

First draft: November 30, 1994  
This draft: September 12, 2000

## MUTUAL FUND SURVIVORSHIP

Mark M. Carhart\*, Jennifer N. Carpenter\*\*, Anthony W. Lynch<sup>×</sup> and David K. Musto<sup>+</sup>

- \* 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 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, 24<sup>th</sup> Floor, New York, NY 10005. This paper was previously entitled, “Survivor Bias and Mutual Fund Performance.”

# MUTUAL FUND SURVIVORSHIP

## ABSTRACT

This paper offers a comprehensive study of survivorship issues, in the context of mutual fund research, using the mutual fund data set of Carhart (1997). We find that funds in our sample disappear primarily because of multi-year poor performance. Then we demonstrate analytically that this survival rule typically causes the survivor bias in average performance to increase in the length of the sample period, though it is possible to construct counterexamples. In the data, we find a strong positive relation between the survivor bias in average performance and sample period length. The bias is economically small at 17 basis points per annum for one-year samples, but a significantly larger one percent per annum for samples longer than fifteen years. We also find evidence of performance persistence in our sample and, consistent with the presence of a multi-period survival rule, we find that the persistence is weakened by survivorship bias. Finally, we 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. Because survivorship issues are relevant for many data sets used in finance, the analysis in this paper has potential applications in areas of financial economics beyond just mutual fund research.

Survivorship bias affects almost every mutual fund study. Most commercially available mutual fund data sets include only funds currently in operation, and many commonly used research methodologies impose additional selection biases. With the exception of a few recent papers, however, researchers frequently ignore selection biases altogether or argue that their effect is insignificant. This attitude is unfortunate, as selection-bias issues pervade almost all empirical studies of panel and time-series data sets.

This paper offers a comprehensive study of survivorship issues, in the context of mutual fund research. We examine how survivorship bias affects mutual fund studies both theoretically and empirically. We study the effect of survivorship on three types of mutual fund studies: (1) estimates of average performance, (2) tests of performance persistence, and (3) cross-section estimates of the relation between performance and fund attributes. The analysis divides survivorship bias into the separate but related issues of survivor bias and look-ahead bias, an important distinction rarely acknowledged in the literature. Our results indicate that survivorship bias substantially alters the inferences from mutual fund studies, but that the effects vary across test type, form of survivorship bias, and sample period length. Because survivorship issues are relevant for many data sets used in finance, the analysis in this paper has potential applications in areas of financial economics beyond just mutual fund research.

A number of recent papers have addressed issues in mutual fund survivorship. Brown, Goetzmann, Ibbotson and Ross (1992) and Carpenter and Lynch (1999) study the effects of survivorship bias on tests of performance persistence using simulation and calibration. Grinblatt and Titman (1989) and Wermers (1997) study the effect of survivorship bias on a database of underlying stock holdings. Malkiel (1995) estimates the effects of survivorship bias in Lipper Analytical

Service's database. Brown and Goetzmann (1995) estimate survivor biases in their ten-year sample of mutual fund returns and find that nonsurvivors underperform the average fund in each of their last three years. Finally, Elton, Gruber and Blake (1996) study survivorship issues in the cohort of larger funds listed in the 1977 issue of Wiesenberger's *Investment Companies*.

Our study contributes to the existing mutual fund literature in several ways. First, it carefully distinguishes between survivor bias and look-ahead bias, particularly in the context of persistence tests where the distinction is most applicable. Second, the paper carefully analyzes how the survival rule affects the average performance bias in survivor-only samples. Third, the paper considers the impact of survivor bias on cross-sectional regressions of performance on fund characteristics, providing a framework for assessing the direction and magnitude of any biases in the coefficients. Finally, the paper measures the effects of survivorship bias in common mutual fund tests using the data set of Carhart (1997). This data set is one of the most complete mutual fund data sets available and is virtually free of survivorship bias.

Our results suggest that nonsurvivors in the U.S. mutual fund industry disappear primarily because of multi-year underperformance. A probit analysis confirms the predictive ability of lagged performance, even in the presence of the most recent year's performance. We show that a survival criterion based on multi-year performance (a multi-period survival rule) typically causes survivor-biased estimates of average performance to increase in the time-length of the sample, but at an ever decreasing rate. In our sample, we measure the bias in annual return at 17 basis points for one-year samples, 43 basis points for five-year samples, and approximately one percent for data sets longer than fifteen years. At the same time, it is possible to construct examples in which the bias in average performance is not increasing as a function of the sample period length. Our analysis provides a

warning to researchers that a multi-period survival rule can have unexpected consequences for estimates of average performance in a survivor-only sample.

We next examine empirically the impact of survivor bias on persistence tests, and find that the bias attenuates performance persistence relative to the full sample. The downward bias in the persistence measures induced by using both look-ahead biased and survivor biased samples is understandable given the multi-period nature of the 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 empirically that a multi-period survival rule can attenuate persistence in these samples despite any heterogeneity in performance volatility across funds in the sample. With persistence in the full sample, the attenuation in the persistence is found to be much larger for the survivor-only sample than the look-ahead biased sample.

Finally, we explain how survivor-only conditioning can affect the cross-sectional relations obtained between fund performance and fund characteristics. In particular, for the cross-sectional slope coefficient to be biased in the survivor-only sample, the fund characteristic in question must affect the survivor bias in performance. In fact, the direction and magnitude of the characteristic's impact on the performance bias determine the direction and magnitude of the slope coefficient bias. We then estimate the slope coefficient biases for commonly-used fund characteristics when using both the survivor-only sample of U.S. mutual funds and the full sample. We find that the magnitude of these biases can be large and their directions can be explained using the intuition we develop.

Section 1 presents definitions while Section 2 describes the U.S. mutual fund data set. Section 3 characterizes survivors relative to non-survivors and quantifies the survival rule. Section 4 considers the impact of using a survivor-only sample on average performance measures, both

theoretically and empirically. Section 5 examines the bias in persistence measures induced by using survivor-only or look-ahead biased samples. Finally, section 6 examines the impact of survivor bias on cross-sectional regressions, and section 7 concludes.

## 1. Definitions

### A. *Selection bias definitions*

To mitigate potential confusion, we define some important terms used in this study. *Survival rules* refer to the criteria which cause funds to disappear from the data set. A one-period survival rule means that only funds with current period returns greater than some threshold are observed at the end of the period. A multiple-period survival rule means that funds appear in the data set only if their past  $n$ -period return exceeds some threshold.

Fund disappearance, or attrition, can lead to two distinct but related problems, *survivor bias* and *look-ahead bias*. *Survivor bias* is the effect of including in the sample only the funds extant at the end of the sample period. *Look-ahead bias* is the effect of requiring funds to survive some minimum length of time by trimming funds that perish during a look-ahead period. The survivor-biased sample trims not only these funds but also funds that perish between the end of the particular look-ahead period and the end of the sample period. Consequently, the survivor-biased sample can be thought of as imposing a look-ahead bias whose look-ahead period is larger for ranking periods that occur earlier in the sample. Survivor bias is solely a property of a data set, whereas look-ahead bias usually results from a test methodology imposing a survival condition. The distinction between these two biases is not always acknowledged in the literature. Some studies consider data sets free of survivor bias but then impose look-ahead-biased methodologies.

An example of a survivor-biased sample is Morningstar OnDisc, which reports performance since January 1976 only for funds still existing at the end of the sample period. In principle, correcting for survivor bias is simply an issue of data collection, although in practice the missing data are often not completely obtainable.

In contrast, the common performance persistence test methodology of regressing future  $n$ -period performance on a measure of past performance suffers from an  $n$ -period look-ahead bias, since the test conditions on survival for another  $n$  periods beyond the evaluation date. Some degree of look-ahead bias is inherent in any test of performance persistence that requires a balanced future and past performance sample.<sup>1</sup> An important issue is how the effects of look-ahead bias vary with the nature of the survival rule, particularly for persistence tests, and the theory sections provide some discussion of this issue. However, the empirical work, by construction, can only characterize the look-ahead bias induced by the survival rule actually in effect in the U.S. mutual fund industry. Mitigation of look-ahead bias requires minimizing the *look-ahead period*, the time period over which future performance is measured. The methodology we use in the persistence tests below requires looking forward only one month.

#### *B. Averaging 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. A third approach, frequently employed in the mutual fund literature, calculates statistics cross-sectionally for each time period and then averages these estimates through time. We rely primarily on this third approach for aggregation.

### *C. Performance measurement*

We employ two methods of performance measurement. The first method simply subtracts from fund returns the equal-weight average return on all funds with the same objective in that period. We call this the “group-adjusted” performance measure. When funds change objectives, they move to a new group-adjusted measure. 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 61 basis points (t-statistic of 1.63), considerably less than the 9.8% reported by Brown and Goetzmann.

The second performance measure is the time-series regression intercept from asset pricing models, commonly called “alphas” after Jensen’s (1968) work. We use two such models: the Capital Asset Pricing Model (CAPM) derived by Sharpe (1964) and Lintner (1965), and the 4-factor model of Carhart (1997). For the CAPM, we use Fama and French’s (1993) market proxy, updated to 1995. The 4-factor model uses Fama and French’s (1993) 3-factor model plus an additional factor capturing Jegadeesh and Titman’s (1993) one-year momentum anomaly. The model is

$$r_{it} = \alpha_{iT} + b_{iT}RMRF_t + s_{iT}SMB_t + h_{iT}HML_t + p_{iT}PRIYR_t + e_{it} \quad (1)$$

where  $r_{it}$  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 PRIYR are returns on value-



weighted, zero-investment, factor-mimicking portfolios for size, book-to-market equity, and one-year momentum in stock returns. Carhart (1997) describes the 4-factor model in greater detail and finds it prices passively-managed portfolios formed on size, book-to-market equity and one-year return momentum considerably better than the CAPM or Fama and French's (1993) 3-factor model. We use the 4-factor model in addition to the CAPM in an effort to adjust fund performance for well-known regularities in stock returns. Finally, Carhart (1995) finds that dynamic performance measurement models like Ferson and Schadt (1994) do not substantially alter his performance estimates.

## 2. Data

Our database covers all known diversified equity mutual funds monthly from January 1962 to December 1995 and excludes sector funds, international funds, and balanced funds. Moreover, the data are virtually free of survivor bias. We obtain data on surviving funds and for funds that disappear after 1989 from Micropal/Investment Company Data, Inc. (ICDI.) For all other nonsurviving funds, the data are collected from *FundScope Magazine*, *United and Babson Reports*, *Wiesenberger's Investment Companies*, the *Wall Street Journal*, and ICDI's past printed reports. We partition the sample into three primary investment objectives using Wiesenberger and ICDI classifications: aggressive growth, growth and income, and long-term growth. All funds in the sample start as general equity funds in one of these three objectives. Funds frequently change objectives during the sample but we never drop a fund once in the sample.

The data set includes monthly returns and annual attributes. Return series are as complete as can practically be obtained, but 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 funds from ICDI, Wiesenberger *Investment Companies, FundScope Magazine* and *Investment Dealer's Digest*. 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.

While our sample does not include a number of nonsurviving fund returns, the bias induced by the last few omitted returns is probably quite small. Since mergers and liquidations need shareholder approval, these reorganizations require at least several months to complete, and probably closer to four to six months. 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.

The sample differs from Carhart (1997) in two primary ways. First, we deal with multiple

share class funds differently. Multiple share class funds divide a common pool of assets into share classes with differing distribution costs. Whereas Carhart (1997) treats each share class as a separate fund, the sample in this paper includes only the original share class for each fund. For fund size, we use the sum of the total net assets over all share classes. This treatment of multiple share class funds eliminates 62 funds from Carhart (1997). Second, we extend the data set two years to 1995 and remove several duplicate and improperly categorized funds from Carhart (1997). This adds 194 new funds and removes 28 funds.

Table 1 reports annual summary statistics on the data set as well as time-series averages over the complete period. Our sample includes a total of 2,071 diversified equity funds, 1,346 of them still operating as of December 31, 1995. In an average year, the sample includes 545 funds with average total net assets (TNA) of \$179.5 million and average expenses of 1.19 percent per year.

To measure net additions and withdrawals, we also measure *Flow* as:

$$Flow_{it} = \frac{TNA_{it} - (1 + R_{it})TNA_{it-1} - MGTNA_{it}}{Avg\ Monthly(TNA_{it-1}, TNA_{it})}, \quad (2)$$

where  $MGTNA_{it}$  is the increase in fund  $i$ 's assets in period  $t$  due to merger and the denominator is the average monthly total net assets of fund  $i$  from  $t-1$  to  $t$ . *Flow* is similar to Sirri and Tufano's (1992) flow measure except that it adjusts for TNA changes due to merger, and it uses average monthly assets instead of beginning assets. On average, the typical fund receives net inflows of 7.0 percent per year as measured by *Flow*.

