Journal Of Investment Management, Vol. 16, No. 3, (2018), pp. 46–57

www.joim.com

PICKING THROUGH THE ALPHA GRAVEYARD

Correcting for Survivorship Bias in Investment Product Universes

Gregory C. Allen[a,*], *Ivan S. Cliff*[b] *and Walter J. Meerschaert*[c]

The authors propose a practical technique to correct for survivorship bias across return distributions for investment product universes. The technique is designed to work efficiently in a large-scale performance measurement environment. It uses all available data for survivors and non-survivors, corrects for bias across the full distribution (from 1st to 99th percentile), and can be applied to other return-based statistics such as Sharpe ratio, standard deviation, and correlation. The technique is applied to a variety of product universes over a 10-year period to highlight the practical ways that it can be used to improve the investment decision-making process.

Survivorship bias is the logical error of focusing only on the things that survived a process while ignoring the things that did not. Famous examples abound across fields ranging from the science of falling cats (Whitney and Mehlhaff, 1987), to World War II aircraft design (Mangel and Samaniego, 1984). In the investment business two survivorship-bias-induced logical errors are pervasive and impact the allocation of billions of dollars. The first logical error results in an overly pessimistic view of the relative skill of managers whose strategies have survived over time. The second logical error results in an overly optimistic view of the potential for active investment strategies to outperform benchmarks. Both errors arise from the same mistake—ignoring the truncated track records of non-surviving strategies and making judgments based solely on the complete track records of survivors.

Prior literature on survivorship bias in the investment industry has generally focused either on measuring the magnitude of the bias for fund universes over specific periods to develop generalized estimates (Elton *et al.*, 1996; Fung and

[a]CEO, CRO, Callan LLC, USA.
[b]CFA, Director of Research.
[c]Manager of Information Technology.
[*]Corresponding author.
Callan LLC, 600 Montgomery St., Suite 800, San Francisco, CA 94111, USA. Office: 415.274.3060, Cell: 415.265.0287. E-mail: allen@callan.com

Hsieh, 1997), or on using "bias-free" universes to measure various attributes such as the persistence of performance (Carhart, 1994; Brown and Goetzmann, 1995). Both types of studies speak to the pervasiveness of the problem, and the need for a practical approach that everyday decision-makers can use in real time to correct for survivorship bias when evaluating returns for investment product universes.

In this article, we propose a technique that corrects for survivorship bias across the full distribution of returns (from 1st through 99th percentile). The correction is made by explicitly incorporating all of the data for truncated track records of non-survivors into the determination of the shape of the overall distribution. We then apply this bias correction technique to a variety of investment product universes over a 10-year period to demonstrate its impact. As expected, the correction technique tends to reduce returns across the entire distribution for each universe. In general, the reduction in return is greater for the bottom quartile of the distribution than it is for the top quartile or the median. Finally, we observe a strong positive relationship between the mortality rate of the universe and the magnitude of the survivorship bias correction. Hedge fund universes demonstrated the highest mortality rates over the analysis period, and consequently demonstrated the highest levels of survivorship bias.

1 Survivorship bias in investment management product universes

Before describing the bias correction technique, it is useful to frame the problem using the example of the HFRI Fund Weighted Hedge Fund Composite Universe. Over the 10-year period ended December 31, 2016, there were 6,660 distinct funds in this universe that reported at least one quarterly return. This resulted in a total of 125,798 individual quarterly returns being reported for the period. The industry standard practice for estimating a 10-year return distribution for this universe would include only the returns for funds that had a full 10-year history. Over this period there were 882 funds that met this criterion, reporting a total of 35,280 individual quarterly returns. Thus, the industry standard approach in this example would systematically ignore the 90,518 data points for non-survivors, and estimate the shape of the distribution using only 38% of the available data.

This biased approach is rarely questioned and is employed on a daily basis to drive asset allocation, manager selection, and manager termination decisions that impact billions of dollars. Arguably, some of the most relevant information contained in a manager universe resides in the track records of strategies that did not survive, yet the industry systematically ignores this information, effectively assuming that any capital invested in these strategies never existed. Finding a way to incorporate this information into a less biased estimate of actual investor experience can only result in better decisions and better investment outcomes over time. This is the promise of the technique discussed in this article.

