SURVIVOR BIAS AND IMPROPER MEASUREMENT: HOW THE MUTUAL FUND INDUSTRY INFLATES ACTIVELY MANAGED FUND PERFORMANCE

Research Paper

March 2006

Sponsored by:

Amy L. Barrett, MBA, CFA, CFP
Brent R. Brodeski, MBA, CPA, CFP, CFA
Savant Capital Management, Inc.
190 Buckley, Dr.
Rockford, IL 61107
(815) 227-0300
abarrett@savantcapital.com
bbrodeski@savantcapital.com

Table of Contents

Executive Summary

The mutual fund industry systematically and significantly overstates fund performance in a way that falsely makes actively managed mutual funds occasionally look competitive with indexes. When the little-understood "survivor bias" factor is taken into account, actively managed mutual funds in **all** nine of the Morningstar Principia® "style boxes" lagged their related indexes from 1995-2004. In all but one of the 42 narrower Morningstar fund categories, the survivor bias effect worked to inflate fund returns. The analysis shows that the purging of the weakest funds from the Morningstar database boosted apparent returns on average by 1.6 percent *per year* over the 10-year period.

"Survivor bias" is a kind of grade inflation for mutual funds that occurs when the funds with the worst performance are made to disappear from the database while strong performers move forward. The result: skewed performance numbers that make the remaining active managers look better since poor performers vanish before they can drag down the overall performance numbers for the indexes.

Very few investors know about survivor bias, but it should be a major concern. For example, over the 10-year period studied, the Mid Blend category returned a whopping cumulative 72 percent **less** than Morningstar data would suggest. The Corporate High Quality Fund category demonstrated the least survivor bias with 0.4% cumulative return difference. The largest evidence of survivor bias exists in the Aggressive Growth Fund category at 116%.

The analysis of 42 of the Morningstar® mutual fund categories included U.S. domestic active equity managers and corresponding indexes for Morningstar's broader "style boxes": Large Growth, Large Blend, Large Value, Mid Growth, Mid Blend, Mid Value, Small Growth, Small Blend and Small Value.   Even when survivor bias is ignored, only the Morningstar Small Growth managers outperformed its index … and, even then, only for the "tech boom" portion (1995-1999) of the study period.

Though earlier researchers have shown that survivor bias exists and that indexing/passive is a superior investment methodology, this analysis is the first to demonstrate conclusively that the main data source (Morningstar) relied upon by individual investors and many financially professionals distorts real historical returns. Unfortunately, the one commercial source of bias-free data does not provide category level returns and is not readily accessible to individual investors. The study also is the first to compare fund manager results to benchmarks custom aligned to specific Morningstar fund categories measured over a 10-year period. The specific period studied (1995-2004) further eliminates most common biases often introduced by style, size, geographic, market cycles, currency and sector factors. Several anomalies are explained by comparing sub-period category data (1995-1999 vs. 2000-2004).

Despite numerous studies that highlight managers' failures, active managers offer half-hearted concessions related to the difficulty of beating "efficient" U.S. large stocks while holding out that managers add value elsewhere — pointing at various "anomalies" purportedly implying that certain market segments remain "inefficient". This study demonstrates that when properly measured—that is, 1) using proper benchmarks, 2) adjusting for survivor bias, and 3) using a time period spanning complete market cycles—nearly all presumed "anomalies" disappear. Furthermore, the few remaining anomalies may be issues of measurement challenge as opposed to manager skill.

Morningstar Principia® Mutual Fund data is available in the CD-Rom format dating back to 1992, with 1994 being the first year for which full style box data is available. The study period covered the 10 years from 1/1/95 to 12/31/04. This period represents a full market cycle, including boom, bust and recovery periods. Researchers compared the data annually, over two five-year periods (1995-1999 and 2000-2004) and the full 10-year period.