In addition, funds trade 82.5 percent of the value of their assets (*Mturn*) in an average year. Since reported turnover is the minimum of purchases and sales over average TNA, we obtain *Mturn* by adding to reported turnover one-half of the absolute value of *Flow*. Also, over the full sample, average maximum load fees are 7.05 percent, and 59.8 percent of funds charge them in a given year.

Maximum load is the total of the maximum initial, rear and deferred sales charges, as a percentage of assets invested.

In an average year, we find that 3.6 percent of funds disappear. Of this total, 2.2 percent per year disappear due to merger and 1.0 percent disappear because of liquidation. By contrast, Elton, Gruber, and Blake (1996) find an attrition rate of only 2.3 percent 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. In the subsamples grouped by investment objective, aggressive growth funds perish at an annual rate of 4.5 percent, which is statistically significantly larger than 2.9 percent for long term growth and 3.3 percent for growth and income funds. In addition, unlike Elton et al. (1996), we find that the annual attrition rate is significantly negatively related to the previous year's market return, with a t-statistic of -2.30.

The annual summary statistics indicate substantial variation in mutual fund properties through time. For example, the rate at which assets enter and leave the industry varies, with alternating periods of high growth/low disappearance rates and low growth/high disappearance rates. In addition, the nominal size of funds, TNA, and average expense ratio mostly increase over the 34-year period, while both the load fees and the proportion of funds charging load fees decrease.

Table 1 also demonstrates that the equity mutual funds in our sample earned reported returns approximately 0.6 percent per year below the value-weighted CRSP index, occasionally under or over performing the CRSP index by as much as 9 percent per year. Reported returns are net of all operating expenses (expense ratios) and security-level transactions costs, but do not include sales charges. Perhaps more surprising, funds only hold 83.2 percent of their portfolios in common stocks in an average year. In the remainder of their portfolios, funds hold 10.2 percent in cash and 6.6

percent in preferred stocks and bonds.

While our sample is probably the most complete survivor-bias-free mutual fund database available, Brown and Goetzmann (1995), Elton, Gruber and Blake (1996), Grinblatt and Titman (1989), Malkiel (1995), and Wermers (1996) study related mutual fund databases. Elton et al. 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. By contrast, our sample includes all funds between 1962 and 1995, adding new funds as they appear.

Grinblatt and Titman (1989) and Wermers (1996) use quarterly data on the mutual funds' underlying stock holdings since 1975 from CDA/Spectrum to estimate returns gross of transactions costs and expense ratios. Wermers' data set therefore permits a more detailed analysis of investment strategies and gross investment performance than ours. However, the CDA data do not permit return calculations on nonsurvivors in their final periods before disappearance. The data set studied by Brown and Goetzmann (1995) is similar to ours, except that it covers only the period from 1977 to 1988 and uses annual returns estimated from Wiesenberger's *Investment Companies*. Finally, Malkiel's (1995) data set uses quarterly returns from 1971 to 1991, obtained from Lipper Analytical Services.

### **3. Characterizing Attrition in the Data**

This section examines the properties of surviving and nonsurviving mutual funds and gives evidence on the cause of fund disappearance.

#### *A. Properties of Surviving and Nonsurviving Mutual Funds*

Table 2 compares the performance, size, expense ratios, and turnover of surviving and

nonsurviving mutual funds. Not surprisingly, nonsurviving funds exhibit considerably poorer performance than surviving funds. After estimating the group-adjusted and 4-factor model performance on individual funds over their complete return series, we calculate the cross-sectional average of these estimates for survivors and nonsurvivors. By these measures, nonsurviving funds underperform survivors by 31 to 36 basis points per month, or about 4 percent per year.

Also not unexpectedly, nonsurviving funds are smaller and have higher expense ratios and turnover than surviving funds. Table 2 contains measures of relative size, expenses, turnover, and money flow for various groups of funds. We measure relative size, Relative TNA, for a given group as follows. For each year in the sample, we first compute the ratio of each fund's TNA to the average TNA for the entire sample in that year, and then we calculate the group average of this ratio for that year. Then, we take the time series average of these annual group averages to obtain the Relative TNA measure for the group. Relative expense ratio and turnover (Mturn) are measured analogously. To measure relative flow, we use the difference in Flow instead of the ratio, again taking the time-series average of annual cross-sectional averages. By these measures, surviving funds are approximately 45 percent larger than the average fund and growing faster by 1.2 percent per year, while nonsurviving funds are less than one-third the size of the average fund and shrinking. Similarly, surviving funds have expense ratios about 11 percent lower than average, while nonsurvivors charge expenses 23 percent above average. Nonsurviving funds also trade about 15 percent more than the average fund, while survivors trade about 4 percent less. These results are consistent with those of Brown and Goetzmann (1995) and Malkiel (1995), who also document the higher expenses and smaller size of nonsurvivors.

We subdivide the defunct mutual fund sample by reason for disappearance into four broad

categories: (1) mergers, (2) liquidations, (3) other self-selected reasons, and (4) not self-selected or unknown reasons. Table 2 shows that about 58 percent of all defunct funds disappear because of merger and 36 percent disappear due to liquidation. A further two percent vanish through other self-selected means, usually at the fund manager's request for removal.

Approximately five percent of nonsurviving funds depart for unknown reasons or are dropped from the sample by the database manager, not the fund itself. Sixteen of these are tax-free exchange (TFE) funds. TFEs permitted a tax-free exchange of an investor's stock portfolio for shares in the fund, allowing investors to defer capital gains recognition.<sup>2</sup> Congress withdrew this tax loophole in 1967 and these funds disappeared from our sources in the same year. Five funds are dropped from the sample because they are variable annuity investment vehicles, and the reason for disappearance is unknown for fifteen funds.

While all nonsurviving fund groups underperform, liquidated funds exhibit the worst relative performance and the smallest size and highest expense ratios. Liquidated funds are only five percent of the average fund's size and have expenses and turnover 85 percent and 53 percent higher than the average fund, respectively. In contrast, funds that subsequently merge earn performance similar to the average nonsurvivor. Not surprisingly, merged funds are larger and have lower expense ratios than the typical nonsurviving fund.

Funds disappearing for reasons other than merger or liquidation mostly underperform as well. However, the performance on split, variable annuity and tax-free exchange funds are not abnormally negative. Performance is significantly negative for funds voluntarily removing themselves from the sample, funds reorganizing to have closed-end status, and funds disappearing for unknown reasons. These findings are not surprising. Since Sirri and Tufano (1992) and others show that investors

respond to past performance, poorly-performing funds may stem the tide in negative flows by changing to closed-end or removing themselves from commercial mutual-fund-ranking services.

In Table 3, we measure factor loadings and performance on equal-weighted portfolios of mutual funds, as in Carhart (1997). The portfolios that include nonsurviving funds keep them in the equal-weighted average until they disappear and then readjust the portfolio weights appropriately.<sup>3</sup> This procedure mitigates look-ahead bias. Table 3 shows that the equal-weighted portfolio of all mutual funds underperforms by five basis points per month relative to the CAPM and 15 basis points relative to the 4-factor model. The 4-factor model estimate amounts to a sizeable underperformance of 1.8 percent per year. As in Carhart (1997), the significant difference in performance estimates between the CAPM and 4-factor model is due to mutual funds holding smaller, lower book-to-market and higher momentum stocks which increases their average return over the sample period by a net of 10 basis points per month.

The performance of the portfolios of survivors and nonsurvivors is considerably different. Survivors achieve abnormal performance of +3 basis point per month relative to the CAPM, and -7 basis points relative to the 4-factor model. Nonsurvivors, however, earn CAPM and 4-factor model performance measures of -24 and -33 basis points per month, respectively. In our sample, the overall survivor bias in average performance does not depend on the performance measurement model. The difference between estimates of performance using survivors only and estimates using the complete sample are 8 basis points per month, using either the simple return, CAPM, or 4-factor model estimates of survivor bias over the complete time period. 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.

The survivor bias does differ significantly across fund objective groupings. In annual returns, aggressive growth survivors outperform all aggressive growth funds by 1.9 percent per year. For growth and income and long-term growth funds, the biases are 0.4 and 1.1 percent, respectively. In summary, omitting nonsurvivors from estimates of average performance downwardly biases factor-model-adjusted mutual fund performance by approximately one percent per year. This applies only to the complete sample period; in the Section 4, we measure survivor bias in annual return estimates for other sample period lengths.

#### *B. Evidence on the Survival rule*

We now address the question of whether funds disappear primarily because of a single poor return, a single-period survival rule, or because of a sequence of poor returns, a multi-period survival rule.

##### *B. 1. Performance Prior to Disappearance for Non-survivors*

This subsection examines the relative performance of nonsurviving funds in their final five years of existence. Figure 1 suggests that nonsurvivors underperform throughout their last five years of existence, but especially in their final year. The figure presents the annual group-adjusted performance on an equal-weight portfolio of nonsurviving funds in each of their last five years.<sup>4</sup> This performance is gross of expense ratios in order to remove the effect of declining fund size on performance. The figure suggests that most nonsurvivors disappear after underperforming for multiple years, and perhaps also that some funds disappear after only one particularly poor final year return. However, the portfolio average does not directly reveal the distribution of individual fund

performance.

The evidence in Table 4 shows that multiple-period performance dominates the selection process. The table reports the proportion of all nonsurviving funds with group-adjusted performance below various performance fractiles of all funds. In their final twelve months, 62 percent of nonsurvivors report performance below the median, and 24.8 percent report performance in the bottom decile of all funds. Similarly, over their last five years, almost 80 percent are below the median, 33 percent in the bottom decile, and 21 percent fall in the bottom 5 percent. In addition to the large proportion of individual funds that underperform over their final five years, the relative performance of nonsurviving funds worsens as the performance measurement periods lengthens. This indicates that most funds vanish after underperforming for several years; if funds disappeared after only a single poor return, the relative performance of nonsurviving funds would increase as the performance measurement period lengthened. However, there is also evidence that funds sometimes disappear after only one poor return. Relatively more funds appear in the bottom one percent performance fractile for their last year than their last two to five years.

### *B. 2. Probit Analysis of Disappearance Predictors for Non-survivors*

To establish the statistical significance of lagged returns as predictors of fund disappearance, we fit a probit model similar to that of Brown and Goetzmann (1995). For each year  $y$  from 1965 to 1994, we collect all funds alive at the end of  $y$ , and set the variable  $DIE$  to zero if the fund disappears before the end of  $y+1$ , and one otherwise. To predict the value of  $DIE$ , we use variables that describe relative fund size, past performance, and new money flow. We calculate the relative size of the fund at the end of  $y$ ,  $E\_TNA_y$ , as the log of the fund's size divided by the average size of

funds of its type at the end of  $y$ . We include the fund's group-adjusted return for years  $y-4$  to  $y$ ,  $E\_RET_{y-4}$  to  $E\_RET_y$ , to capture the effects of past performance. Finally, the variables  $E\_NM_{y-4}$  to  $E\_NM_y$  are included to capture the effects of the relative new money invested in the fund in each of the last five years. For a given year, this variable is defined to be a fund's new money divided by the average new money for funds of its type that year. Consistent with Brown and Goetzmann (1995), new money is calculated as the percentage change in a fund's total net asset value, net of fund return. To avoid throwing out funds that disappear within five years of inception, and consequently do not have data for some of the lags, we set  $E\_RET_{y-L}$  and  $E\_NM_{y-L}$  to zero if one of them does not exist for lag  $L$ , and set the indicator variable  $MISS_{y-L}$  to one. Otherwise, we set  $MISS_{y-L}$  to zero. So there are sixteen explanatory variables in all: relative size, five lags of relative returns, five lags of relative new money, and five dummies.

Results for an unconstrained version of the model are reported in Table 5. Negative coefficients indicate that the probability of disappearance decreases as that variable increases. The results are broadly consistent with those of Brown and Goetzmann (1995). The likelihood of disappearing goes up as size goes down, as recent and lagged relative returns go down, and as relative new money goes down, though this last relation is generally not statistically significant.

We test three hypotheses using log likelihood ratio tests and report the results in the last three rows of Table 5. The first test is whether the lagged returns matter at all, that is, whether the coefficients on  $E\_RET_{y-4}$  through  $E\_RET_{y-1}$  are all equal to zero. The log likelihood ratio has a p-value less than 0.001 so the null that the lagged returns do not enter is rejected. The second test is whether the lagged returns enter identically, that is, whether the coefficients on  $E\_RET_{y-4}$  through  $E\_RET_{y-1}$  are all equal. The resulting test statistic is 5.06, which is below the 90% rejection level,

so the null that the lagged returns enter identically is not rejected. The third test is whether the coefficient on  $E\_RET_y$  is the same as the average coefficient on the other four lags. The test statistic of 4.19 exceeds the 95% cutoff of 3.84 for one degree of freedom, so the null is rejected.