2 Survivorship bias correction technique

For purposes of exposition we refer to the survivorship bias correction technique described in this article as "SUBICO". In simple terms, the SUBICO technique assumes that upon the termination of a strategy's track record, the capital that was invested in that strategy is reallocated equally across all of the available strategies in the universe at the time of the termination. This includes both surviving strategies and any new strategies created since the beginning of the measurement period. The result is a distribution of track records for each terminated strategy which are the same up until the termination date but then diverge post-termination based on the performance patterns of the available replacement strategies on the termination date. Each of these concatenated track records is then assigned the appropriate weight within the overall distribution such that the sum of their weights equals the weight of a single strategy that survived the entire period.

Intuitively this technique mimics what happens to an investor's capital in the real world. When an investment manager terminates a strategy the investors are forced to go out and select an alternative. The process typically involves looking at all of the available alternatives in the universe at the time of the termination and selecting the one that seems most suitable to carry on with the remaining capital. The SUBICO technique essentially replicates this process for all potential investors by assigning an equal probability to every possible choice that an investor might make in this situation.[1]

3 Simplified correction technique

In order to understand how the SUBICO technique works it is helpful to build up to it by first discussing a simpler correction technique which has been referenced periodically in the literature. The simpler technique results in a reasonable estimate of survivorship bias at the median for a distribution, but is inferior to the SUBICO technique in a number of other important dimensions which we discuss below.

Strategy	03/14	06/14	09/14	12/14	03/15	06/15	09/15	12/15	03/16	06/16	09/16	12/16	Annual Return
101	11.8	9.7	0.9	1.0	-5.5	2.6	8.6	4.4	-18.3	2.9			
102	7.0	8.1	-1.3	3.5	-2.0	10.6	2.7	1.7	-11.7	3.4	2.8	4.2	**9.3**
103	-4.5	7.0											
104	7.6	13.2	2.3	10.8	-9.6	-4.0	3.8	-0.3	-11.0	-1.0			
105	14.0	11.4	4.0	3.1	-6.5	10.1	2.1	-1.6	-9.7	4.1	6.3	5.1	**13.9**
106	13.0	11.4	3.6	1.1	-4.3	7.2	6.9	2.7	-12.5	3.7	-5.9	0.0	**8.3**
107						3.8	7.1	-0.1	-9.5	4.2	0.6	0.3	
108	12.8	11.5	-0.6	-0.7	2.0	7.5	3.9	2.2	-13.2	4.5	-5.4	2.8	**8.4**
109	11.1	12.3	3.7	-4.0	-2.0	11.6	7.5	2.5	-10.3	3.0	2.7	4.3	**14.0**
110	10.6	8.4	-0.6	-1.3									
111	11.0	10.5	-1.0	2.2	-3.3	5.9	5.5	0.8	-10.6	1.8	-2.2	3.5	**7.6**
112	9.0	7.8	1.2	-1.1	-4.4	8.8	7.9	-1.4	-11.4	6.2	-2.4	1.9	**6.7**
113	10.1	10.6	0.0	2.2	-2.2	8.3	4.6	1.2	-13.0	6.8	-0.1	1.2	**9.4**
114	19.1	12.5	1.9	0.4	-5.5	7.8	7.4	2.9	-15.5	-1.4	-10.1	2.5	**5.8**
115	12.0	8.5	1.7	0.9									
116	10.5	8.4	1.4	2.5	-9.1	7.0	2.1	-0.3	-11.4	4.1	-3.6	0.9	**3.4**
117							5.5	-0.1	-13.9	4.0	-0.4	4.1	
118	12.0	1.4	1.4	4.7	-4.5	6.2	3.4	-0.3	-7.9	-1.8	-3.6	0.3	**3.3**
119	8.9	9.0	2.8	3.5	-7.9	8.3	2.9	-0.4	-9.5	3.8	3.5	2.6	**8.9**
120	7.9												
121	9.7	-2.0	4.7	5.1	-10.5								
122	7.4	9.5	4.4	5.4	-7.8	3.2	3.9	-0.9	-15.1	1.1	3.7	1.6	**4.6**
123	11.1	10.6	0.0	-0.5	-5.8	7.1	2.7	-0.7	-9.3	4.6	1.2	4.4	**8.0**
124	7.8	5.3	6.2	5.3	-10.3	-2.5	-4.5	0.6	-10.2				
125	7.7	8.4	3.5	5.0	-1.7	5.9	4.0	-1.5	-8.0	3.6	-2.0	3.2	**9.2**
126	9.6	6.1	0.8	-4.4	-0.4	6.4	3.6	0.0	-7.4	2.9	2.9	2.8	**7.5**
												75th Percentile	9.3
												Median	**8.1**
												25th Percentile	6.5
Count	24	23	22	22	20	20	21	21	21	20	18	18	16
Average	**9.9**	**8.7**	**1.9**	**2.0**	**-5.1**	**6.1**	**4.4**	**0.5**	**-11.4**	**3.0**	**-0.7**	**2.5**	**6.8**

Exhibit 1 Return distribution for stylized small-cap equity universe.