Overview: Past Passive Management vs. Active Management Studies

"The Arithmetic of Active Management", written by William F. Sharpe (1991), is a simple three page paper with a powerful punch. Sharpe makes the case that "(1) before fund costs, the return on the average actively managed dollar will be equal the return on the average passively managed dollar and, (2) after costs, the return on the average actively managed dollar will be less than the return on the average passively managed dollar." He "uses only the most rudimentary notions of simple arithmetic" to make his case. He concludes that "These assertions will hold for any time period. Moreover, they depend only on the laws of addition, subtraction, multiplication and division." The rationale is that while any given active manager holding a different set of securities than those of their passively managed counterpart will vary from the market index, in aggregate, the market return must equal a weighted average of all active managers. Thus, some actively managed funds will outperform the index but some will under perform. But, in aggregate, active managers cannot outperform a passive strategy, and after costs, will under perform the passive strategy. He points out that "properly measured, the actively managed dollar must underperform the average passively managed dollar, net of costs. Empirical analyses that appear to refute this principal are guilty of improper measurement."

Despite numerous studies that highlight managers' failure to "beat the market" (including Jensen: 1968, Blake, Elton Gruber: 1993, Mark Carhart: 1997, and James Davis: 2000) at best, investment professionals and the fund industry offer half-hearted concessions related to the difficulty of beating "efficient" U.S. large stocks while holding out that managers add value elsewhere—pointing at various "anomalies" purportedly implying that certain market segments remain "inefficient". Jensen, considered by many to have begun the active vs. passive discussion, said in his 1968 paper: "…[t]he results indicate very little evidence that any *individual* fund was able to earn significantly higher returns than those which could have been expected from random choice." In 2001, Richard Ennis of Ennis, Knupp & Assoc., wrote Small-Cap-Alpha Myth to illustrate that Small Cap funds do not outperform. To make their case, active managers point to selective, and improperly measured data that, on the surface, appears to highlight "kinks in the armor" related to market efficiency. **Sadly, such presumed "anomalies" provide investors a false sense of confidence that the high fees charged by active managers are justified given that such managers intend to exploit supposed "inefficient" areas of the market and well known anomalies with the hope of achieving excess returns (alpha).**

A Closer Look: Arguments Defending Active Management

For example, at the end of the 1990s, active managers frequently cited two apparently "obvious" examples of market inefficiency. Active managers referenced data such as the unadjusted Morningstar U.S. equity style box data as presented in Table One (showing S&P size and style index data minus Morningstar style box fund manager returns—a negative return hints to manager alpha since the manager returns exceed index returns), active managers often conceded the challenge in beating large stock indices but were quick to point out the significant alpha opportunity available in small capitalization growth stocks. This was generally rationalized by pointing out that small stock managers were able to benefit from the lack of good public research (stock analyst coverage) on small stocks. Thus, market mispricings existed and could be unearthed by diligent and hardworking active managers. In addition, small growth stocks also "supposedly" offered a "momentum" affect that good active managers were able to capture. The second oft cited "kink in the armor" was the outperformance of foreign stock managers relative to the MSCI EAFE index. As illustrated in Table Two, raw unadjusted (biased) Morningstar objective data suggested that foreign active managers provided investors 1.9% alpha annually, after fund expenses—obvious manager skill at work. Managers imply such alpha generation due to inefficiency is even more prevalent in less researched foreign markets.

Table One: 1995 - 1999 Index minus the Bias Morningstar Style Box
Data source: Principia disk - Jan. 2000 with returns as of 12/99

	VALUE	BLEND	GROWTH
US LARGE	3.3%	3.0%	4.3%
US MID	2.2%	2.8%	1.4%
US SMALL	3.7%	0.6%	-4.9%

Table Two: 1995 - 1999. MSCI EAFE Index vs. Bias Foreign Funds

	MSCI	Foreign	Difference
1995-1999	13.2%	15.1%	-1.9%

Referencing such examples, active managers regularly leveraged these apparently "obvious" flaws in efficient market theory and Sharpe's reasoning to both attract investors to active mandates for U.S. small growth and foreign stocks as well as provide a further glimmer of hope that other styles of active management might actually be able to defy the simple rules of arithmetic.

We use Sharpe's logic as a framework for our study. Sharpe's simple premise that the sum of all active management must be a losing proposition is validated by our study. *This study demonstrates that when properly measured: 1) using proper benchmarks, 2) adjusting for survivor bias, and 3) using a time period spanning complete market cycles, nearly all presumed anomalies bandied about by active managers and the fund industry disappear.* **We propose that not only were the two previously popular anomalies invalid, but when properly measured, most remaining anomalies can be explained away with logical explanations suggesting improper measurement (as opposed to manager skill).** We agree with Sharpe's assertion that investment professional's conclusions "can only be justified by assuming that the laws of arithmetic have been suspended for the convenience of those who choose to pursue careers as active managers."