Taken together, the results of the hypothesis tests indicate that a multi-period survival rule is applicable (since the first null is rejected), that the coefficients on the four lagged returns are the same (since the second null is not rejected), and that the coefficient on the year- $y$  return variable is significantly larger in magnitude than the coefficient on the lagged return variables (since the third null is rejected). These results suggest that single period and multi-period survival rules operate simultaneously in the U.S. mutual fund industry.

#### **4. Survivor Bias Effects on Estimates of Average Performance**

This section examines the properties of the bias in estimates of average performance created by eliminating nonsurvivors from mutual fund samples using either a single or multi-period survival rule. Focusing on the time-series average of the sample's cross-section mean performance each period, subsection A shows that the bias induced by a single-period survival rule is invariant to sample length, and then provides compelling intuition for why the bias induced by a multi-period rule is typically increasing with sample length. Section B then presents an example that illustrates why the multi-period rule does not always create a bias that is increasing in the sample length. Section C simulates biases for  $m$ -period survival rules and different sample lengths using sample growth rates and fund attrition rates that match those in the U.S. mutual fund data. The simulation shows that when  $m > 1$ , the survivorship bias grows with the length of the sample period, consistent with the intuition presented in subsection A. Finally, subsection D examines the relation between

the average performance bias and sample period length empirically. We find that the relation is positive, which suggests that a multi-period survival rule operates in the U.S. mutual fund industry.

### *A. Theory*

Every period, each mutual fund in existence generates a performance measure. For convenience, we shall call the periods years, though they could be any length of time, and the performance measures returns, though they could be any measure of performance, such as factor-model-adjusted returns or group-adjusted returns. An  $m$ -year survival rule causes fund disappearance,  $m \in \{1, 2, 3, \dots\}$ . That is, each year, funds at least  $m$  years old disappear through liquidation or merger if the sum of their returns over the preceding  $m$  years falls below a threshold  $b$ . Younger funds do not disappear. In addition, fund returns are cross-sectionally and intertemporally independent and identically distributed with mean  $\mu$ .<sup>5</sup> Let  $g \geq 0$  be the annual growth rate of the number of funds in the mutual fund industry.

Consider a survivor-only sample of funds for a  $k$ -year sample period. This is the set of all funds in existence prior to the end of the sample period that survive the selection process in every year from their date of inception to the end of the sample period. Notice that the sample includes newer funds with fewer than  $k$  periods of performance, who survive until the end of the sample period. Define the survivor-biased estimate of average performance as the time-series average of the yearly equal-weight cross-sectional mean returns of these funds. Notice that, by assumption,  $b$ ,  $g$ ,  $\mu$ , and the variance of fund return are all independent of  $k$ .

***Proposition 1:*** If a single-year survival rule causes fund disappearance (i.e.,  $m=1$ ), the bias in the

average performance estimate for the whole sample 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

$$E[R|R > b] - \mu \quad (3)$$

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

Now suppose a multi-year survival rule determines fund survival (i.e.,  $m > 1$ ). Consider the collection of funds that survive through some time  $T$ . Each of these funds 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:

$$C_t \equiv [ (\sum_{\tau=t-m+1}^t R_\tau) > b ], \quad (4)$$

and let  $\mu_{ij}$  be the one-period return conditioned on a set of  $j+1$  consecutive performance cuts with the last cut occurring  $i-1$  periods after the return:

$$\mu_{ij} \equiv E[R_{t+1-i} | C_{t-j}, \dots, C_t]. \quad (5)$$

When thinking about the conditional mean of a given period's return, we define a cut whose return window includes the given return as a *direct* cut and one whose window does not as an *indirect* cut. The conditional mean in (5) imposes a direct cut only when  $j \geq 0$  and  $j+m > i > 0$ . If this condition does not hold for a given  $i$  and  $j$ ,  $\mu_{ij}$  only involves indirect cuts, and so must be equal to  $\mu$ .<sup>6</sup>

Conditional on surviving through time  $T$ , the mean time- $t$  return of each fund that is  $j$  years old at time  $T$  is  $E[R_t | C_{T-j+m}, C_{T-j+m+1}, \dots, C_T] = \mu_{T+1-t, j-m}$ . These age- $j$  funds have returns from time  $T+1-j$  to  $T$ , and so only funds that are at least  $T-t+1$  years old at  $T$  have a return at time  $t$ . It follows that the cross-sectional mean time  $T+1-\tau$  return of funds that survive through time  $T$  is

$$\mu_{T+1-\tau}^T \equiv \frac{\sum_{j=\tau}^J \hat{w}_{j,T}^T \mu_{\tau,j-m}}{\sum_{j=\tau}^J \hat{w}_{j,T}^T} \quad (6)$$

where  $J$  is the age of the oldest funds alive at  $T$  and  $\hat{w}_{j,T}^T$  is the fraction of time- $T$  survivors (after the time- $T$  cut) that are  $j$  years old at time  $T$ . As  $\tau$  increases, the time period becomes earlier, and increasingly younger cohorts are omitted from the summation. Finally, the survivor-biased estimate of average performance across the  $k$ -period sample ending at  $T$  is

$$\bar{\mu}_k^{-T} \equiv \frac{\sum_{t=T-k+1}^T \mu_t^T}{k}. \quad (7)$$

We are interested in characterizing the behavior of  $\bar{\mu}_k^{-T}$  as a function of  $k$ .

Before we proceed, a few comments are in order. It might seem that indirect cuts should not affect a conditional mean return. In particular, when calculating conditional means for  $R_t$ , it might seem that conditions associated with cuts before time  $t$  and after time  $t+m-1$  can be disregarded. However, in general, this is not the case, as our example in subsection B below demonstrates. 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. To pursue the implications of this idea, we ignore the impact of indirect cuts for now and let  $\mu_\tau$  denote any conditional mean of  $R_t$  which imposes exactly  $\tau$  direct cuts. Note that  $\tau$  can range from 1 to  $m$  since  $m$  is the maximum number of  $m$ -period cuts that can include a given single-period return. We also use  $\mu_\tau$  to denote the set of conditional means ( $\mu_{i,j}$ 's) with exactly  $\tau$  direct cuts. With this notation, the cross-sectional mean time- $(T+1-\tau)$  return of funds that survive through time  $T$  can be written as follows for  $J \geq 2m$ :

$$\begin{aligned}
\mu_{T+1-\tau}^T &\equiv \kappa_{T+1-\tau}^T \left[ \sum_{j=\tau}^{m-1} \hat{w}_{j,T}^T \mu_j + \sum_{j=m}^{m-1+\tau} \hat{w}_{j,T}^T \mu_{j-(m-1)} + \sum_{j=m+\tau}^J \hat{w}_{j,T}^T \mu_{j-m} \right], \quad \tau = 1, 2, \dots, m-1 \\
&\equiv \kappa_{T+1-\tau}^T \left[ \sum_{j=\tau}^{m-1+\tau} \hat{w}_{j,T}^T \mu_{j-(\tau-1)} + \sum_{j=m+\tau}^J \hat{w}_{j,T}^T \mu_{j-m} \right], \quad \tau = m, \dots, J-m \quad (8) \\
&\equiv \kappa_{T+1-\tau}^T \left[ \sum_{j=\tau}^J \hat{w}_{j,T}^T \mu_{j-(\tau-1)} \right], \quad \tau = J-m+1, \dots, J
\end{aligned}$$

where

$$\kappa_{T+1-\tau}^T = \frac{1}{\sum_{j=\tau}^J \hat{w}_{j,T}^T} \quad (9)$$

is a scaling factor to ensure that the weights of the various time- $T$  cohorts born prior to  $T+1-\tau$  sum to one.

Intuition suggests that the conditional mean of  $R_t$  is increasing in the number of direct cuts: in other words,  $\mu_\tau$  is increasing in  $\tau$ . Equation (8) can be used to understand why this intuition implies a survivorship-biased  $k$ -period sample mean  $\bar{\mu}_k^T$  that is increasing in  $k$ . Since the sample mean  $\bar{\mu}_k^T$  is obtained by averaging  $\mu_{T+1-\tau}^T$  terms from  $\tau=1$  to  $\tau=k$ , increasing the sample period length from  $k$  to  $k+1$  involves adding  $\mu_{T+1-(k+1)}^T$  to the set of terms used to calculate the sample mean. Thus, it is enough to explain why  $\mu_{T+1-(k+1)}^T$  is typically larger than the average of  $\mu_{T+1-k}^T \dots \mu_T^T$  for any  $k$ .

The first line of equation (8) suggests that the cross-sectional mean  $\mu_{T+1-\tau}^T$  increases as  $\tau$  increases from 1 to  $m$ . For any given  $\tau \leq m$ ,  $\mu_{T+1-\tau}^T$  is an average of  $\mu, \mu_1, \dots, \mu_\tau$  because when the time period,  $t$ , is  $\tau-1$  periods from time  $T$ , the maximum number of cuts that can include  $R_t$  is  $\tau$ . Since  $\mu_\tau$  is increasing in  $\tau$ , this average increases in  $\tau$ , because the relative weights on the  $\mu, \mu_1, \dots, \mu_\tau$  are unchanged as we go from  $\tau$  to  $\tau+1$ . For  $\tau$  beyond  $m$  but not too close to  $J$ ,  $\mu_{T+1-\tau}^T$  is an average of  $\mu_1, \dots, \mu_m$ . It follows that, for any  $k$  beyond  $m$  but not too close to  $J$ ,  $\mu_{T+1-(k+1)}^T \gg \mu_{T+1-k}^T \gg \dots \gg \mu_T^T$ .



Thus, for any  $k$  not too close to  $J$ ,  $\bar{\mu}_{T+1-(k+1)}^T$  is likely to be greater than  $\bar{\mu}_k^T$  and so increasing the sample length from  $k$  to  $k+1$  increases the sample mean.

The following proposition formalizes this intuition using a set of assumptions guaranteeing that  $\bar{\mu}_k^T$  is increasing in  $k$ . Before introducing the proposition, a little more notation is needed. Let  $w_{j,t}$  be the fraction of funds with a time- $t$  return that are age  $j$  at time  $t$ . Let  $x_i$  denote survival probabilities conditional on survival in previous periods:  $x_i \equiv \text{pr} \{C_t \mid C_{t-1}, \dots, C_{t-i}\}$  for  $i > 0$ ,  $x_0 \equiv \text{pr} \{C_t\}$ , and  $x_i \equiv 1$  for  $i < 0$ .

It follows immediately from these definitions that  $w_{j+1,t}$  can be recovered from  $w_{j,t-1}$  using the  $x_i$ :

$$\begin{aligned} w_{j+1,t} &= w_{j,t-1} \frac{x_{j-m}}{1+g}, & j &= 1, 2, \dots \\ &= 1 - \sum_{i=2}^{J-T+t} w_{i,t-1}, & j &= 0. \end{aligned} \quad (10)$$

where the oldest fund at time  $t$  is  $J-T+t$  years old. Moreover, the  $\hat{w}_{j,T}^T$  can be recovered from  $w_{j,T}$  using the following relation:

$$\hat{w}_{j,T}^T = \frac{w_{j,T} x_{j-m}}{\sum_{i=1}^J w_{i,T} x_{i-m}}, \quad j = 1, 2, \dots, J. \quad (11)$$

We are now ready for the proposition:

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

- 1) the conditional mean of  $R_t$  only depends on and is strictly increasing in the number of direct cuts: for any given  $\tau = 1, 2, \dots, m$ , all elements of  $\boldsymbol{\mu}_\tau$  are the same, and  $\mu < \mu_1 < \mu_2 < \dots < \mu_m$ ;
- 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  $\tau$ ;

then the bias in the average performance estimate for the whole sample is increasing in the length of the sample period,  $k$ .

**Proof:** Start with a  $k$ -period sample ending at time  $T$ . Its survivorship-biased mean is  $\bar{\mu}_k^T$ . In general, we can increase the sample period length  $k$  by adding a year to the end of the sample period ( $\bar{\mu}_{k+1}^{T+1}$ ) or by adding a year to the beginning of the sample period ( $\bar{\mu}_{k+1}^T$ ). 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. Also note that in a steady-state, the industry must have a countably infinite number of cohorts of funds sorted by age (fund-age cohorts).