Exhibit 1 details the quarterly returns for a stylized universe of investment products over a three-year period. In this example there are a total of 24 strategies in the universe at the beginning of the measurement period, eight of them terminate before the end of the measurement period, and two new strategies are created along the way (for a total of 26). Of the 26 strategies only 16 have a full 3-year track record. The last column in the exhibit shows the annualized 3-year return for each of these 16 strategies. The industry standard approach for calculating a distribution of returns for this universe over this time period would include only the returns for these 16 strategies (ignoring the data for the other 10). The median return for the universe using this approach would be 8.1%.

The bottom row of Exhibit 1 shows the average return for each quarter for all of the strategies in the universe that had a return in that quarter. Importantly, these averages include both the returns for the non-survivors prior to their termination, as well as the returns for any new strategies. The annualized return of this average return series (6.8% shown in the last row, last column) is lower than the median return of 8.1% for the 16 surviving strategies. The difference between these two measures (roughly 1.3% in this case) is often used to provide an estimate of survivorship bias at the median for an investment product universe (Swensen, 2009).

While useful and unbiased at the median, this approach of linking average returns and comparing them with the median return for survivors does not provide insight into the impact of survivorship bias throughout the rest of the distribution. Common practice has been to assume that the estimated adjustment at the median is constant across the whole distribution. As both of the techniques outlined below will demonstrate, however, this assumption underestimates the impact of non-survivors on the bottom half of the distribution, and thus systematically underestimates the downside risk of implementation.

In Exhibit 2 we employ a simple technique which aims to correct for survivorship bias across the entire distribution (from 1st to 99th percentile) by applying the information contained in the average return series to each of the records for the non-survivors.

Under this technique the missing returns for a non-survivor are filled in with the average return for the other products in the universe for each quarter. Intuitively this can be seen as taking the remaining capital from each of the non-survivors on their termination dates and investing it into an equally-weighted multi-strategy portfolio managed by all of the members of the universe on the date of the termination. As the last column of Exhibit 2 illustrates, this concatenation technique results in a 3-year return being calculated for all 24 of the products that were in universe at the beginning of the measurement period.

This technique has the advantage of providing an unbiased estimate of the corrected median (note that the median annual return of the new distribution closely approximates the annual return calculated by linking the quarterly average returns),[2] as well as providing estimates of the survivorship bias impact across the rest of the distribution. In this example the 25th percentile corrected return drops by 80 basis points (from 9.3% to 8.5%), and the 75th percentile corrected return drops by 210 basis points (from 6.5% to 4.4%) relative to the uncorrected distribution shown in Exhibit 1.

While this technique is unbiased at the median, and is computationally quite efficient, it has some notable shortcomings in practice. Foremost of these is the fact that non-surviving funds all end up with the same set of returns for the quarters

