtain only annual returns on many nonsurvivors. Excluding these funds from our monthly portfolio returns upwardly biases performance estimates. To mitigate this potential bias, we compare the average annual return on all funds to those with only monthly returns. If they differ for any year, we add one-twelfth of this difference equally to all months of that year (using continuously compounded returns.) The difference in mean annual return is typically less than 0.20%.

8. The "group-adjusted" measure employed above exhibits cross-correlations by construction. However, if the sizes of the groups are large enough, these cross-correlations are likely to be small.

9. The analysis in this section continues to hold for negative growth rates. A negative growth rate means new funds arrive more slowly than existing funds disappear. At the extreme, no new funds enter the industry, and the industry growth rate is at a minimum determined by the attrition rates. To avoid violating this lower bound, we assume the growth rate here is non-negative.

10. An appendix is available from the authors that constructs examples in which the sample mean is not increasing in the sample length.

11. We assume the database is compiled one year after the last year of the database which simplifies the categorization of survivors and nonsurvivors. The standard error of the survivor

bias for a sample period of length k is calculated as $\left(\frac{1}{T-n+1}\right)\left(2\sum_{i=1}^{n-1}\left(\frac{i}{n}\right)^2 + T-2(n-1)\right)^{\frac{1}{2}} std(R)$,

where T is the number of years in the database, $n = k$ if $k \le \frac{T+1}{2}$, $n = T$+1-k if $k > \frac{T+1}{2}$, and $std(R)$ is the standard deviation of the survivor bias in annual sample equal-weighted return.

12. Myers (2001) finds this in pension funds, too.

13. This is the bias simulated by Brown, Goetzmann, Ibbotson and Ross (1992).

14. A group-adjusted return can be regarded as the return on a zero-investment portfolio that is long the fund and short the group.

15. A good summary of recent theoretical work addressing incidental truncation can be found in Greene (2000). An important finance application is the event study literature since many firm events are discretionary (see, for example, Prabhala, 1997).

16. Turnover is the minimum of purchases and sales divided by average TNA, while our modified turnover measure adds one half of the absolute value of our flow variable to turnover. Our flow variable is similar to Sirri and Tufano's (1998) flow measure except that it adjusts the numerator for TNA changes due to merger, and it uses average monthly assets instead of beginning assets in the denominator. We use average monthly assets in the denominator so that small, rapidly growing funds are not outliers.

17. For the SUR model, the t-statistic is derived from the Wald statistic for the hypothesis that the summed-up coefficient equals zero. By taking into account cross-regression correlation, the SUR is accounting for autocorrelations in monthly fund return within the year when calculating the Wald statistic.

Table 1
Performance of surviving and nonsurviving mutual funds

Survivors are those funds still operating December 31, 1995 and nonsurvivors are funds disappearing before this date. Cross-sectional average monthly performance is the cross-sectional average of the performance estimates of individual funds based on the complete time series of their returns. Group-adjusted performance is the difference between a fund's return and the average return on all other funds with the same declared fund objective. 4-Factor alpha is the intercept from a time-series regression of a fund's excess returns on the 4-factor model factor-mimicking portfolios over the fund's complete history. The four factors are *RMRF*, *SMB*, and *HML* Fama and French's (1993) market proxy and factor-mimicking portfolios for size and book-to-market equity, and *PR1YR*, a factor-mimicking portfolio for one-year return momentum. The right-hand panel contains 4-factor model estimates for portfolios of funds, with t-statistics in parentheses.

	Number of funds	Cross-sectional average monthly performance		4-Factor model estimates for equal-weighted mutual fund portfolios					
		Group-adjusted	4-Factor alpha	Alpha	Loadings on *RMRF*	Loadings on *SMB*	Loadings on *HML*	Loadings on *PR1YR*	Adjusted R-square
All funds	2,071	-0.03%	-0.14%	-0.15%	0.89	0.33	-0.06	0.09	0.978
				(-4.17)	(105.65)	(25.66)	(-4.36)	(9.09)	
Survivors	1,346	0.10%	-0.03%	-0.07%	0.90	0.29	-0.05	0.09	0.984
				(-2.34)	(117.57)	(26.21)	(-4.62)	(10.87)	
Nonsurvivors	725	-0.26%	-0.34%	-0.33%	0.88	0.37	-0.07	0.09	0.966
				(-7.42)	(74.96)	(21.05)	(-3.11)	(6.30)	
Survivors - all funds		0.13%	0.11%	0.08%	0.00	-0.04	0.01	0.00	0.208
				(6.52)	(0.59)	(-9.14)	(2.85)	(0.05)	
Nonsurvivors by reason									
Merged with another fund	417	-0.19%	-0.29%						
Liquidated	258	-0.45%	-0.54%						
Other, self-selected	14	-0.18%	-0.28%						
Other, not self-selected	36	0.03%	NA						