The definition of a steady-state together with (10) and (11) can be used to obtain the following expression for  $\hat{w}_{j+1,T}^T$  as a function of  $\hat{w}_{j,T}^T$ :

$$\hat{w}_{j,T}^T = \hat{w}_{j-1,T}^T \frac{x_{j-m}}{1+g}, \quad j = 2, 3, \dots \quad (12)$$

Using the assumption that  $x_{j-m} = x$  for any  $j-m \geq 0$  together with (12) allows us to write  $\hat{w}_{j,T}^T$  in the following way:

$$\begin{aligned} \hat{w}_{j,T}^T &= \hat{w}_{1,T}^T \left( \frac{1}{1+g} \right)^{j-1}, & j &= 1, 2, \dots, m-1, \\ &= \hat{w}_{i,T}^T \left( \frac{x}{1+g} \right)^{j-i} & j, i &\geq m. \end{aligned} \quad (13)$$

Substituting this expression into (8), which holds exactly because of assumption 1), gives the

following expressions for the cross-sectional survivorship-biased mean for time  $T+1-\tau$ :

$$\mu_{T+1-\tau}^T \equiv \frac{\hat{w}_{1,T}^T \left( \sum_{j=\tau}^{m-1} \frac{1}{(1+g)^{j-1}} \mu + \sum_{j=m}^{m-1+\tau} \frac{x^{j-(m-1)}}{(1+g)^{j-1}} \mu_{j-(m-1)} + \sum_{j=m+\tau}^{\infty} \frac{x^{j-(m-1)}}{(1+g)^{j-1}} \mu_{\tau} \right)}{\hat{w}_{1,T}^T \left( \sum_{j=\tau}^{m-1} \frac{1}{(1+g)^{j-1}} + \sum_{j=m}^{m-1+\tau} \frac{x^{j-(m-1)}}{(1+g)^{j-1}} + \sum_{j=m+\tau}^{\infty} \frac{x^{j-(m-1)}}{(1+g)^{j-1}} \right)} \quad (14)$$

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

$$\mu_{T+1-\tau}^T = \frac{\hat{w}_{\tau-1,T}^T \left( \sum_{j=1}^m \left( \frac{x}{1+g} \right)^j \mu_j + \sum_{j=m+1}^{\infty} \left( \frac{x}{1+g} \right)^j \mu_m \right)}{\hat{w}_{\tau-1,T}^T \left( \sum_{j=1}^{\infty} \left( \frac{x}{1+g} \right)^j \right)} \quad (15)$$

for  $\tau > m-1$ . Under the assumption that the  $\mu_i$  are increasing in  $i$ , it follows from (14) that  $\mu_{T+1-\tau}^T$  is increasing in  $\tau$  for  $\tau = 1, 2, \dots, m$ . Moreover, equation (15) shows that  $\mu_{T+1-\tau}^T$  is constant for  $\tau \geq m$ . Thus,  $\bar{\mu}_k^T$  must be increasing in  $k$  for all  $k$ .

***Q.E.D.***

This proof demonstrates exactly the intuition driving both the proposition and the fact that the result can be expected to hold more generally. It does so by characterizing the cross-sectional mean as a function of the number of periods until time  $T$ . Moving back in time from  $T$ , the cross-sectional mean increases for the first  $m$  periods, at which point it reaches a steady-state value and becomes invariant to moving further back in the sample. Increasing  $k$  adds to the number of these steady-state means relative to the  $m$  means at the end of the sample which are all lower than the steady-state mean. Consequently, the greater weight on the steady-state means increases the sample average. More generally, the  $m$  means at the end of the sample can be expected to be lower than earlier means since these returns in the last  $m$  periods are subjected to fewer direct cuts than earlier returns in the sample. Hence,  $\bar{\mu}_{k+1}^T$  is larger than  $\bar{\mu}_k^T$  given the assumptions of the proposition and is typically larger in more general settings.<sup>7</sup>

At the same time, none of the three assumptions holds in general, and it is possible to construct examples in which the sample mean is no longer increasing in the sample length. Most important, indirect cuts to  $R_t$  can affect the conditional mean of  $R_t$ . Moreover, although direct cuts are generally expected to increase the conditional mean of  $R_t$ , indirect cuts can actually reduce this

conditional mean. Roughly speaking, the reason for this is that, 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. An example of this effect is provided in the next section.

Another complication is that funds of different ages disappear at different rates, causing the weights of the different aged cohorts to change over time. Funds younger than  $m$  at time  $t$  do not disappear at time  $t$  while funds aged  $j$  disappear at rate  $1-x_{j-m}$ . Typically,  $x_{j-m}$  is varying as a function of  $j$  for all  $j \geq m$ . Recalling that  $x_i = pr \{C_t \mid C_{t-1}, \dots, C_{t-i}\}$ , 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, even though cuts prior to  $C_{t-m+1}$  do not overlap with  $C_t$ ,  $x_i$  also varies as a function of  $i$  for  $i > m-1$ . For such an  $x_i$ , the cuts  $C_{t-i}, \dots, C_{t-m}$  still affect the probability of  $C_t$ , despite the lack of overlap, because of their interaction with the cuts  $C_{t-m+1}, \dots, C_{t-1}$ , which do overlap with  $C_t$ .

Finally, if the assumption of a steady state is relaxed, it matters whether the sample length is increased by adding a year to the end of the sample period ( $\bar{\mu}_{k+1}^{T+1}$ ) or by adding a year to the beginning of the sample period ( $\bar{\mu}_{k+1}^T$ ). In particular, if we increase the sample period length by adding a year to the beginning of the sample period, the cross-sectional mean  $\mu_{T-k}^T$  is added into the time series average. With a maximum fund age at time  $T$  that is finite, the cross-sectional survivorship-biased mean for time  $T+1-k$  ( $\mu_{T+1-k}^T$ ) may start declining in  $k$  for  $k$  sufficiently large, since the earliest periods of the industry have only very young funds whose early-period returns have very few direct cuts. At the extreme, there are only one-year-old funds in the first period of the industry's existence, and those that survive until  $T$  have a first period return subjected to only one direct cut. Thus,  $\bar{\mu}_k^T$  may be hump-shaped as a function of  $k$ , rather than increasing.

The next two subsections present an example which shows that the survivorship bias in average performance can decrease in the sample period length and a simulation of survivorship biases in an industry calibrated to match that of U.S. mutual funds.

### *B. Counterexample*

Using the formulas developed in the previous subsection, Table 6 presents two examples in which  $m = 2$ , a fund's return  $R_t$  can take on the values 1 (win) or -1 (loss) with equal probability. The critical return level is  $b = -1.5$ , so a fund disappears after two straight losses. The fund industry has a continuum of funds with independent returns. Panel A lists the survival rates  $x_0, \dots, x_4$  while Panel B gives a matrix of conditional mean returns, the  $\mu_{ij}$ 's, for  $j$  ranging from 0 to 4. The example on the right, labeled Counterexample, uses the actual survival rates and conditional means obtained from the assumed return generating process and survival rule. The example on the left, labeled Proposition 2 example, fixes all the conditional survival probabilities ( $x_1, x_2, \dots$ ) equal to the unconditional survival probability,  $x_0$ , and sets conditional means with the same number of direct cuts equal. Any conditional mean with one direct cut is set equal to  $E[R_t|C_t]$  while any conditional mean with two direct cuts is set equal to  $E[R_{t-1}|C_t, C_{t-1}]$ .

The right-hand side of Panel A shows that for the assumed return generating process and survival rule, a fund's survival probability depends on the number of previous periods survived. Similarly, the right-hand side of Panel B shows that conditional mean returns with the same number of direct cuts are not always the same. As the left-hand side shows, conditional means with just one direct cut, (elements of  $\boldsymbol{\mu}_1$ ), lie in the first row and along the main diagonal of the matrix. In the Counterexample, on the right-hand side, we observe variation in conditional means going across the

first row and down the main diagonal. Similarly, the remaining conditional means in the matrix have two direct cuts (i.e., are elements of  $\mu_2$ ), and we see considerable variation across these means on the right-hand side as well.<sup>8</sup> At the same time, the intuition that the conditional means are increasing in the number of direct cuts is also borne out by the Counterexample. An examination of the right-hand side of Panel B shows that, for this return generating process and survival rule, the smallest conditional mean with 2 direct cuts (0.385) is still larger than the largest conditional mean with only one such cut (0.333).

We can use the Counterexample to demonstrate precisely how indirect performance cuts can reduce conditional mean returns. Consider a sequence of four returns  $R_1, R_2, R_3$  and  $R_4$  and the three associated performance cuts  $C_2, C_3$ , and  $C_4$ . Our focus is on the conditional mean of  $R_2$ . For this return,  $C_2, C_3$  are direct cuts while  $C_4$  is an indirect cut. Applying the two direct cuts gives us  $E[R_2 | C_2, C_3]$ , which corresponds to  $\mu_{2,1} = 0.6$  from the Counterexample in Panel B. Notice that applying  $C_3$  eliminates any path with  $R_2$  and  $R_3$  both equal to -1. If we then add the indirect cut  $C_4$ , we obtain  $E[R_2 | C_2, C_3, C_4] = \mu_{3,2} = 0.5$  from the Counterexample in Panel B. Adding the indirect cut reduces the conditional mean of  $R_2$ . To see why, recall that  $C_4$  cuts paths with both  $R_3$  and  $R_4$  equal to -1. But after the application of the direct cuts  $C_2$ , and especially  $C_3$ , the only paths left with  $R_3$  equal to -1 must have  $R_2$  equal to 1. Thus, adding the indirect cut  $C_4$  only removes paths with  $R_2$  equal to 1, which must reduce the conditional mean of  $R_2$ . Similar intuition can be applied to  $R_3$  in the five return sequence  $R_1, R_2, R_3, R_4$  and  $R_5$  to explain the drop in the conditional mean from 0.5 to 0.385 going from  $\mu_{2,2}$  to  $\mu_{3,3}$ .

The Proposition 2 example is designed to satisfy the three assumptions of the proposition, so the cohort weights at time  $T$  reported in Panels C (prior to the time- $T$  cut) and D (just after the

time- $T$  cut) are chosen to be the steady-state weights. The definition of a steady-state together with (10) are used to obtain the time- $T$  steady-state weight for 1-year-olds before the time- $T$  cut ( $w_{1,T}$ ). As a result, the 1-year-old weights are the same for all possible  $T$  values, both before and after the cut. Under the Proposition 2 assumptions, the precise distribution of funds aged 2 and older at time-1 does not affect the cross-sectional mean calculations and so for ease of exposition, 2-year-olds get all the weight not assigned to the 1-year-olds. For the Counterexample, 50% of the funds are 1 year old at time 1 before the time-1 cut while the remaining 50% are 3 years old.

Panel E reports a matrix of cross-sectional means: the element in row  $t$ , column  $T$ , is the cross-sectional mean time- $t$  return for the sample of time- $T$  survivors,  $\mu_t^T$ . For each example, the elements of Panel E are obtained from the elements of Panels B and D using equation (6). The last panel, F, lists time series averages of the cross-sectional means for all possible subperiods in the four-year history of this mutual fund industry. The element in row  $t$ , column  $T$ , is the survivor-biased time-series average mean return for the subperiod starting in year  $t$  and ending in year  $T$ . Moving across a row of the matrix in Panel F of Table 6 means increasing the sample period length by adding a year to the end of the sample period. Moving up a column means increasing the sample period length by adding a year to the beginning of the sample period.

Since the assumptions of Proposition 2 hold for the left-hand-side example, we see that the only determinant of the time series averages in Panel F is the length of the sample period. Moreover, the average is increasing in the sample length. Consistent with the proof of the proposition, the left-hand side of Panel E shows that the cross-sectional mean increases rolling back in time from  $T$ , and reaches a steady-state value  $m-1$  periods prior to time  $T$ : for example, when  $T=4$ , the cross-sectional mean increases going from time 4 (0.250) to time 3 (0.533) and remains at this steady-state value

for all earlier points in time.

In the Counterexample, non-monotonicities are apparent in both directions. For example, in the top row, adding a year to the end of the sample running from time 1 to time 2 reduces the average return estimate from 0.337 to 0.335. In the third column, adding a year to the start of the sample running from time 2 to time 3 reduces the average return estimate from 0.360 to 0.335. So while it matters whether the sample is extended forward or backward, it is possible to obtain non-monotonicities in the sample average by extending the sample in either direction.

There are two main sources of these non-monotonicities. The first is the variation in conditional means with the same number of direct cuts. In particular, Panel B shows that with 4 conditioning statements ( $j=3$ ), there is a substantial dip in the conditional mean,  $\mu_{i,j} \equiv E[R_{t+1-i} | C_{t-j}, \dots, C_t]$ , as the return is moved back in time by increasing  $i$ . This dip occurs at  $i=3$  even though the number of direct cuts remains at 2 as  $i$  goes from 2 to 4. The second source is the large weight that the Counterexample assigns to 1-year-olds at time 1. As discussed above, the time-1 return for these 1-year-old funds is subject to only one direct cut, and so the large weight for these funds at time 1 drags down the survivorship-biased time-1 cross-sectional mean. This causes the observed hump-shaped pattern in the sample average as periods are added to the start of the sample (by moving up a column).