Strategy	03/14	06/14	09/14	12/14	03/15	06/15	09/15	12/15	03/16	06/16	09/16	12/16	Annual Return
101	11.8	9.7	0.9	1.0	-5.5	2.6	8.6	4.4	-18.3	2.9	-0.7	2.5	5.6
102	7.0	8.1	-1.3	3.5	-2.0	10.6	2.7	1.7	-11.7	3.4	2.8	4.2	9.3
103	-4.5	7.0	1.9	2.0	-5.1	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	1.4
104	7.6	13.2	2.3	10.8	-9.6	-4.0	3.8	-0.3	-11.0	-1.0	-0.7	2.5	3.6
105	14.0	11.4	4.0	3.1	-6.5	10.1	2.1	-1.6	-9.7	4.1	6.3	5.1	13.9
106	13.0	11.4	3.6	1.1	-4.3	7.2	6.9	2.7	-12.5	3.7	-5.9	0.0	8.3
107						3.8	7.1	-0.1	-9.5	4.2	0.6	0.3	
108	12.8	11.5	-0.6	-0.7	2.0	7.5	3.9	2.2	-13.2	4.5	-5.4	2.8	8.4
109	11.1	12.3	3.7	-4.0	-2.0	11.6	7.5	2.5	-10.3	3.0	2.7	4.3	14.0
110	10.6	8.4	-0.6	-1.3	-5.1	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	4.9
111	11.0	10.5	-1.0	2.2	-3.3	5.9	5.5	0.8	-10.6	1.8	-2.2	3.5	7.6
112	9.0	7.8	1.2	0.1	-4.4	8.8	7.9	-1.4	-11.4	6.2	-2.4	1.9	7.2
113	10.1	10.6	0.0	2.2	-2.2	8.3	4.6	1.2	-13.0	6.8	-0.1	1.2	9.4
114	19.1	12.5	1.9	0.4	-5.5	7.8	7.4	2.9	-15.5	-1.4	-10.1	2.5	5.8
115	12.0	7.5	1.7	0.9	-5.1	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	6.7
116	10.5	8.4	1.4	2.5	-9.1	7.0	2.1	-0.3	-11.4	4.1	-3.6	0.9	3.4
117							5.5	-0.1	-13.9	4.0	-0.4	4.1	
118	12.0	1.4	1.4	4.7	-4.5	6.2	3.4	-0.3	-7.9	-1.8	-3.6	0.3	3.3
119	8.9	9.0	2.8	3.5	-7.9	8.3	2.9	-0.4	-9.5	3.8	3.5	2.6	8.9
120	7.9	8.7	1.9	2.0	-5.1	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	6.2
121	9.7	-2.0	4.7	5.1	-10.5	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	3.1
122	7.4	9.5	4.4	5.4	-7.8	3.2	3.9	-0.9	-15.1	1.1	3.7	1.6	4.6
123	11.1	10.6	0.0	-0.5	-5.8	7.1	2.7	-0.7	-9.3	4.6	1.2	4.4	8.0
124	7.8	5.3	6.2	5.3	-10.3	-2.5	-4.5	0.6	-10.2	3.0	-0.7	2.5	0.1
125	7.7	8.4	3.5	5.0	-1.7	5.9	4.0	-1.5	-8.0	3.6	-2.0	3.2	9.2
126	9.6	6.1	0.8	-4.4	-0.4	6.4	3.6	0.0	-7.4	2.9	2.9	2.8	7.5
75th Percentile													8.5
Median													6.9
25th Percentile													4.4
Count	24	24	24	24	24	25	26	26	26	26	26	26	24
Average	9.9	8.7	1.9	2.0	-5.1	6.1	4.4	0.5	-11.4	3.0	-0.7	2.5	6.8

Exhibit 2 Return distribution showing simple correction technique.

in which they did not survive. To put this problem into perspective, think back to our example using the HFRI Fund Weighted Composite universe. Only 882 of the 4,066 funds who were in the universe at the beginning of the period in that example had a full 10-year record. Using this technique would result in over 3,100 of the funds in the corrected universe having the exact same return pattern by the end of the analysis period. The result, for universes with high mortality rates like this one, is that the corrected distribution tends to collapse inward toward the average return series that is used to fill in the data for the non-survivors.

The second problem with this technique is that the pattern of the average return series over time does not accurately capture the variance and covariance patterns of the individual strategies within the universe. This is because the average return series behaves like a highly diversified portfolio made up multiple strategies meaning that it generally has a higher correlation to the index, a lower standard deviation, and lower tracking error compared with the single strategies that make up the universe. The result is that this simplified technique cannot be used to estimate distributions for return-based statistics that depend on variance or covariance measures such as Sharpe ratio, information ratio, standard deviation, or correlation.

4 The SUBICO technique

The SUBICO (survivorship bias correction) technique is conceptually similar to the simple technique described above, but adds an additional layer of complexity. Rather than assuming that a terminating strategy's capital is invested into a single equal-weighted multi-strategy portfolio (the average), the SUBICO technique assumes that the capital is invested individually into every one of the available strategies at the time of the termination. Exhibits 3 and 4 are helpful in understanding the difference between the two techniques. Both exhibits show the distribution of cumulative return paths for the strategies in our stylized universe. Time is shown on the horizontal axis and cumulative (not annualized) return is