Though earlier researchers have routinely demonstrated the failure of active management and existence of survivor bias, this analysis is the first to demonstrate conclusively that the main data source (Morningstar) relied upon by individual investors and many financial professionals distorts real historical returns. Unfortunately, the one commercial source of bias-free data (CRSP Database) does not provide detailed sub-asset class returns and is not readily accessible to individual investors. The study also is the first to compare fund manager results to benchmarks custom-aligned to specific Morningstar fund categories measured over a 10-year period. The specific period studied (1995-2004) further eliminates most common biases often introduced by style, size, geographic, market cycles, currency and sector factors. Several anomalies are explained by comparing sub-period category data (1995-1999 vs. 2000-2004).

Understanding the Role of Survivor Bias

Despite the numerous studies that support Bill Sharpe's passive over active approach to investment management, non-believers point to various anomalies or apparent inconsistencies in index funds' superior returns (such as small growth managers' excess returns at the end of the 1990's) to convince themselves and others that active management adds value. One fault in this argument is that the commonly touted data has been systematically purged of the "badly managed funds." This phenomenon has been dubbed **survivor bias**. Survivor bias occurs when funds with poor performance are wiped out or made to disappear from the database as the result of fund mergers or outright fund liquidations while strong performers continue to exist creating a skewed statistical database. Survivor bias makes it appear as though the poor performers never existed at all.

Several researchers have quantified the effects of survivor bias. Burton Malkiel of Princeton University reviewed Lipper data from 1982 to 1991. During the study period, the average return of

the funds was 17.09% per year. When the database was adjusted for survivor bias, the funds' return was 15.69% per year. Clearly, survivor bias skewed the real story. Malkiel's survivor bias free data shows that the funds underperformed the S&P 500's return (17.52%) by 1.83%.

Mark Carhart conducted the most comprehensive study of survivor bias. He found that actual returns (for capital appreciation funds) to investors were 2% per year below the returns when the data was not adjusted to survivor bias.

Morningstar and Lipper, the two heavyweights in fund databases, acknowledge survivor bias in their products. A study by Lipper Analytical Services showed that the return for all equity funds was 15.87% during the study period. When the returns were adjusted by adding back the dead funds, the return dropped by 1.5%. Lipper's president, Michael Lipper, admits that this is a shortcoming but considers it a minor issue.

Morningstar also conducted a study of survivor bias using their database. They continue telling the investment community that it does not change the average U.S. Stock fund performance at all. We disagree with their conclusion. While it is difficult to precisely quantify the affects of survivor bias given the variance across fund categories and the cumulative nature of survivor bias (it gets worse as the time period of a study expands), a simple average of the 42 categories shows that purging the dead funds from the Morningstar database adds return of 1.6% per year over a ten-year period–see bottom of Table 8 on page 17.

Though this is an estimate, it is indicative of the pervasive nature of survivor bias. Survivor bias is also incredibly consistent across fund categories. In other words, the survivor biased data overstates fund category returns in eight out of the nine "style boxes" (see Table One below for an example of the "style box") and 41 of 42 Morningstar categories. **We suspect the survivor bias effect is even larger over a longer time frame given the cumulative effect of survivor bias as an increasing quantity of funds are purged from historical data, due to mergers and fund liquidations, as the time horizon expands.** We illustrate the phenomenon in the Results section–Survivor Bias and Time Period Results.

A New Look at Mutual Fund Performance: Study Methodology

The data collection process involved identifying the appropriate time period, data source and indices. Morningstar Principia Mutual Fund database was selected as the provider of the mutual fund data. The historical data had been stored in-house. Morningstar data is available in CD-Rom format going back to 1991. 1995 was the first complete calendar year for which data was readily available for both style boxes and many other Morningstar categories. The study period began in 1/1/95 and ended 12/31/04–a ten-year period.