That both sources are playing a role is confirmed by the following unreported results. First, we replace the conditional mean matrix for the Counterexample with that for the Proposition 2 example and find that non-monotonicities remain. Second, the time-1 fund weights are modified so that all the funds are 2-years old. This modification eliminates the hump-shape discussed above, but again, non-monotonicities are still present.



### C. Calibration

The counterexample above demonstrates some of the counterintuitive effects of increasing the length of the sample period. However, the most important effect is driven by the fact that a return near the end of the survivorship-biased sample is conditioned on fewer direct cuts than other returns in the sample. Because of this, average performance estimates are typically lower in the last  $(m-1)$  periods than the rest of the survivorship-biased sample and the time-series mean is increasing in the sample period length. To illustrate how this main effect typically swamps the counterintuitive effects in more realistic settings, we generate a mutual fund history designed to match the U.S. mutual fund industry. In particular, 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. The first way, used in Panel A of Table 7, 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. The second way, used in Panels B and C of Table 7, fixes  $b/\sqrt{m}$  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 while Panel C assumes they are all only 1 year old. 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, according to the formulas supplied above. Finally, for each given sample period length  $k$ , we average the

performance estimate across all possible subperiods of length  $k$ . Table 7 reports this average (in percentage) 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 7, 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 1-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. Even so, Panel C shows that the decline is very small in magnitude, with the largest decline for a 1-year increase in  $k$  of 0.04 basis points. This result is not surprising since the cross-sectional mean for the incremental year associated with a sample size increase has only a small impact on the sample average when the sample size is large. To summarize, these calibration results suggest that the counterintuitive effects of increasing the sample period length illustrated in the binomial example above are not likely to play an important role in the U.S. mutual fund industry.

#### *D. Evidence*

We now measure the bias in average performance estimates due to survivor bias as a function of the sample period length. Our objective is to obtain a rough rule of thumb on appropriate corrections for researchers using survivor-biased U.S. equity mutual fund samples. As shown above, the bias in average performance estimates depends on the length of the sample period if a multi-period performance survival rule is in effect. Further, the results in Section 3 demonstrate that funds disappear primarily due to multiple-year performance. In our data set, we find a strong positive

relation between survivor bias and sample time length, as predicted by the theory and calibration above.

We consider all the possible survivor-biased samples 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 return induced by including only survivors. We report the average survivor bias across all possible  $k$ -year samples for various sample lengths. We also calculate correlation-adjusted standard errors assuming independent and identically distributed annual returns.<sup>9</sup>

Table 8 shows that survivor bias strongly increases in the sample time length. For a survivor-biased sample of only one year, the bias in average return is only 17 basis points, whereas the bias is 43 basis points per year for survivor-biased samples of five years. For samples greater than fifteen years, the hypothesis that survivor bias is one percent 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. Interestingly, over the complete 34-year period, the survivor-biased sample outperforms the value-weighted CRSP index by 0.5 percent per year while the unbiased sample underperforms the index by 0.6 percent annually.

Figure 2 plots the survivor bias in average performance estimates over all sample time lengths. The figure suggests that survivor bias levels off at about one percent per year for intervals of around fifteen years or longer. Thus, while there is no single rule of thumb on the magnitude of survivor bias, for time periods of fifteen years or longer, one percent is probably a good approximation of the bias in mean annual return estimates.

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

### A. *Theory*

Brown, Goetzmann, Ibbotson and Ross (1992) and Carpenter and Lynch (1999) provide a detailed theoretical analysis of the effect of survivorship bias and look-ahead bias on estimates of persistence in mutual fund performance, where persistence is defined as a positive relation between performance in an initial ranking period and a subsequent evaluation period. Brown et al. (1992) show that if mutual funds 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 cut, higher volatility funds have higher means. Thus, in a sample of survivors, first (ranking) period winners tend to be high volatility funds. The same high volatility funds tend to win in the second (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. In simulations of typical persistence tests on samples of funds calibrated to match U.S. mutual fund data, 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.

While the spurious persistence and reversal effects are present in both look-ahead biased and survivor-biased samples to some degree, Carpenter and Lynch (1999) distinguish the effects of the look-ahead and survivorship biases. They show that the additional cut associated with the survivor-biased sample has offsetting effects on persistence measures. On one hand, the incremental cut of funds that perish after the look-ahead period tends to remove high volatility funds, reducing the cross-sectional dispersion of volatility in the sample, and therefore reducing the spurious persistence. On the other hand, with a multi-period survival rule, the incremental cut can also reduce the spurious reversal effect. As they explain it, we can consider three periods, the ranking period, the evaluation period, and the period after that. The incremental cut of the survivor-biased sample removes funds that survive the first two periods but perish in the third. These tend to have a performance pattern good-bad-bad or bad-good-bad. Funds that win in both the first two periods are too strong to perish in the third. Funds that lose in the first two periods are already eliminated from the look-ahead biased sample. In this way, the incremental cut tends to remove reversers. Which of these various effects dominates in the data depends on the nature of real survival rules and the degree of cross-sectional dispersion in volatility. We examine the impact of look-ahead bias and survivor-bias on persistence measures empirically below.

### *B. Evidence*

This section studies the effect of survivor 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. As discussed earlier, this empirical section can only characterize these biases for a particular survival rule, namely the one in effect for the U.S. mutual fund industry.

Annually, we form ten equal-weighted portfolios of mutual funds sorted on a lagged performance measure. We hold the portfolios for one year, then re-form them. This yields a time-series of monthly returns on each decile portfolio over the complete time period, 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.

It is worth taking a few sentences to compare the look-ahead bias associated with this method to that associated with the “follow the money” approach of Elton, Gruber and Blake (1996a and b) and Gruber (1996). Exploiting the fact that almost all disappearing funds are merged with other funds, the “follow the money” approach generates a return series for each fund in the sample by using a disappearing fund’s returns up until the time of its disappearance and the merging fund’s returns thereafter. Given a sample with complete fund return histories except for any partial month just prior to disappearance, both the Carhart (1997) and the “follow the money” approach induce at most, a one month look-ahead bias in the evaluation period returns. The reason is that both methods include all nonsurvivor returns up until the time of disappearance in the calculation of evaluation period performance.<sup>10</sup>

At the same time, the methods impose different conditioning in the ranking period. In particular, a fund that disappears during the ranking period would be eliminated in the Carhart (1997) approach, but would be included in the portfolio formation stage in the “follow the money” approach. One might regard this feature of the Carhart (1997) methodology as a form of look-ahead

bias in the ranking period. However, when fund returns are independent through time, the absence of any look-ahead bias in the measurement of evaluation period performance implies the absence of any bias in the methodology's persistence measure, irrespective of the look-ahead bias associated with ranking period performance. The lack of any look-ahead bias in the evaluation period means that the distribution of evaluation period performance is not truncated. Therefore, the evaluation period performances of funds in winner and loser groups are all equal to the unconditional fund-performance mean, irrespective of exactly how the winner and loser groups are formed. Consequently, the Carhart (1997) approach used here and the "follow the money" approach can both be regarded as being virtually free of look-ahead bias.

The results in Table 9 suggest that survivor bias attenuates the evidence of persistence in mutual fund performance. Panel A reports persistence test statistics using the complete sample. Consistent with Carhart (1997), the portfolios demonstrate strong persistence in mean return, most of which is explained by the 4-factor model and expense ratios. For the portfolios sorted on one-year return, the post-formation spread in monthly returns between deciles 1 and 10 is a sizeable and statistically significant 63 basis points per month. The 4-factor model explains all but 24 basis points per month of this spread, and this remainder is insignificantly different from zero. The difference in average annual expense ratios of 52 basis points between deciles 1 and 10 explains a further 4 basis points of this spread in performance on these portfolios. With a p-value of 14.8 percent, the Spearman rank ordering test also fails to reject the hypothesis that the 4-factor alphas are randomly ordered. The results for the lagged five-year return and 4-factor alpha portfolios offer similar conclusions on performance persistence, except that the 4-factor model explains smaller amounts of the measured persistence in expected return.

The evidence favoring persistence in mutual fund returns is weaker in the sample of survivors. Panel B, which repeats the tests of Panel A using the survivor-biased sample of funds, shows that spreads in mean return and 4-factor model performance shrink considerably relative to the complete sample, and the statistical significance diminishes as well. In some cases, an econometrician using the survivor-biased sample may incorrectly reject persistence. Evidently, excluding nonsurvivors attenuates persistence because nonsurvivors consistently underperform. While the 4-factor alphas are somewhat higher for all portfolios in the survivor-biased sample, decile 10's performance is especially increased, amounting to approximately 20 basis points per month.

Panel C examines the effect of look-ahead bias separately from the effect of survivor bias. These tests use the full sample of survivors and nonsurvivors, but require 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. This is the bias simulated by Brown, Goetzmann, Ibbotson and Ross (1992).

Although the results are still downwardly biased, the look-ahead biased tests in Panel C do not impact the results from persistence tests as much as survivor bias does. Using the look-ahead biased sample changes the inference only for the five-year returns-sorted portfolios, the longest look-ahead period. The downward bias in the persistence measure induced by using both look-ahead biased and survivor biased samples relative to the full sample is understandable given the multi-period nature of the survival rule documented earlier.

Finally, we undertake Hendricks, Patel and Zeckhauser's (1996) test for spurious persistence



due to survivorship. Hendricks et al. (1996) show that when performance is categorized finely, the relation between pre- and post-period rankings will be J-shaped in a survivor-biased sample or using a look-ahead biased methodology. They devise a regression test for this convexity, which we employ in our survivor- and look-ahead-biased samples. Under the hypothesis that performance persists spuriously due to survivorship, the HPZ J-shape t-statistic should be reliably negative. Moreover, Carpenter and Lynch (1999) present simulation evidence that the HPZ J-shape t-statistic is rarely reliably positive unless performance is truly persistent. Given this result, our finding that the HPZ J-shape t-statistics are all positive and often significant in our survivor-biased and look-ahead-biased samples supports the conclusion that U.S. mutual funds exhibit true persistence.

## **6. Effects of Survivor Bias on Cross-Section Tests**

In this section, we demonstrate that survivor bias can affect estimates of the cross-sectional relations between fund performance and fund characteristics, but only when the fund characteristics in question affect the bias in performance induced by survivor-only conditioning. We also show that these effects can be substantial when using our survivor-biased sample of U.S. mutual funds and some fund characteristics that are commonly used in such cross-sectional regressions.

### *A. Theory*

This subsection establishes the link between survivor bias and cross-sectional regressions of performance on fund characteristics. Consider the following cross-sectional regression of time- $t$  performance on a fund characteristic:

$$\alpha_{it} = a_t + b_t q_{it} + \varepsilon_{it} \tag{16}$$

where  $\alpha_{it}$  is an individual fund performance estimate and  $q_{it}$  is a fund characteristic. This regression implies the following expression for expected fund- $i$  performance conditional on the value of the characteristic for fund  $i$ :

$$E[\alpha_{it} | q_{it}] = a_t + b_t q_{it} \quad (17)$$

which is linear in  $q_{it}$ . We next need to quantify the bias in the estimate of  $b_t$  induced by a requirement that fund  $i$  survives until the end of the sample, a condition we denote by  $S_{iT}$ , where  $T$  is the end of the sample period. This bias can depend on  $q_{it}$ . For simplicity we assume that any such dependence is linear:

$$E[\alpha_{it} | S_{iT}, q_{it}] - E[\alpha_{it} | q_{it}] = c_t + d_t q_{it}. \quad (18)$$

This bias is analogous to that in the conditional  $\mu_{ij}$ 's introduced in section 4, except here we are allowing the possibility that the return distribution and the survival rule can depend on fund characteristics. Adding these two expressions together gives an expression for the expected fund- $i$  performance conditional on the fund- $i$  value of the characteristic and the survival of fund  $i$  until the end of the sample:

$$E[\alpha_{it} | S_{iT}, q_{it}] = (a_t + c_t) + (b_t + d_t) q_{it} \quad (19)$$

which is also linear in  $q_{it}$ . Given this expression, it follows that when the cross-sectional regression in (16) is run using the survivor-only sample, the intercept and slope become  $(a_t + c_t)$  and  $(b_t + d_t)$  respectively.<sup>11</sup>

Consequently, the expression for the survivor bias in mean performance (18) determines the bias imparted to the cross-sectional regression coefficients. Moreover, the sign and magnitude of the coefficient relating the survivor bias in performance to the fund characteristic,  $d_t$ , determines the sign and magnitude of the survivor bias in the slope coefficient in the cross-sectional regression. In particular, if the survivor bias in performance does not vary with the fund characteristic,  $q_{it}$ , then the

slope coefficient in the cross-sectional regression is unaffected by the use of the survivor-only sample rather than the full sample.