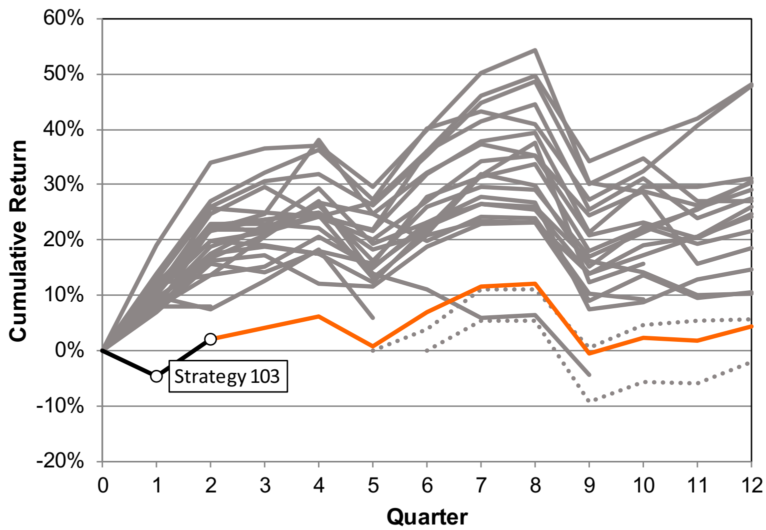


Exhibit 3 Cumulative return distribution — simple correction.

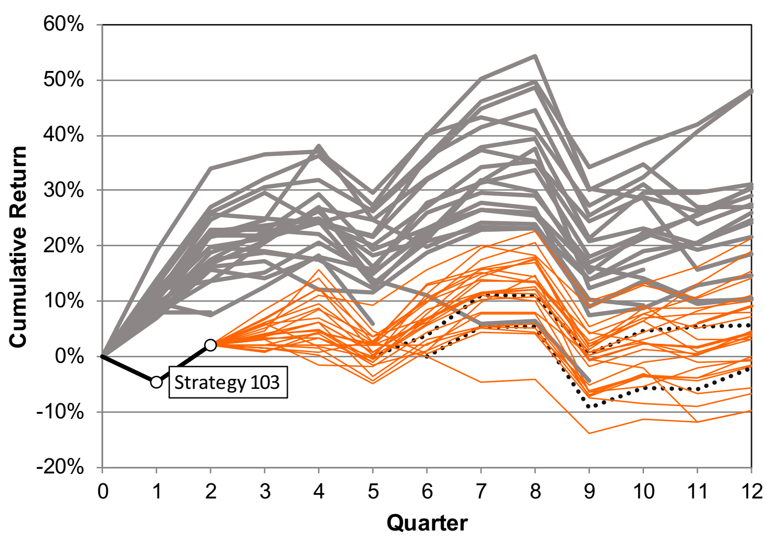


Exhibit 4 Cumulative return distribution — SUBICO correction.

shown on the vertical axis. Strategy 103, whose track record terminates at the end of the second quarter, is highlighted in each exhibit.

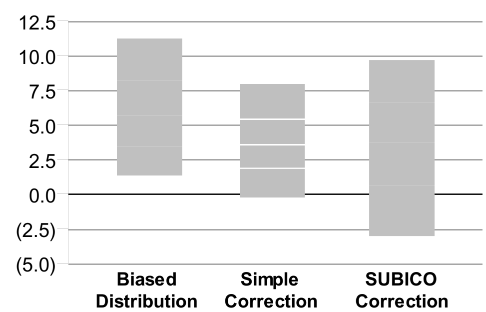
In Exhibit 3 the orange line illustrates the simple correction technique where the quarterly average return series for the distribution is appended to Strategy 103's two-quarter track record to fill in the remainder of the post-termination period. While difficult to verify visually, the variance of the post-termination track record is muted relative to the individual track records of the other strategies in the universe over this period, and its tracking error versus the Russell 2000 index is lower.

In Exhibit 4 we illustrate the SUBICO technique, where instead of simply creating one track record for Strategy 103, we create a distribution of 23 track records (the number of available strategies in the universe at the time of the termination). Each of these individual track records is created by starting with Strategy 103's two-quarter record and appending the track record for one of the available strategies to fill out the rest of the series. Each of these concatenated track records is then given an equal weight in the resulting distribution (1/23rd of the weight of a survivor) such that the sum of all of their weights adds up to the weight of a single survivor.

This technique is applied to every terminating strategy in the universe over the measurement period resulting in an annualized return (or any other return-based statistic) and an associated probability weight for each of the concatenated track records. The returns and probability weights for the concatenated track records are then merged with the returns and probability weights for the survivors in a two-column array (column 1 is return, column 2 is probability weight). The array is sorted from high to low on return and a cumulative probability distribution is calculated for the return column (by summing the probabilities downward from the top of the array to the associated return). Finally, this cumulative probability distribution is used to estimate the 1st through 99th percentile returns. For smaller populations, interpolation and extrapolation techniques are required to solve for the individual percentile values.[3]

In practice the SUBICO technique results in more reasonable distributions than the simple technique (and the biased standard), particularly in the case of high mortality universes. Exhibit 5 examines the distributions of annualized returns for the HFRI Fund Weighted Composite Universe over the 10-year period ended December 31, 2016 using three different methodologies. The first column shows the results using the industry standard approach. The second column applies the simple correction technique, whereas the third column applies the SUBICO technique.