Table 2
Survivor bias in average performance as a function of the sample period length: Calibration to the U.S. mutual fund industry

For each *m*-year attrition rule, $m \in \{1, 2, 3, 4, 5, 10\}$, and each subperiod of length *k* years in the 34-year history, we compute average performance measures for survivor-only samples using simulated conditional mean returns and attrition rates. We assume that annual fund returns are normally distributed with mean zero and standard deviation 5%. We set the growth rate in the industry equal to 5.5% to match the growth rate in the data. Then, for each given sample period length *k*, we average the performance estimates across all possible subperiods of length *k*. This average is reported for $k = 1, 2, 3, 4, 5, 10$, and 30, while the change in this average is reported for *k* going from 30 to 31, from 31 to 32, from 32 to 33 and from 33 to 34.

Panel A: Mutual fund industry consists entirely of *m*-year-olds at time 1. Scaled cutoff $b/\sqrt{m}$ varies with *m* to maintain a sample attrition rate of 3.5%.

m	Bias (in percent) for sample length of:							Change in bias (in basis points) from increasing sample length from:				$b/\sqrt{m}$	Death rate (in percent)
	1	2	3	4	5	10	30	30 to 31	31 to 32	32 to 33	33 to 34		
1	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.00	0.00	0.00	0.00	-9.06	3.5
2	0.29	0.42	0.45	0.47	0.48	0.50	0.51	0.03	0.03	0.04	0.09	-8.45	3.5
3	0.24	0.35	0.45	0.49	0.51	0.55	0.58	0.06	0.06	0.09	0.19	-7.80	3.5
4	0.22	0.30	0.39	0.48	0.51	0.58	0.63	0.09	0.10	0.15	0.29	-7.18	3.5
5	0.20	0.27	0.34	0.41	0.50	0.60	0.66	0.12	0.14	0.21	0.39	-6.59	3.5
10	0.14	0.18	0.21	0.25	0.29	0.53	0.73	0.35	0.40	0.52	0.91	-4.03	3.5

Table 2 - continued

Panel B: Mutual fund industry consists entirely of m-year-olds at time 1. Scaled cutoff $b/\sqrt{m}$ is fixed and sets the attrition rate for m=1 to 3.5%.													
m	Bias (in percent) for sample length of:							Change in bias (in basis points) from increasing sample length from:				$b/\sqrt{m}$	Death rate (in percent)
	1	2	3	4	5	10	30	30 to 31	31 to 32	32 to 33	33 to 34		
1	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.00	0.00	0.00	0.00	-9.06	3.5
2	0.23	0.34	0.37	0.38	0.39	0.41	0.42	0.02	0.02	0.03	0.07	-9.06	2.7
3	0.16	0.23	0.31	0.33	0.35	0.38	0.40	0.04	0.04	0.06	0.12	-9.06	2.2
4	0.12	0.17	0.22	0.28	0.30	0.34	0.37	0.06	0.06	0.09	0.15	-9.06	1.8
5	0.10	0.14	0.17	0.21	0.25	0.31	0.34	0.07	0.09	0.11	0.17	-9.06	1.5
10	0.05	0.06	0.07	0.09	0.10	0.18	0.25	0.15	0.16	0.18	0.23	-9.06	0.8