### *B. Evidence*

It is worth examining whether the magnitude of the bias in the cross-sectional regressions described above can be large for a survivor-biased sample of U.S. mutual funds. The cross-section methodology follows Carhart (1997). As in Fama and MacBeth (1973), we estimate the cross-sectional relation in **(16)** each month, then average the coefficient estimates across the complete sample period. To mitigate look-ahead bias, we estimate  $\alpha_{it}$  as a one-month abnormal return from the 4-factor model, where the 4-factor model loadings are estimated over the prior three years. We consider expense ratio, turnover, modified turnover (Mturn),  $\ln(\text{TNA})$ , and maximum load fees as explanatory variables. As in Carhart (1997),  $\ln(\text{TNA})$  and load fees are lagged one year.

We report the results from these tests for the full and survivor-only samples in Table 10 and find that using the survivor-only sample causes the slope coefficients to go from being positive to negative on fund size  $\ln(\text{TNA})$ , and to become less negative on all the other characteristics. The magnitude of the bias can be large. For example, the slope on  $\ln(\text{TNA})$  goes from being 0.16 in the full sample to -0.03 in the survivor-only sample.

The directions of these induced biases can be explained using the intuition described in the previous subsection. In particular, for a given fund characteristic, the direction and magnitude of the bias in the slope coefficient is determined by the impact of the characteristic on the survivor bias in performance. The results of the probit analyses in section 3 and in Brown and Goetzmann (1995) can be used to infer the impact of the characteristics on the survivor bias in performance. From

Table 4, we see that the probability of disappearance is negatively related to fund size, holding performance fixed. This suggests that the bias in performance induced by survival is likely to be decreasing in fund size (a negative  $d_i$ ). Based on the analysis above, this implies a lower slope coefficient on fund size, which is exactly what Table 10 finds.

An analogous argument that exploits the probit results in Brown and Goetzmann (1995), can be used to explain other results in Table 10. They find that an expense variable has a positive impact on the probability of disappearance, again holding performance fixed. Since turnover is highly correlated with expenses, it is likely that a turnover variable would also enter such a probit positively. This suggests a positive impact for both these variables on the bias in performance induced by survival, which implies that using the survivor-only sample imparts a positive bias to the slope coefficient for these variables.

## 7. Summary and Conclusions

Using Carhart's (1997) sample of U.S. mutual funds, we find that funds disappear primarily because of poor multi-year performance, rather than a single poor annual return. We demonstrate both analytically and empirically that this survival rule typically causes the bias in estimates of average annual return to increase in the sample length. In our sample, the bias is economically small at 17 basis points for one-year samples, but a significantly larger one percent for samples longer than fifteen years. At the same time, we show that it is possible to construct examples in which the average performance bias is not increasing as a function of the sample length.

In tests of mutual fund performance persistence using U.S. data, both survivor-only and look-ahead conditioning weaken the evidence of persistence. This sometimes results in rejections of

persistence when the evidence is statistically significant in the full sample. This result is consistent with the use of a multi-period survival rule in the U.S. mutual fund industry. However, the evidence favoring persistence does not necessarily support the existence of skilled or informed portfolio managers; Carhart (1997) shows that persistence is mostly explained by investment expenses. Finally, we 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.

This paper attempts to provide a comprehensive analysis of survivorship issues in mutual fund studies. We document a few rules of thumb for incorporating the effects of survivorship, but also provide a warning that the nature of the biases imparted can be quite complicated, especially since the fund industry appears to use a multi-period survival rule.

More generally, 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 survivorship biases induced by a multi-period survival rule 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.

In particular, many areas of finance run cross-sectional regressions with performance as the independent variable. Our study cautions that 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 study warns that data unavailability for failed equity markets can have important ramifications for such cross-country comparisons, particularly if the characteristics in question affect the survival rule applied to a given equity market.

One final point is worth making. While the researcher may have little choice as to the data set, the researcher does get to choose the methodologies employed on the data set. Our paper stresses the importance of choosing methodologies that are free of selection bias.

## 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.
- 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.
- Person, Wayne E., and Rudi W. Schadt, 1994, "Measuring Fund Strategy and Performance in

- Changing Economic Conditions,” working paper, University of Washington School of Business Administration, Seattle, WA.
- 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.
- Grinblatt, Mark, and Sheridan Titman, 1989, “Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings,” *Journal of Business*, 62, 393-416.
- Grinblatt, Mark, and Sheridan Titman, 1992, “The Persistence of Mutual Fund Performance,” *Journal of Finance*, 47, 1977-1984.
- 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.
- Jensen, Michael C., 1968, “The Performance of Mutual Funds in the Period 1945-1964,” *Journal of Finance*, 23, 389-416.
- Lintner, John, 1965, “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economics and Statistics*, 47, 13-37.



- Malkiel, Burton G., 1995, "Returns from Investing in Equity Mutual Funds 1971-1991," *Journal of Finance*, 50, 549-572.
- Sharpe, William F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under conditions of Risk," *Journal of Finance*, 19, 425-42.
- Sirri, Erik R., and Peter Tufano, 1992, "The Demand for Mutual Fund Services by Individual Investors," working paper, Harvard Business School, Boston, MA.
- United and Babson Mutual Fund Selector*, semi-monthly, January 1976 to December 1983, Wellesley, MA.
- Wermers, Russ, 1997, "Momentum Investment Strategies of Mutual Funds, Performance Persistence, and Survivorship Bias," working paper, University of Colorado at Boulder, Boulder, CO.
- Wiesenberger, Arthur & Co., 1962-1993, *Investment Companies*, Warren, Gorham & Lamont, Inc., Boston, MA.

## ENDNOTES

1. We thank Will Goetzmann for this clarification.
2. Exchange was the only method of acquiring shares in these funds, although shares were redeemable for cash.
3. 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 20 basis points.
4. When a fund's termination date is unknown, we assume the fund terminates in the month after its last return observation in our sample.
5. 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.
6. Also note that  $\mu_{i,j} = \mu_{j+m+1-i,j}$  since for a fixed number of consecutive conditioning statements, all that matters for the conditional one-period return is its position relative to the statements, with time going forward or backward.
7. This intuition also applies to the methodology requiring funds to be in existence from the start of the sample period until the end to be included in the sample. With this methodology, new funds are not added to the sample. Consequently, sample funds receive  $m$  direct cuts in all

periods except the last  $(m-1)$  periods and possibly the first  $(m-1)$ , depending on the fund-age distribution at the start of the sample. Increasing the sample period length by extending the ending date increases the number of periods with  $m$  direct cuts, and so can be expected to increase the sample's average performance.

8. Notice that incremental conditions can actually decrease the conditional mean of  $R_{t+1-j}$ . Moving across a row in this matrix (increasing  $j$ ) means adding conditions associated with an incremental performance cut, but does not necessarily increase  $\mu_{i,j}$ .

9. 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 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  $n$  is the sample time length in years,  $T$  is the number of samples in the database, and  $std(R)$  is the standard deviation in annual returns.

10. It is worth noting that the Elton, Gruber and Blake sample includes partial-month returns just prior to disappearance, whereas our sample does not. However, this is a data difference rather than a methodology difference.

11. One complication that we ignore here is that  $c_i$  and  $d_i$  are likely to depend on  $i$ .

**Table 1**

## Mutual Fund Database Annual Summary Statistics, 1962 to 1995

Annual cross-sectional averages from 1962 to 1995. Total funds are the total number of funds at the beginning of the year and new funds are funds started during that year. Merg is the number of funds that merged, Liq is the number of funds that liquidated and Oth is the number of funds that disappeared for a reason other than merger or liquidation. Attrition rate is the number of disappearing funds divided by the total number of funds at the beginning of the year. TNA is total net assets, Exp ratio is the total annual management and administrative expenses divided by average TNA. Mturn is modified turnover and represents reported turnover plus 0.5 times the absolute value of Flow. Flow is the percentage change in TNA adjusted for investment return and mergers. Maximum load is the total of maximum front-end, rear-end, and deferred sales charges as a percentage of the investment. Percent stock is the proportion of total net assets invested in equity securities. EW fund return is the equally-weighted average annual mutual fund return and CRSP return is the value-weighted average annual return on all NYSE, AMEX, and NASDAQ stocks in the CRSP database.

Year	Beg Total Funds	New Funds	Disappearing Funds			Attrition Rate (%/yr)	Avg TNA (\$ mil)	Avg Exp Ratio (%/yr)	Avg Mturn (%/yr)	Avg Flow (%/yr)	Percent with Load	Avg Total Load	Percent Common Stock	EW Fund Return	CRSP Return
			Merg	Liq	Oth										
1962	213	16		1		0.5%	72.0	0.82%	NA	9.8%	82.9%	7.70	86.9%	-15.8%	-10.3%
1963	228	13	2	5	1	3.5%	82.0	0.94%	NA	0.4%	80.7%	7.73	88.4%	18.4%	20.9%
1964	233	12	2	2		1.7%	95.1	0.82%	NA	5.5%	80.5%	7.76	89.0%	12.4%	16.3%
1965	241	21	1			0.4%	110.0	0.84%	NA	6.7%	78.0%	7.80	88.1%	23.0%	14.4%
1966	261	26	3	2		1.9%	101.0	0.84%	72.6%	11.4%	78.4%	7.91	84.4%	-5.8%	-8.7%
1967	282	40	2	3		1.8%	121.7	0.89%	75.1%	14.0%	77.9%	7.86	84.2%	36.7%	28.6%
1968	316	66	1			0.3%	135.1	0.95%	81.4%	27.8%	78.4%	7.96	81.5%	16.3%	14.1%
1969	382	94		3		0.8%	100.2	1.04%	86.6%	20.8%	76.7%	8.10	80.6%	-14.1%	-10.8%
1970	473	68	8	12	3	4.9%	83.3	1.19%	89.5%	10.4%	74.5%	8.06	81.5%	-9.4%	0.1%
1971	517	39	6	14	2	4.3%	93.2	1.38%	87.6%	5.5%	71.6%	8.04	85.5%	19.8%	16.2%
1972	535	19	12	19		5.8%	102.7	1.27%	79.5%	1.4%	66.9%	8.10	87.1%	10.8%	17.3%
1973	523	11	26	8	3	7.1%	81.8	1.26%	67.1%	-1.0%	66.2%	8.10	82.2%	-24.5%	-18.8%
1974	497	3	23	10	2	7.0%	62.3	1.36%	57.7%	1.0%	65.7%	8.11	78.1%	-24.9%	-27.8%