As this exhibit illustrates both of the correction techniques lower the distributions relative to the biased industry standard. The simple correction technique, however, collapses the results towards the center of the distribution, thereby understating the dispersion of possible outcomes for an



	Biased Distribution	Simple Correction	SUBICO Correction
10th Percentile	11.2	7.8	9.7
25th Percentile	8.2	5.4	6.6
Median	**5.7**	**3.8**	**3.8**
75th Percentile	3.4	2.2	0.6
90th Percentile	1.3	(0.1)	(3.0)

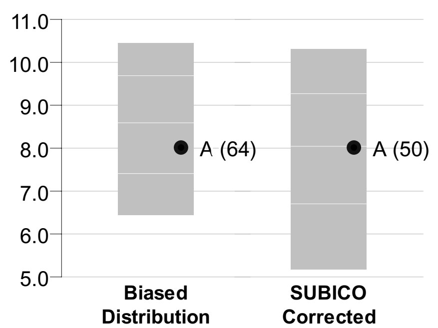
Exhibit 5 Return distributions — HFRI Fund weighted composite universe for 10-year ended December 31, 2016.

investor's capital. This is due to the fact that the concatenated path is the same (the average) for every terminating strategy post-termination. The SUBICO technique captures the dispersion of potential post-termination outcomes and therefore better represents the range of possibilities.

5 Applications of the SUBICO technique—Product rankings

The SUBICO technique allows us to address both of the logical errors referenced at the beginning of this article. Exhibit 6 is designed to illustrate the technique's potential to address the first error, i.e., taking an overly pessimistic view of the relative skill of managers whose strategies have survived over time.

Exhibit 6 shows the rankings of a stylized small-cap product over the 10-year period ended December 31, 2016. The rankings are calculated relative to the Callan Total Institutional Small Cap ("TISC") separate account universe.[4] The left bar shows the ranking of this product using the standard approach. The right bar illustrates the SUBICO-corrected ranking. To help illustrate this particular point, the stylized product was assumed to earn the average return for the TISC universe in each quarter of the measurement period.



	Biased Distribution	SUBICO Corrected
10th Percentile	10.4	10.3
25th Percentile	9.7	9.3
Median	**8.6**	**8.0**
75th Percentile	7.4	6.7
90th Percentile	6.4	5.2
Average Product	**8.0**	**8.0**

Exhibit 6 Return Rankings — Callan total institutional small-cap universe for 10-years ended December 31, 2016.

Using the standard approach, the stylized product ends up ranking in the 64th percentile relative to the products that survived. This is in spite of the fact that it outperformed roughly half of the available products in every quarter during the measurement period (i.e., it generated the average return). When the distribution is corrected to include the data for the non-survivors, the 90th percentile return drops by 120 basis points, the median drops by 60 basis points, and the stylized strategy ends up ranking at the median (a more intuitive result, given its performance pattern).

In the competitive world of institutional investment management, this simple example illustrates how the standard ranking approach might lead decision-makers to draw the wrong conclusions about the relative skill of one of their managers. It is not uncommon for Investment Committees or Boards to conclude that a product that ranks below the 60th percentile after 10-years is a candidate for replacement. Given that roughly 5–6% of institutional equity products drop out of the population per year, this error compounds with each year that is added to the measurement period. The net result when you multiply this error across the entire industry is a downward spiral of unrealistic expectations, leading to higher manager turnover, leading to higher product mortality rates, and leading to even more unrealistic expectations.

Exhibit 7 illustrates the mortality rates for a wide variety of product universes[5] over the 10-year period ended December 31, 2016. The mortality rate is defined as the percentage of products in the starting universe that terminated before the end of the period. For each universe the magnitude of the survivorship bias correction at the