Time Period Selection Process

This period is significant for several reasons. First and foremost, this ten-year period is a reasonably sufficient period to review because it represented a full market cycle. In the late 1990's, the market was very strong and key segments of the market performed extraordinarily well. This euphoric period ended with the "bursting" of the technology "bubble" in early 2000 and a continuing bear market into 2002. Small Cap, Large Cap, Growth, Value, Foreign and Emerging Market asset categories went in and out of style during this ten-year period–meaning demand for these asset classes ebbed and flowed. Additionally, the strength of the dollar relative to competing currencies came fairly close to completing a full cycle, implying that the majority of total return over the entire ten years is primarily from asset return as opposed to currency return. The time period is also significant in that the major world markets went full circle. In the late 1990s, the U.S. and European markets rallied while the Japanese market languished. Recently, Japan has staged a market recovery.

The ten-year period is important in that Morningstar evolved its data collection process. During the period prior to the ten years, mutual fund data was collected for only a handful of mutual fund categories such as Aggressive Growth and Equity Income. These categories were based on prospectus objectives. The Morningstar Principia product developed significantly in 1993 with the addition of the familiar style box (see Table One on page 4 for an example). In 1995, additional asset categories were added to the lineup. Our study covers 42 asset categories. Table Three on page 9 illustrates fund categories studied.

Data Source Selection

The data collection process involved collecting biased, survivor bias free (SBF) and index data. Index data selection is discussed in the next section. The biased and SBF data was taken from the Morningstar Principia mutual fund disks. The difference between the two data sets reflects the fact that the default "biased" data set offered by Morningstar is a "purified" data base that our study contends results in improper measurement of historical manager returns. This means that, by default, all funds that were liquidated or merged into other funds (typically due to poor results) had their returns purged (purified) by Morningstar from the default biased data base. In contrast, the SBF series that we re-constructed still INCLUDES the "dead" funds' returns. Our SBF construction methodology captures and adds back the return of purged funds that typically struggled. For example, had a fund been liquidated or been merged into another fund during July, the dead fund's returns were added back to the SBF data through June. Said differently, default biased data (typically used by investors and advisors) on the 12/31/04 disk does not include all the funds that were liquidated or were merged into another fund during the ten years between 1/1/1995 and 11/30/04. Instead, the default biased data is an average of only those funds that still existed as of 12/31/04.

Identifying Survivor Bias in Data

The biased data collection process involved taking the average mutual fund returns from the 12/31/04 disk for all 42 categories (which includes the nine domestic style boxes). The biased data is available on an annual basis. For example, the Large Growth category had the returns listed below as found on the Principia disk ending 12/04. The appropriate formulas were applied to compound the data.

1995	32.0%
1996	19.7%
1997	27.3%
1998	32.6%
1999	38.3%
2000	-12.2%
2001	-21.2%
2002	-27.4%
2003	28.5%
2004	7.5%

Survivor Bias Free Data Collection Process

Historical quarterly data was pulled from the disks for each of the 40 separate calendar quarters between 3/31/1995 – 12/31/2004. Capturing quarterly fund category returns from each of the 40 different calendar quarter-end disks published during the ten-year period assured that the re-constructed SBF quarterly return series included all existing funds, for each measured fund category, as of the last day in each calendar quarter.

Monthly vs. Quarterly Data: Though mutual fund return data is available from Morningstar on a monthly basis, the study used quarterly return data. Though our quarterly re-construction

methodology still omits returns on funds that died between quarters, our review of a sample of more frequent monthly (versus quarterly) data showed that the return difference is slightly understand, if anything. In the design phase of the study, we compared the survivor bias free results from monthly data versus quarterly data in a single year: 2001. 2001 was selected because it followed the 1999 technology bubble crash. Many equity funds disappeared during the year. Thus, this was the most extreme year during the study for the survivor bias effect. Even so, the monthly versus quarterly analysis showed that the return difference was insignificant. For example, in the Large Growth category, the annualized difference between using monthly and quarterly return data was .04%. Furthermore, the two series—monthly and quarterly–are highly correlated with an R-Squared of .99. We tested 16 different fund categories to discern whether using monthly data would enhance the study. The biggest difference between the monthly and quarterly data was for the emerging markets. This annualized difference was a mere .05%. Since it was apparent that the study would not be materially affected by using quarterly data, we simplified the data collection process by using only quarterly returns. The implication of using quarterly data (that measures returns in approximately 90 day increments) is that the typical liquidated or merged fund likely fell out of the database after 45 days (on average) between quarters, whereas monthly data would have re-captured purged fund returns up until about 15 days prior to each fund's respective death: a difference of 30 days. There are 10 years of data (each year having 4 quarters). Thus, the study consists of 40 quarterly measurement periods. We applied the appropriate cumulative formulas to the quarterly data to construct annual returns. For example, the data collected for the 4 quarters (ending March, June, September and December) of 1995 was compounded to determine the total return for that year.