**Table 1 - continued**

Year	Beg Total Funds	New Funds	Disappearing Funds			Attrition Rate (%/yr)	Avg TNA (\$ mil)	Avg Exp Ratio (%/yr)	Avg Mturn (%/yr)	Avg Flow (%/yr)	Percent with Load	Avg Total Load	Percent Common Stock	EW Fund Return	CRSP Return
			Merg	Liq	Oth										
1975	465	4	31	9		8.6%	86.4	1.41%	59.7%	-1.5%	64.5%	8.11	83.4%	33.6%	37.4%
1976	429	7	19	5	10	7.9%	98.8	1.27%	65.1%	-10.5%	63.7%	7.90	87.5%	24.9%	26.8%
1977	402	15	18	5	1	6.0%	89.3	1.31%	53.7%	-6.8%	59.7%	7.77	82.1%	2.3%	-3.0%
1978	393	11	21	4	3	7.1%	88.5	1.32%	71.3%	-8.5%	57.4%	7.69	82.7%	11.2%	8.5%
1979	376	6	6	4	3	3.5%	101.1	1.28%	71.4%	-10.5%	55.3%	7.67	83.6%	30.0%	24.4%
1980	369	18	5	2	1	2.2%	124.5	1.21%	84.2%	-3.9%	53.6%	7.69	84.4%	32.2%	33.2%
1981	379	17	9	3		3.2%	116.4	1.17%	77.5%	-0.5%	51.1%	7.68	79.1%	-1.2%	-4.0%
1982	384	34	9	2		2.9%	141.3	1.29%	90.6%	7.6%	48.8%	7.64	81.6%	25.9%	20.4%
1983	407	52		2	2	1.0%	177.3	1.15%	97.9%	17.2%	46.2%	7.50	82.9%	19.6%	22.7%
1984	455	52	2	2	3	1.5%	171.5	1.12%	85.6%	7.8%	44.0%	7.36	80.7%	-1.4%	3.3%
1985	500	91	1	2	4	1.4%	203.5	1.17%	97.5%	12.0%	42.7%	7.12	81.6%	26.8%	31.4%
1986	584	100	2	3	1	1.0%	218.5	1.20%	96.9%	18.8%	44.0%	6.53	80.9%	13.3%	15.6%
1987	678	116	3	3		0.9%	206.8	1.28%	104.7%	12.2%	45.1%	6.06	80.8%	0.8%	1.8%
1988	788	86	10	8	3	2.7%	203.6	1.41%	87.4%	-1.4%	45.5%	5.64	78.9%	14.3%	17.6%
1989	853	57	20	12	5	4.3%	254.2	1.39%	83.5%	4.1%	46.2%	5.42	80.1%	23.5%	28.5%
1990	873	83	18	18	3	4.5%	232.1	1.42%	101.2%	2.4%	48.1%	5.07	79.2%	-6.0%	-6.0%
1991	917	88	28	17	2	5.1%	332.3	1.32%	NA	15.1%	46.9%	4.96	82.0%	34.5%	33.6%
1992	958	159	34	11	11	5.8%	389.8	1.36%	91.3%	21.5%	52.0%	4.74	82.0%	8.3%	9.0%
1993	1061	178	25	12	6	4.1%	466.4	1.31%	93.1%	22.2%	49.2%	4.78	83.3%	13.0%	11.3%
1994	1196	217	23	19	1	3.6%	478.0	1.31%	83.5%	12.3%	45.7%	4.61	87.1%	-1.7%	-0.6%
1995	1370	55	47	36	12	6.9%	681.0	1.30%	99.3%	14.7%	45.3%	4.62	88.6%	30.4%	35.7%
Mean	545.2	55	13	8	4	3.6%	179.5	1.19%	82.5%	7.0%	59.8%	7.052	83.2%	11.1%	11.7%
Std	287.2	52	12	7	3	2.4%	137.8	0.19%	13.3%	9.7%	13.6%	1.234	3.0%	16.9%	16.3%

**Table 2**

## Properties of Surviving and Nonsurviving Mutual Funds

Average performance and attributes of individual mutual funds by survival category. Survivors are those funds still operating December 31, 1995 and nonsurvivors are funds disappearing before this date. Group-adjusted performance is the time-series average of the difference between a fund's return and the average return on all other funds with the same declared fund objective over the fund's complete history. 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. Relative TNA for a given group is calculated as follows. For each year in the sample, the ratio of each fund's TNA to the average TNA for the entire sample in that year is averaged across funds in the group. Then, the time series average of these annual group averages gives the Relative TNA measure for the group. Relative expense ratio and Mturn are calculated in the same manner. To obtain Relative Flow for a group, we use the difference in Flow instead of the ratio, again taking the time-series average of annual cross-sectional within-group averages, where Flow measures the change in TNA adjusted for investment returns and fund mergers.

Group	Number of Funds	Cross-Sectional Average Abnormal Monthly Performance		Relative TNA	Relative Exp	Relative Mturn	Relative Flow
		Group Adjusted	4-Factor Alpha				
By current status							
Survivors	1,346	0.10%	-0.03%	1.45	0.89	0.96	1.2%
Nonsurvivors	725	-0.26%	-0.34%	0.28	1.23	1.15	-4.7%
Nonsurvivors by reason for disappearance							
Merged with another fund	417	-0.19%	-0.29%	0.32	1.13	1.14	-4.6%
Liquidated	258	-0.45%	-0.54%	0.05	1.85	1.53	-4.0%
Other, self-selected							
Removed at fund request	11	-0.10%	-0.25%				
Changed to closed-end fund	2	-0.86%	-0.64%				
Split into multiple funds	1	0.26%	0.09%				
	14	-0.0018	-0.0028	0.34	1.19	0.77	-0.1%
Other, not self-selected							
Tax-Free Exchange Fund	16	0.23%	NA				
Variable Annuity	5	-0.07%	0.04%				
Unknown	15	-0.14%	-0.33%				
	36	0.03%	NA	0.52	1.16	0.75	-1.9%

**Table 3**

## Performance of Equal-weight Mutual Fund Portfolios by Survival Category and Fund Objective

Performance of equal-weighted portfolios of mutual funds by survival and objective category from 1962 to 1995. RMRF, SMB, and HML are Fama and French's (1993) market proxy and factor-mimicking portfolios for size and book-to-market equity. PR1YR is a factor-mimicking portfolio for one-year return momentum. The t-statistics are in parentheses.

Portfolio	Mean Monthly Return	Standard Deviation	CAPM			4-Factor model					Adjusted R-square
			Alpha	Beta	Adjusted R-square	Alpha	Loadings on				
							RMRF	SMB	HML	PR1YR	
All funds	0.93%	4.36%	-0.05% (-0.96)	0.97 (76.91)	0.938	-0.15% (-4.17)	0.89 (105.65)	0.33 (25.66)	-0.06 (-4.36)	0.09 (9.09)	0.978
By current status											
All survivors	1.00%	4.31%	0.03% (0.56)	0.97 (84.23)	0.952	-0.07% (-2.34)	0.90 (117.57)	0.29 (26.21)	-0.05 (-4.62)	0.09 (10.87)	0.984
All nonsurvivors	0.74%	4.38%	-0.24% (-3.65)	0.97 (63.12)	0.917	-0.33% (-7.42)	0.88 (74.96)	0.37 (21.05)	-0.07 (-3.11)	0.09 (6.30)	0.966
Survivors - all funds	0.08%	0.25%	0.08% (6.39)	-0.01 (-2.12)	0.009	0.08% (6.52)	0.00 (0.59)	-0.04 (-9.14)	0.01 (2.85)	0.00 (0.05)	0.208
By fund objective											
Aggressive growth	1.04%	5.16%	-0.01% (-0.06)	1.11 (47.66)	0.863	-0.16% (-3.06)	0.94 (69.62)	0.61 (31.33)	-0.16 (-6.80)	0.18 (11.20)	0.965
Growth and income	0.89%	3.68%	-0.02% (-0.54)	0.84 (109.46)	0.967	-0.12% (-3.92)	0.82 (113.44)	0.13 (11.27)	0.08 (5.39)	0.04 (4.87)	0.978
Long-term growth	0.92%	4.41%	-0.07% (-1.43)	0.99 (83.31)	0.951	-0.13% (-3.99)	0.91 (105.18)	0.28 (21.35)	-0.10 (-6.47)	0.09 (8.25)	0.981

**Table 4**

## Relative Performance of Nonsurviving Mutual Funds in their Final Years

Proportions of nonsurviving funds with group-adjusted performance below various performance fractiles of all funds in their final 60 months of operation and the number of nonsurviving funds included in each estimate. Returns are measured over the final 12-, 24-, 36-, 48-, and 60-month periods prior to fund termination, not the final calendar years.

Performance Group	Group adjusted Performance				
	Last year	Last 2 years	Last 3 years	Last 4 years	Last 5 years
Bottom 50%	62.0%	69.1%	75.1%	79.0%	76.8%
Bottom 10%	24.8%	29.8%	32.8%	32.4%	33.0%
Bottom 5%	15.7%	20.4%	21.7%	20.3%	20.8%
Bottom 1%	6.3%	4.8%	5.6%	5.4%	5.9%
Number of Funds	637	598	531	463	409



**Table 5**  
Probit Model of Fund Disappearance

Estimates of probit models that give the probability that a fund disappears in year  $y+1$  as a function of information from years  $y-4$  through  $y$ . A positive coefficient indicates that the probability of disappearance goes *down* as that variable goes *up*.  $E\_TNA_y$  is the log of the fund's total net assets at the end of  $y$  divided by the average total net assets of other funds of its type.  $E\_RET_t$  is the fund's group-adjusted return in year  $t$ .  $E\_NM_t$  is the new money invested in the fund in year  $t$  minus the average new money that year of funds of its type. If there is insufficient data in year  $t$  to calculate either  $E\_RET_t$  or  $E\_NM_t$ , then both variables are set to 0 and  $MISS_t$  is set to 1. Otherwise,  $MISS_t$  is 0. The number of observations is 16777 and the number of fund deaths is 598. The only requirement for a fund's year  $y$  to be included in the regressions is that  $E\_TNA_y$  be available for that fund. Results for two probits are reported. The first allows each of the four lagged return variables to enter separately while the second forces the coefficient on each to be the same (the  $E\_RETLAG$  coefficient). P-values for significant difference from zero are to the right, and the results of hypothesis tests using the log likelihood ratio are reported at the bottom.

Independent Variable	Coefficient	p-value	
<i>Intercept</i>	-2.348	0.000	
$E\_TNA_y$	-0.165	0.000	
$E\_RET_y$	-1.564	0.000	
$E\_RET_{y-1}$	-1.311	0.000	
$E\_RET_{y-2}$	-1.130	0.000	
$E\_RET_{y-3}$	-0.441	0.132	
$E\_RET_{y-4}$	-0.962	0.004	
$E\_NM_y$	-0.001	0.602	
$E\_NM_{y-1}$	-0.008	0.218	
$E\_NM_{y-2}$	-0.004	0.519	
$E\_NM_{y-3}$	-0.011	0.180	
$E\_NM_{y-4}$	-0.014	0.131	
$MISS_y$	0.083	0.213	
$MISS_{y-1}$	0.054	0.507	
$MISS_{y-2}$	-0.117	0.176	
$MISS_{y-3}$	-0.015	0.872	
$MISS_{y-4}$	0.192	0.010	
Hypothesis	df	$\chi^2$ Value	p-value
$E\_RET_{y-1} = E\_RET_{y-2} = E\_RET_{y-3} = E\_RET_{y-4} = 0$	4	64.37	0.000
$E\_RET_{y-1} = E\_RET_{y-2} = E\_RET_{y-3} = E\_RET_{y-4}$	3	5.06	0.167
$(E\_RET_{y-1} + E\_RET_{y-2} + E\_RET_{y-3} + E\_RET_{y-4}) / 4 = E\_RET_y$	1	4.19	0.041

**Table 6**  
Survivor Bias Examples: Proposition 2 and a Counterexample

Two survivor bias examples. The table presents two examples in which  $m = 2$ , a fund's return  $R_t$  can take on the values 1 or -1 with equal probability, and the critical return level is  $b = -1.5$ . The fund economy has a continuum of funds with independent returns. The Counterexample (on the right-hand side) uses the actual attrition rates,  $x_0, x_1, \dots$ , and conditional means, the  $\mu_{i,j}$ 's, obtained from the assumed generating process for fund returns. The Proposition 2 example (on the left-hand side) fixes all the conditional survival probabilities ( $x_1, x_2, \dots$ ) equal to the unconditional survival probability,  $x_0$ , and sets conditional means with the same number of direct cuts equal. Any conditional mean with one direct cut is set equal to  $E[R_t|C_t]$  while any conditional mean with two direct cuts is set equal to  $E[R_{t-1}|C_t, C_{t-1}]$ .