Panel C: Mutual fund industry consists entirely of 1-year-olds at time 1. Scaled cutoff $b/\sqrt{m}$ is fixed and sets the attrition rate for m=1 to 3.5%.													
m	Bias (in percent) for sample length of:							Change in bias (in basis points) from increasing sample length from:				$b/\sqrt{m}$	Death rate (in percent)
	1	2	3	4	5	10	30	30 to 31	31 to 32	32 to 33	33 to 34		
1	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.00	0.00	0.00	0.00	-9.06	3.5
2	0.23	0.34	0.37	0.38	0.39	0.41	0.42	0.01	0.00	-0.02	-0.04	-9.06	2.7
3	0.15	0.23	0.30	0.33	0.34	0.38	0.40	0.02	0.01	-0.01	-0.03	-9.06	2.1
4	0.11	0.16	0.21	0.27	0.29	0.34	0.37	0.03	0.01	0.01	-0.02	-9.06	1.6
5	0.09	0.12	0.16	0.20	0.25	0.30	0.34	0.04	0.03	0.02	-0.00	-9.06	1.3
10	0.03	0.05	0.06	0.07	0.09	0.17	0.24	0.09	0.08	0.07	0.05	-9.06	0.6

Table 3
Estimates of survivor bias in average performance as a function of the mutual fund sample period length

Mean annual group-adjusted return estimates from a survivor-biased sample and from a complete sample and the implied survivor bias. The table averages all possible biased and unbiased samples of a given sample period length that might be assembled from our database over the 1962 to 1995 period. Survivor bias is the difference between the mean annual group-adjusted return estimates in the two samples. A fund's group-adjusted return for a month is its total return that month minus the equal-weighted average return of funds with the same objective. The table also reports correlation-adjusted standard errors in the estimate of survivor bias, assuming independent and identically distributed annual returns.

		Mean annual return estimate			
Sample period length (years)	Number of samples	Survivor-biased sample	Unbiased sample	Survivor bias	Standard error
1	34	0.31%	0.24%	0.07%	0.02%
5	30	0.59%	0.21%	0.37%	0.06%
10	25	0.78%	0.12%	0.66%	0.09%
15	20	0.93%	0.08%	0.85%	0.12%
20	15	1.02%	0.08%	0.94%	0.14%
25	10	1.19%	0.12%	0.99%	0.14%
30	5	1.25%	0.21%	1.04%	0.13%
34	1	1.30%	0.24%	1.06%	0.12%

Table 4
The effects of survivorship on persistence tests

Persistence measures for full, end-of-sample conditioned, and look-ahead conditioned samples. Mutual funds are sorted on January 1 each year into decile portfolios based on a lagged performance measure. The performance measures are 1-year return, 5-year return and 4-factor alpha measured over the prior 3 years. The portfolios are equal-weighted monthly so the weights are readjusted whenever a fund disappears. Funds with the highest lagged performance measure comprise decile 1 and funds with the lowest comprise decile 10. The Spearman nonparametric test measures the correlation in rank ordering of post-formation portfolio performance measures. Here, the null hypothesis is that the performance measures are randomly ordered. The HPZ J-shape t-statistic is the t-statistic for the linear term in a quadratic regression of post-formation rank on pre-formation rank. A reliably negative t-statistic is consistent with spurious performance persistence due to survivorship. In Panel A, holding-period returns are total returns. In Panel B, holding-period returns are group-adjusted: from each fund return we subtract the average return of its type (e.g. Aggressive Growth) that month.

Panel A: Returns

	Decile 1-10 spread		Monthly 4-factor model alphas					
Portfolio sorting variable	Mean monthly return	t-stat	Decile 1	Decile 10	Decile 1-10 spread	t-stat	Spearman test p-value	HPZ J-shape t-stat
Full sample								
1-Year returns	0.63%	4.52	-0.13%	-0.37%	0.24%	1.79	0.148	
5-Year returns	0.23%	2.09	-0.10%	-0.34%	0.24%	2.06	0.025	
3-Year 4-factor alpha	0.36%	5.04	-0.01%	-0.36%	0.36%	4.60	0.000	
End-of-sample conditioned sample								
1-Year returns	0.52%	3.93	-0.05%	-0.15%	0.10%	0.84	0.204	1.74
5-Year returns	0.18%	1.85	-0.07%	-0.19%	0.12%	1.15	0.027	1.48
3-Year 4-factor alpha	0.19%	2.66	0.01%	-0.17%	0.18%	2.30	0.000	2.40
Look-ahead conditioned sample								
1-Year returns	0.62%	4.44	-0.14%	-0.36%	0.21%	1.60	0.174	1.76
5-Year returns	0.20%	1.84	-0.11%	-0.29%	0.17%	1.34	0.052	0.38
3-Year 4-factor alpha	0.34%	4.73	0.00%	-0.34%	0.33%	4.07	0.000	1.80