Data Collection Process: Principia permits the user to design screens for certain mutual fund characteristics. For example, to select (screen) all the funds in the Large Growth category, one can select the appropriate tabs within Principia and a list of all Large
Growth funds appear. The data for this subset of the Morningstar database can then be averaged. This returns a simple average: not a dollar weighted average. Researchers screened the database for 42 mutual fund categories including, but not limited to: Large U.S. Equity funds, Technology funds (as defined by Morningstar), Aggressive Growth funds (as defined by the fund's prospectus) and U.S. Municipal National Short Bond funds. We developed screens for 2 major categories: Morningstar Category and Prospectus Objective. The Morningstar Category, as defined by Morningstar, "identifies funds based on their actual investment styles, as measured by their underlying portfolio holdings (portfolio statistics and compositions over the past three years). In contrast, Prospectus Objective is "based on the wording of the fund's prospectus and is the particular fund's investment objective." Morningstar's database has expanded over the years. Furthermore, additional mutual fund categories are now available that did not exist over the entire 10 year period we studied.

We limited the study to 42 categories because many of the additional categories (now available) were not in existence for the entire ten-year period. (There was one exception to the ten-year rule. We include the emerging markets category although the data was not available for the full period. Emerging data became available beginning in the third quarter 1995. As such, this fund category encompasses a shorter 9 ½ years. Still, we choose to include this category, despite a shorter time horizon, because some emerging markets mutual fund managers aggressively tout their ability to beat the market). In some additional categories not analyzed in our study where data was available for the entire time period, the number of funds in the category was limited. Thus, we eliminated such fund categories from the study. For example, in 1995 the number of funds in the Health category was 19. We eliminated this category from our study due to concern that sample bias could affect results due to the small universe of health funds. For example, had only 2 funds disappeared, it would represent 11% (2/19th) impact on the fund category average return.

Six of the 42 mutual fund categories studied further consisted of combined series of categories. As an illustration, we combined the Large Growth, Large Blend, and Large Value categories to create a Large category. This allowed us to view the return patterns of Large, Mid, and Small stocks. The 3 other blended categories are described in the Index section.

Table Three: Principia Mutual Fund Categories

Morningstar Category	Index Data	A R- Squared	B SBF Cum.	C Biased Cum.	D B-C Cum.	E B-C Annl.
U.S. Equities						
Large Growth	S&P/Barra LG	0.92	142.1%	155.9%	-13.8%	-0.6%
Large Blend	S&P 500	0.98	134.5%	178.4%	-43.9%	-1.9%
Large Value	S&P/Barra LV	0.98	177.5%	202.4%	-24.9%	-1.0%
Mid Growth	S&P/Barra MG	0.73	172.5%	210.2%	-37.7%	-1.4%
Mid Blend	S&P 400	0.91	206.4%	278.6%	-72.2%	-2.4%
Mid Value	S&P/Barra MV	0.94	233.1%	296.4%	-63.3%	-2.0%
Small Growth	S&P/Barra SG	0.74	193.3%	235.7%	-42.4%	-1.5%
Small Blend	S&P 600	0.94	270.2%	256.3%	13.8%	0.4%
Small Value	S&P/Barra SV	0.95	285.7%	323.1%	-37.4%	-1.1%
Large	S&P 500	0.98	159.7%	185.2%	-25.5%	-1.0%
Mid	S&P 400	0.77	190.8%	255.7%	-64.9%	-2.3%
Small	S&P 600	0.81	233.9%	272.8%	-39.0%	-1.3%
Technology	S&P 500 Information Technology	0.78	139.1%	184.9%	-45.8%	-1.9%
Utilities	S&P 500 Utilities	0.55	129.9%	157.9%	-28.0%	-1.3%
International Equities						
Foreign	MSCI EAFE	0.90	77.6%	113.2%	-35.7%	-2.0%
Europe	MSCI Europe	0.93	137.9%	209.5%	-71.6%	-2.9%
Diversified Pacific	MSCI Pacific	0.90	-4.4%	21.5%	-25.9%	-2.4%
Diversified Emerging**	S&P/IFCI Emerging Markets**	0.98	60.8%	76.6%	-15.8%	-1.0%
U.S. Fixed Income						
U.S. Treasury Long	Lehman Brothers Long Treasury	0.96	131.6%	146.1%	-14.5%	-0.7%
U.S. Treasury Intermediate	Lehman Brothers Intermediate Treasury	0.92	85.5%	88.4%	-2.9%	-0.2%
U.S. Treasury Short	Lehman Brothers 1-3 Year Government	0.96	66.4%	70.1%	-3.7%	-0.2%
Municipal National Long	Lehman Brothers Municipal 15 Year	0.99	75.9%	79.8%	-3.9%	-0.2%
Municipal National Intermediate	Lehman Brothers Municipal 7 Year	0.98	69.1%	72.6%	-3.5%	-0.2%
Municipal National Short	Lehman Brothers Municipal 3 Year	0.95	47.8%	49.7%	-1.9%	-0.1%
International Fixed Income						
International Fixed Income	Citigroup Non-dollar World Government	0.71	106.0%	110.6%	-4.6%	-0.2%