Universe	Mortality Rate	Biased Median Return	Correction at Median	Correction at 90th Percentile	Biased Ranking of Average Product
Separate Account (Callan)					
Large Cap US Equity	55%	7.44%	-0.30%	-0.84%	56
Small Cap US Equity	52%	8.59%	-0.55%	-1.26%	64
International Equity	50%	2.71%	-0.28%	-0.71%	54
Fixed Income	46%	4.90%	-0.02%	-0.05%	51
Mutual Fund (Lipper)					
Large Cap US Equity	42%	6.25%	-0.37%	-0.79%	58
Small Cap US Equity	40%	7.13%	-0.41%	-1.05%	60
International Equity	36%	1.39%	-0.30%	-0.86%	54
Fixed Income	33%	4.34%	-0.12%	-0.40%	52
Hedge Fund (HFRI)					
Total HFRI Universe	78%	5.62%	-1.99%	-4.99%	76
Hedged Equity	74%	5.62%	-1.77%	-4.33%	74
Event Driven	72%	5.51%	-0.91%	-4.06%	74
Global Macro	72%	5.26%	-2.29%	-5.96%	81
Relative Value	78%	6.54%	-1.37%	-2.89%	63
Fund of Funds	77%	3.12%	-1.33%	-1.68%	79

Exhibit 7 Statistics for 10-year ended December 31, 2016.

median and 90th percentile over this period is also shown. The last column shows the ranking that a product that generated the average return for each universe every quarter would receive using the standard approach.

Rankings for the average product range from 51st percentile for the separate account fixed income universe (virtually no survivorship bias) to 79th percentile for the HFRI fund of funds universe (80% mortality rate). The fact that a hedge fund that generated average performance every quarter over this period can rank in the bottom quartile of its respective universe illustrates just how high the penalty can be when biased results are used to analyze relative skill over long periods of time.[6]

Exhibit 7 highlights another interesting fact. In every case the correction at the 90th percentile is greater than the correction at the median. This indicates that the standard ranking technique not only overstates the return at the median, but also tends to understate the magnitude of the downside risk associated with poor (or unlucky) implementation. Tools that both overstate return and understate risk can be especially problematic when used as inputs into an investment decision-making process.

6 Applications of the SUBICO technique—Excess return expectations

Exhibit 8 is designed to illustrate the technique's potential to address the second logical error referenced at the outset of this article, i.e., the tendency to take an overly optimistic view of the potential for active investment strategies to outperform their benchmarks. The first two columns in the table illustrate the biased and corrected median excess returns for each universe. The excess returns are calculated relative to typical benchmarks that institutional investors would use to evaluate the effectiveness of a strategy in each universe. A positive excess return at the median would indicate that more than half of the strategies

Universe	Index	Median Biased Excess Return	Median Corrected Excess Return	Biased Ranking of Index	Corrected Ranking of Index
Separate Account (Callan)					
Large Cap US Equity	Russell 1000	0.36%	0.06%	61	52
Small Cap US Equity	Russell 2000	1.52%	0.97%	80	68
International Equity	MSCI ACWI ex-US	1.75%	1.47%	89	80
Fixed Income	Bloomberg Aggregate	0.56%	0.54%	71	68
Mutual Fund (Lipper)					
Large Cap US Equity	Russell 1000	-0.83%	-1.20%	31	25
Small Cap US Equity	Russell 2000	0.06%	-0.35%	51	42
International Equity	MSCI ACWI ex-US	0.43%	0.13%	61	53
Fixed Income	Bloomberg Aggregate	0.00%	-0.12%	50	47
Hedge Fund (HFRI)					
Total HFRI Universe	LIBOR + 4%	0.41%	-1.58%	60	38
Hedged Equity	LIBOR + 4%	0.41%	-1.36%	61	40
Event Driven	LIBOR + 4%	0.30%	-0.61%	55	41
Global Macro	LIBOR + 4%	0.05%	-2.24%	56	34
Relative Value	LIBOR + 4%	1.33%	-0.04%	70	51
Fund of Funds	LIBOR + 4%	-2.09%	-3.42%	13	7

Exhibit 8 Statistics for 10-year ended December 31, 2016.

in the universe have outperformed their benchmark. The third and fourth columns show the ranking of each benchmark relative to the biased and corrected distributions respectively.

Using hedge funds as an example, the biased distribution might lead one to conclude that LIBOR plus 4% is a reasonable expectation for a hedge fund portfolio over a 10-year period since more 60% of the products in the Total HFRI Fund Weighted Hedge Fund Composite universe (that survived) achieved or beat this hurdle over this period. Adjusting for survivorship bias, however, results in a negative (-1.58%) excess return at the median, and a ranking of 38th percentile for the benchmark, indicating that on an adjusted basis only 38% of hedge funds in this universe would have beaten this hurdle over this period.

Using the results from the corrected distributions would clearly help investors to develop more realistic expectations and make better decisions over time. These results, for example, suggest that it would not be unreasonable for an investor to expect positive excess returns from actively managed small-cap US equity, international equity, and fixed income. This might compel them to focus their active risk budget in these areas. Furthermore, the results suggest that LIBOR $+2.5\%$ might be a more reasonable expectation for a hedge fund program than LIBOR $+4\%$. This might influence an investor's decision as to whether to include or exclude the category in an asset allocation exercise. In both cases starting with more realistic expectations can only foster better investment decisions which in turn can only lead to better investment outcomes over time.