Table 4 - continued

Panel B: Group-adjusted returns							
	Decile 1-10 spread		Monthly 4-factor model alphas for group-adjusted returns				
Portfolio sorting variable	Mean monthly group-adjusted return	t-stat	Decile 1	Decile 10	Decile 1-10 spread	t-stat	Spearman test p-value
Full sample							
1-Year returns	0.66%	6.27	0.07%	-0.29%	0.36%	3.71	0.002
5-Year returns	0.29%	3.95	0.07%	-0.20%	0.27%	3.69	0.000
3-Year 4-factor alpha	0.31%	5.23	0.17%	-0.20%	0.37%	5.90	0.000
End-of-sample conditioned sample							
1-Year returns	0.48%	5.50	0.14%	-0.04%	0.18%	2.38	0.077
5-Year returns	0.21%	3.34	0.12%	-0.03%	0.15%	2.36	0.000
3-Year 4-factor alpha	0.20%	3.96	0.20%	-0.04%	0.24%	4.54	0.037
Look-ahead conditioned sample							
1-Year returns	0.60%	5.89	0.07%	-0.24%	0.31%	3.33	0.003
5-Year returns	0.22%	2.81	0.10%	-0.06%	0.16%	2.15	0.000
3-Year 4-factor alpha	0.28%	4.39	0.19%	-0.15%	0.34%	5.21	0.002

Table 5
The effects of survivorship on cross-section regressions

Simple regressions: For each of five explanatory variables we run three pooled cross-sectional, time-series regressions in which the dependent variable is a fund's group-adjusted return in year $y+1$, with $y+1$ ranging from 1966 though 1991. Group-adjusted return is return minus the average return of funds with the same objective. The first regression uses the "full" sample and is a Seemingly Unrelated Regression. The "full" sample includes any non-missing return for which we also observe the value of the indicated explanatory variable together with the fund's total net assets (TNA), expense ratio, and sales loads, if any, in year y. The independent variable is listed in the first column, and the dependent variables are the twelve separate monthly group-adjusted returns in year $y+1$, excluding months with missing returns. The values reported are the sums of the twelve slope coefficients, and the t-statistic derived from the Wald test that the slopes sum to zero is reported below. The second and third regressions use the "5-year look-ahead" sample and use the fund's annual $y+1$ return as the dependent variable. The "5-year look-ahead" sample includes a fund's year-$y+1$ return if the fund did not disappear in years $y+1$ through $y+5$ and if its total net assets, expense ratio, and sales loads in year y are available. The second regression is OLS. The third regression is the same as the second except we include the λ-function from the two-stage Heckman-correction procedure, Heckit, using the second probit model of Table 5. This probit model predicts disappearance in years $y+1$ through $y+5$ using group-adjusted return and relative flow in years $y-4$ through y, multiple regression coefficients on the four factors of Carhart (1997), and expense ratio, sales loads, and relative TNA in y. In years when lagged return, lagged flow, or Carhart factor coefficients are unavailable, these variables are set to zero and a dummy variable indicates the data are missing. The coefficient on λ is reported in the subsequent column. The three regressions use the following explanatory variables: the expense ratio in year $y+1$, the relative turnover in year $y+1$, the relative TNA at the end of year y, the maximum load in year y, and the group-adjusted return in year y. Turnover is the minimum of purchases and sales divided by average TNA and relative turnover is a fund's modified turnover over the average modified turnover of all funds that year. Relative TNA is the fund's TNA divided by the average TNA of all funds that year. The maximum load is the sum of maximum front-end, back-end and deferred sales charges. The t-statistics are below the coefficient estimates.
Multiple regressions: Each of three sets of simple regressions described above is arranged into a multiple regression, with the five explanatory variables entering independently. The Selected Heckit regression also includes the λ-function from the probit; its coefficient is at the bottom. The t-statistics are below the coefficient estimates.