*Proposition 2 Example*

*Counterexample*

Panel A: Survival rate at  $t$  conditioned on  $j$  periods of prior survival,  $x_j$

$j$	0	1	2	3	4
	0.750	0.750	0.750	0.750	0.750

	0	1	2	3	4
	0.750	0.833	0.800	0.812	0.808

Panel B: Conditional mean return at time  $t+1-i$ , with  $j+1$  cuts, the last at time  $t$ ,  $\mu_{i,j} \equiv E[R_{t+1-i}|C_{t-j}, \dots, C_t]$

$i$	$j$	0	1	2	3	4
1		0.333	0.333	0.333	0.333	0.333
2		0.333	0.600	0.600	0.600	0.600
3			0.333	0.600	0.600	0.600
4				0.333	0.600	0.600
5					0.333	0.600
6						0.333

	0	1	2	3	4
	0.333	0.200	0.250	0.231	0.238
	0.333	0.600	0.500	0.538	0.524
		0.200	0.500	0.385	0.429
			0.250	0.538	0.429
				0.231	0.524
					0.238

Panel C: Fraction of funds at time  $T$  that are  $j$  years old prior to the time- $T$  cut,  $w_{j,T}$

$j$	$T$	1	2	3	4
1		0.200	0.200	0.200	0.200
2		0.800	0.200	0.200	0.200
3			0.600	0.150	0.150
4				0.450	0.112
5					0.338
6					

	1	2	3	4
	0.500	0.083	0.208	0.146
	0.000	0.500	0.083	0.208
	0.500	0.000	0.375	0.062
		0.417	0.000	0.312
			0.333	0.000
				0.271

**Table 6 - continued**

*Proposition 2 Example*

*Counterexample*

Panel D: Fraction of time- $T$  survivors (after the time- $T$  cut) that are  $j$  years old,  $\hat{w}_{j,T}^T$

$j$	$T$	1	2	3	4
1		0.250	0.250	0.250	0.250
2		0.750	0.188	0.188	0.188
3			0.562	0.141	0.141
4				0.422	0.105
5					0.316
6					

	1	2	3	4
	0.545	0.105	0.244	0.177
	0.000	0.474	0.073	0.190
	0.455	0.000	0.366	0.063
		0.421	0.000	0.304
			0.317	0.000
				0.266

Panel E: Cross-sectional mean time- $t$  return for the sample of time- $T$  survivors,  $\mu_t^T$  (the elements of Panel E are obtained from the elements of Panels B and D using equation (6))

$t$	$T$	1	2	3	4
1		0.250	0.533	0.533	0.533
2			0.250	0.533	0.533
3				0.250	0.533
4					0.250

	1	2	3	4
	0.091	0.412	0.286	0.333
		0.263	0.548	0.440
			0.171	0.477
				0.215

Panel F: Average mean return for a sample starting at  $t$  and surviving until  $T$ ,  $\bar{\mu}_{T+1-t}^T$  (sample length is  $T+1-t$ )

$t$	$T$	1	2	3	4
1		0.250	0.392	0.439	0.462
2			0.250	0.392	0.439
3				0.250	0.392
4					0.250

	1	2	3	4
	0.091	0.337	0.335	0.366
		0.263	0.360	0.377
			0.171	0.346
				0.215

**Table 7**

## Survivor Bias in Average Performance as a Function of the Sample Period Length: Calibration to the U.S. Mutual Fund Industry

Survivor bias in average performance as a function the sample period length  $k$  for a return generation process calibrated to the U.S. mutual fund industry. For each  $m$ -period attrition rule,  $m \in \{1, 2, 3, 4, 5, 10\}$ , we simulate values for the conditional means  $\mu_{i,j}$  and the attrition 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. and choose the cutoff return level,  $b$ , in one of two ways. The first way, used in Panel A of Table 7, 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. The second way, used in Panels B and C of Table 7, fixes  $b/\sqrt{m}$  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 while Panel C assumes they are all only 1 year old. 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. Finally, for each given sample period length  $k$ , we average the performance estimate across all possible subperiods of length  $k$ . This average is reported (in percentage) for  $k = 1, 2, 3, 4, 5, 10$ , and 30, while the change in this average is reported (in basis points) for  $k$  going from 30 to 31, 31 to 32, 32 to 33 and 33 to 34.

Panel A: Mutual fund industry consists of 100%  $m$ -year-olds at time 1. Scaled cutoff  $b/\sqrt{m}$  is chosen so that the average death rate is 3.5%.

m	Bias (in %) for Sample Length of:							Change in Bias (in bp) increasing Sample Length from:				$b/\sqrt{m}$	Death Rate (in %)
	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 7 - continued**

**Panel B:** Mutual fund industry consists of 100%  $m$ -year-olds at time 1. Scaled cutoff  $b/\sqrt{m}$  is fixed so that the average death rate for  $m=1$  is 3.5%.

m	Bias (in %) for Sample Length of:							Change in Bias (in bp) increasing Sample Length from:				$b/\sqrt{m}$	Death Rate (in %)
	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 of 100% 1-year-olds at time 1. Scaled cutoff  $b/\sqrt{m}$  is fixed so that the average death rate for  $m=1$  is 3.5%.

m	Bias (in %) for Sample Length of:							Change in Bias (in bp) increasing Sample Length from:				$b/\sqrt{m}$	Death Rate (in %)
	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 8**

Estimates of Survivor Bias in Average Performance as a Function of the Mutual Fund Sample Period Length

Mean annual 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 return estimates in the two samples. The table also reports correlation-adjusted standard errors in the estimate of survivor bias, assuming independent and identically distributed annual returns.

Sample Time Length (years)	Number of Samples	Mean Annual Return Estimate		Survivor Bias	Standard Error
		Survivor-biased Sample	Unbiased Sample		
1	34	11.17%	11.00%	0.17%	0.02%
5	30	11.51%	11.08%	0.43%	0.02%
10	25	11.51%	10.80%	0.71%	0.02%
15	20	12.17%	11.25%	0.91%	0.02%
20	15	12.30%	11.25%	1.04%	0.03%
25	10	11.91%	10.80%	1.10%	0.06%
30	5	12.27%	11.08%	1.19%	0.12%
34	1	12.15%	10.91%	1.24%	0.33%

**Table 9**  
The Effects of Survivorship on Persistence Tests

Persistence measures for full, survivor-biased, and look-ahead-biased 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 t-statistic on the HPZ J-shape measures the convexity in the relation between pre- and post-formation portfolio ranks. A reliably negative t-statistic is consistent with spurious performance persistence due to survivorship.

Portfolio Sorting Variable	Decile 1-10		Monthly 4-Factor Model Alphas				Decile 1-10				
	Spread						Spearman		HPZ		
	Mean	Monthly	Decile 1	Decile 10	Spread	t-stat	Test	p-value	Monthly Expense Ratios	Monthly Turnover (Mturn)	J-shape t-stat
<b>A. Full Sample</b>											
1-Year Returns	0.63%	4.52	-0.13%	-0.37%	0.24%	1.79	0.148		-0.04%	0.4%	
5-Year Returns	0.23%	2.09	-0.10%	-0.34%	0.24%	2.06	0.025		-0.06%	-0.2%	
3-Year 4-Factor Alpha	0.36%	5.04	-0.01%	-0.36%	0.36%	4.60	0.000		-0.04%	-1.0%	
<b>B. Survivor-Biased Sample</b>											
1-Year Returns	0.52%	3.93	-0.05%	-0.15%	0.10%	0.84	0.204		-0.01%	1.7%	1.74
5-Year Returns	0.18%	1.85	-0.07%	-0.19%	0.12%	1.15	0.027		-0.04%	1.0%	1.48
3-Year 4-Factor Alpha	0.19%	2.66	0.01%	-0.17%	0.18%	2.30	0.000		-0.02%	-0.4%	2.40
<b>C. Look-Ahead Biased Sample</b>											
1-Year Returns	0.62%	4.44	-0.14%	-0.36%	0.21%	1.60	0.174		-0.04%	0.4%	1.76
5-Year Returns	0.20%	1.84	-0.11%	-0.29%	0.17%	1.34	0.052		-0.04%	0.3%	0.38
3-Year 4-Factor Alpha	0.34%	4.73	0.00%	-0.34%	0.33%	4.07	0.000		-0.03%	-0.7%	1.80

**Table 10**  
The Effects of Survivorship on Cross-Section Regressions

Estimated univariate cross-sectional regressions for each month from July 1966 to December 1995. The dependent variable is the monthly residual, net of expenses, from the 4-factor model, where the factor loadings are estimated on the prior 3 years of monthly returns. The independent variables are expense ratio, turnover, modified turnover (Mturn), the natural log of TNA lagged one year, and maximum load fees lagged one year. Expense ratio is management, administrative, and 12b-1 expenses divided by average TNA. TNA is total net assets. Turnover is the minimum of purchases and sales divided by average TNA. Modified turnover represents reported turnover plus 0.5 times the absolute value of percentage change in portfolio TNA adjusted for investment returns and mergers. Maximum load is the sum of maximum front-end, back-end and deferred sales charges. All estimates are annualized by multiplying the monthly estimates by 12. The reported estimates are time-series averages of monthly cross-sectional regression slope estimates as in Fama and MacBeth (1973). The t-statistics are on the time-series means of the coefficients.

Independent Variables (coefficients x 100)	Full Sample Estimates (t-statistic)	Survivor-biased Sample Estimates (t-statistic)
Expense Ratio (t)	-1.36 (-4.88)	-0.80 (-2.69)
Turnover (t)	-1.36 (-3.57)	-1.14 (-2.70)
Modified Turnover (t) (Mturn)	-1.06 (-2.87)	-0.85 (-2.11)
ln TNA (t-1)	0.16 (2.39)	-0.03 (-0.41)
Maximum Load (t-1)	-0.06 (-2.44)	-0.01 (-0.54)



## FIGURE LEGENDS

**Figure 1.** Final 5-Year Performance on Non-survivors. The figure presents the average annual group-adjusted performance, gross of expense ratios, on nonsurviving funds in each of their five years prior to termination. Group-adjusted performance is the difference between the fund's return and the equal-weighted portfolio of all funds with the same objective in that period. The dotted lines represent two-standard-error boundaries.

**Figure 2.** Survivor Bias as a Function of the Sample Time Length. The figure reports the bias in average annual return estimates when using a survivor-biased sample instead of the complete, as a function of the time-length of the sample. The bias is the average over all possible samples of a given time 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. Final 5-Year Performance on Nonsurvivors

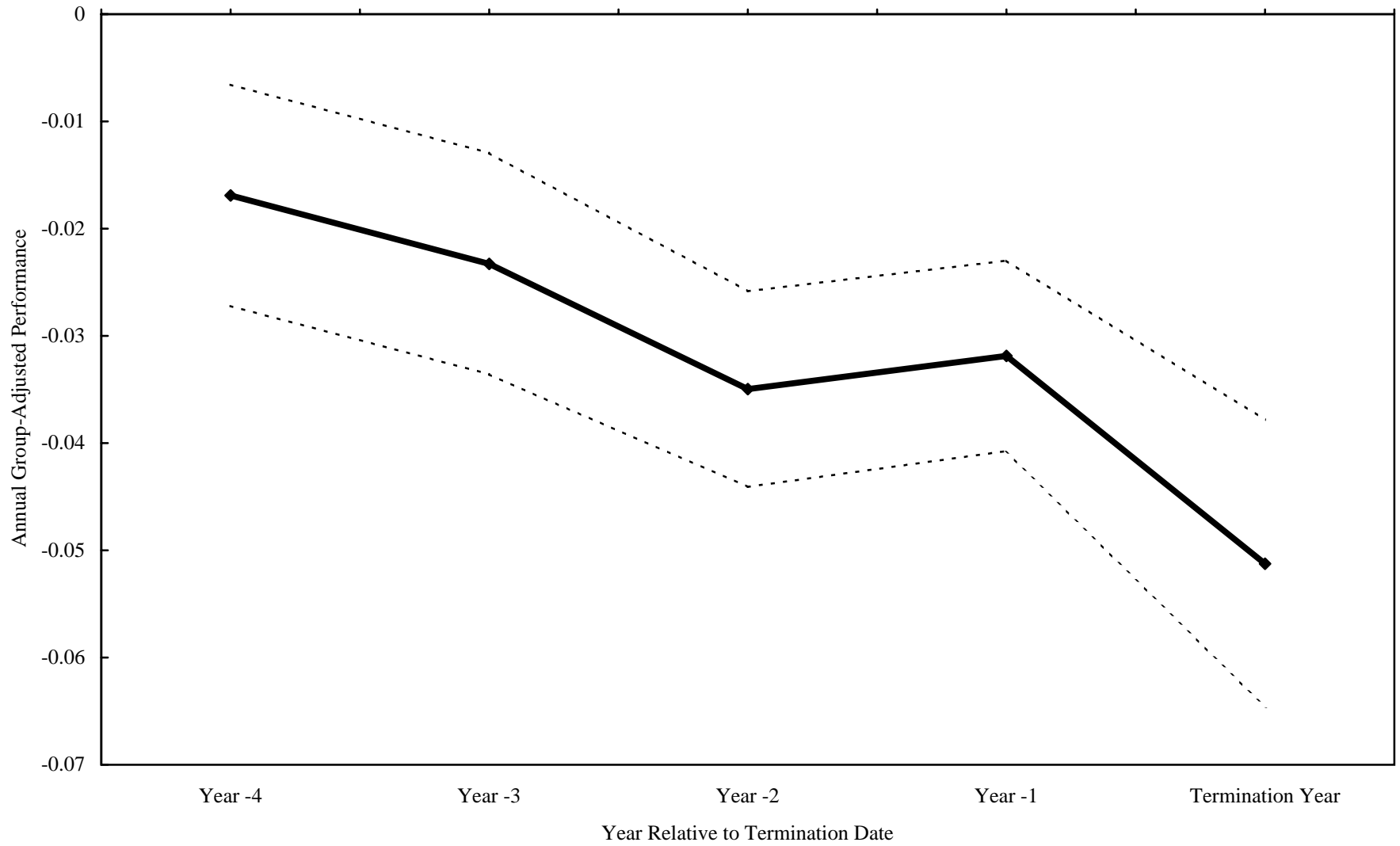


Figure 2. Survivor Bias as a Function of the Sample Time Length