7 Conclusion

The correction technique proposed in this article is a practical way to address survivorship bias in the day-to-day calculations of return distributions for investment product universes used by institutional investors. Because it retains the variance and covariance patterns of the underlying return series, it can also be applied to other important return-based decision variables such as correlation, information ratio, and Sharpe ratio. By quantifying the impact of survivorship bias across the full distribution from the 1st to 99th percentile, the technique allows for a more fair and accurate ranking of surviving products. It also provides valuable insight into the true risk associated with below median outcomes. The adoption of this technique (or something similar) by the industry as a standard would result in better informed fiduciaries and ultimately better investment decisions. It would also create the incentive for the industry to exhume and preserve the records of terminated products allowing us to learn from both the triumphs and the mistakes of the past.

Notes

1 It is important to note that this technique is not designed to replicate observed institutional behavior, but rather to generate an unbiased estimate of a distribution of potential outcomes using all of the available data on survivors and non-survivors. An extension of this approach might be to assign a higher probability of selection to products that meet certain criteria (i.e., better track records, greater assets under management, etc.) used by institutions to select strategies.

2 In small samples, or distributions with significant skewness, the two values can deviate meaningfully. In large samples with non-skewed distributions the two values will generally converge.

3 In practice large universes with high mortality rates can create computational challenges for the SUBICO technique, particularly in a performance measurement production environment where there is a premium on speed. Spawned track records can end up terminating multiple times in high mortality universes. Assuming no limits on computational bandwidth, the ideal approach for dealing with terminating spawned paths is to repeat the spawning process, ultimately creating multiple generations of spawned paths with increasingly tiny associated probabilities. In practice this may be computationally impractical in which case placing a limit on the number of spawns per path is a reasonable way to strike a balance between speed and accuracy. The trade-off between accuracy and computational efficiency can also be managed algorithmically based on variables such as sample size, length of time period, and mortality rate.

4 The Callan separate account universes are used in our analysis because Callan has retained and makes available all of the return data for non-surviving strategies through time.

5 Separate account return data represents self-reported composite returns submitted to the Callan investment manager database by the managers of the products. Managers generally assert GIPS compliance for their return series. Data is not audited. Returns for separate account products are shown gross-of-fees. Mutual fund return data is supplied by Thomson Reuters Lipper. Universe definitions are maintained separately to insure that they continue to include terminated products. Returns for mutual fund products are shown net-of-fees. Returns represent the return for the cheapest share class for that product each quarter. Hedge fund data is supplied by Hedge Fund Research (HFR). Hedge fund universe definitions are maintained separately to insure that they continue to include all terminated products. Returns for hedge fund products are shown net-of-fees.

6 It is interesting to note that mutual fund universes generally have lower mortality rates than separate account universes (roughly 20% lower), and are therefore somewhat less subject to survivorship bias in practice. The primary reason for the lower mortality rates is likely the fact that mutual funds are required by the SEC to report their data as a condition of registration. Managers of separate account products, on the other hand, can opt out of self-reported databases at their discretion.

References

Carhart, M. M. (March 1997). "On Persistence in Mutual Fund Performance," *Journal of Finance* **52**(1), 57–82.

Elton, E. J., Gruber, M. J., and Blake, C. R. (Winter 1996). "Survivorship Bias and Mutual Fund Performance," *The Review of Financial Studies*, Oxford University Press **9**(4), 1097–1120.

Fung, W. and Hsieh, D. A. (Fall 1997). "Survivorship Bias and Investment Style in the Returns of CTAs," *Journal of Portfolio Management* **24**(1), 30–41.

Mangel, M. and Samaniego, F. J. (June 1984). "Abraham Wald's Work on Aircraft Survivability," *Journal of the American Statistical Association* **79**(386), 259–267.

Swensen, D. F. (January 2009). "Pioneering Portfolio Management, an Unconventional Approach to Institutional

Investment," *Business and Economics*, Simon and Schuster, 79–80.
Whitney, W. O. and Mehlhaff, C. J. (1987). "High-Rise Syndrome in Cats," *Journal of the American Veterinary Medical Association* **19**(11), 1399–1403.

Keywords: Survivorship bias; ranking; peer comparison; peer group ranking; peer group distribution; universe ranking; universe distribution; return distribution