	Simple regressions				**Multiple regressions**		
	Full	5-year look-ahead			Full	5-year look-ahead	
Independent variables	SUR slope	OLS slope	Heckit slope	Heckit λ	SUR slope	OLS slope	Heckit slope
Net expense ratio (y+1)	-1.142	-0.850	-0.715	-0.012	-1.150	-0.440	-0.708
	-11.41	-5.95	-4.52	-2.01	-8.13	-1.75	-2.68
Relative turnover (y+1)/100	0.036	0.436	0.469	-0.013	0.110	0.407	0.388
	0.47	3.20	3.41	-1.80	1.24	2.91	2.78
Relative TNA (y)/100	0.046	-0.003	-0.129	-0.035	-0.002	-0.039	0.015
	1.92	-0.07	-3.42	-5.88	-0.10	-1.02	0.37
Maximum load (y)/100	0.020	0.011	0.017	-0.027	-0.034	-0.020	-0.044
	1.03	0.42	0.65	-4.92	-1.64	-0.72	-1.50
Group-adjusted return(y)	0.169	0.170	0.170	0.001	0.147	0.150	0.165
	20.44	15.60	14.79	0.15	16.48	12.27	12.62
Multiple regression Heckit λ							0.029
							3.27

Table 6
Probit model of fund survival: Survival through the next five years

Estimates of a probit model that predicts survival in years *y*+1 through *y*+5, for *y* from 1965 to 1990. A positive coefficient on a variable indicates that the probability of survival goes up as that variable goes up. The probit employs the following predictive variables: relative total net assets (TNA) at the end of year *y* , group-adjusted returns in years *y* through *y*-4, relative net flow in years *y* through *y*-4, net expense ratio in year *y*, maximum load in year *y*, the coefficients from regressing fund returns on the four factors of Carhart (1997) over the five years ending with *y*-1, and the dummy variables *MISS* (*y*), ..., *MISS* (*y*-4), and *CMISS* (*y-1*). A fund's relative TNA at the end of *t* is its TNA divided by the average TNA of all other funds on that date. Its group-adjusted return in year *t* is its year-*t* return minus the return of other funds with the same stated objective. Its relative net flow in year *t* is net new investment in the fund minus the average net new investment in other funds that year, while its maximum load is the sum of the maximum front-end, back-end and deferred sales charges in *t*. If there is insufficient data in year *t* to calculate either group-adjusted return or relative net flow then both variables are set to zero and *MISS* (*t*) is set to one. Otherwise, *MISS* (*t*) is zero. We require 30 observations for the Carhart four-factor regression and if there are not enough observations then the coefficients are set to zero and *CMISS* (*y*-1), which is otherwise zero, is set to one. The number of observations in the probit is 10704. The only requirements for a fund's year *y* to be included in the probit are that relative total net assets at the end of *y,* net expense ratio in year y and maximum load in year *y* be available for that fund. P-values for significant difference from zero are to the right.

Independent variable	Coefficient	p-value
Intercept	1.113	<0.0001
Relative TNA (*y*)	0.440	<0.0001
Group-adjusted return (*y*)	1.800	<0.0001
Group-adjusted return (*y*-1)	1.271	<0.0001
Group-adjusted return (*y*-2)	1.044	<0.0001
Group-adjusted return (*y*-3)	0.713	0.0007
Group-adjusted return (*y*-4)	0.697	0.0017
Relative net flow (*y*)	0.389	<0.0001
Relative net flow (*y*-1)	0.293	0.0048
Relative net flow (*y*-2)	0.212	0.0015
Relative net flow (*y*-3)	-0.005	0.8094
Relative Net Flow (*y*-4)	0.032	0.6212
MISS (*y*)	-0.359	<0.0001
MISS (*y*-1)	-0.103	0.0749
MISS (*y*-2)	-0.212	<0.0001
MISS (*y*-3)	-0.351	<0.0001
MISS (*y*-4)	-0.074	0.1590
Net expense ratio (*y*)	-4.819	0.0009
Maximum load (*y*)	-0.029	<0.0001
Coefficient on *RMRF*: *b*(*y*-1)	0.243	0.0089
Coefficient on *SMB*: *s*(y-1)	-0.103	0.0686
Coefficient on *HML*: *h*(*y*-1)	0.182	0.0021
Coefficient on *PR1YR*: *p*(*y*-1)	-0.132	0.2341
CMISS (*y*-1)	-0.269	0.0024

FIGURE LEGEND

Figure 1
Survivor bias as a function of the sample period length
The figure plots the bias in average annual return estimates, introduced by conditioning on fund survival to the end of the sample period, as a function of the length of the sample period. The bias is the average over all possible sample periods of a given length that might be assembled from our database over the 1962-1995 period. The dotted lines represent two-standard error boundaries in the average bias.

Figure 1. Survivor Bias in Average Performance as a Function of the Sample Time Length