Prospective Objective

	Index Data	A R- Squared	B SBF Cum.	C Biased Cum.	D B-C Cum.	E B-C Annl.
U.S. Equities						
Aggressive Growth	Blend (S&P/Barra LG, MG, SG)	0.70	74.1%	190.4%	-116.3%	-5.6%
Growth	Blend (S&P/Barra LG, MG, SG)	0.97	168.1%	209.4%	-41.3%	-1.6%
Growth & Income	S&P 500	0.96	172.0%	199.5%	-27.5%	-1.1%
Equity Income	S&P/Barra LV	0.97	171.4%	198.1%	-26.7%	-1.0%
Small Company	S&P 600	0.67	255.9%	271.2%	-15.3%	-0.5%
U.S. Equities excluding small*	S&P 500	0.94	167.7%	206.1%	-38.4%	-1.5%
Technology	S&P 500 Information Technology	0.79	141.5%	196.7%	-55.2%	-2.3%
Utilities	S&P 500 Utilities	0.54	127.8%	152.8%	-25.1%	-1.1%
International Equities						
Foreign	MSCI EAFE	0.87	74.1%	113.6%	-39.5%	-2.2%
Europe	MSCI Europe	0.93	136.3%	208.5%	-72.2%	-2.9%
Diversified Pacific	MSCI Pacific	0.93	3.3%	29.5%	-26.3%	-2.3%
Diversified Emerging**	S&P/IFCI Emerging Markets**	0.83	64.4%	81.1%	-16.7%	-1.0%
Balanced Fund						
Balanced Fund	Blend (60% S&P/B 500, 40% LB Agg.)	0.95	126.1%	138.4%	-12.3%	-0.6%
U.S. Fixed Income						
U.S. Treasury/ High Quality Corporate	Lehman Brothers Government/Credit	0.98	81.5%	81.9%	-0.5%	0.0%
Corporate High Quality	Lehman Brothers Credit	0.90	79.6%	79.9%	-0.4%	0.0%
Treasury	Lehman Brothers Government	1.00	88.3%	89.6%	-1.3%	-0.1%
Municipal National	Lehman Brothers Municipal	0.99	72.0%	73.3%	-1.3%	-0.1%

* Average of Aggressive Growth, Growth, Growth & Income and Equity Income
** The Diversified Emerging Fund and the S&P/IFCI EM Index data covers the period 6/1/95 - 12/31/04.

SBF = Survivor Bias Free Data
Biased = Bias Data
Cum. = Cumulative
Annl. = Annualized

Index Selection Process

Table Three on the previous page identifies the 42 fund categories we studied. In order to illustrate the performance of active fund management relative to a benchmark, appropriate indices needed to be identified. In several cases (i.e. Utilities and International Fixed Income), fund categories do not track published indices as shown by the low R-Squares (see Table Three). We elaborate on these special cases in the potential variances (anomalies) discussion section. However, in most cases, indices were selected (and in some cases blended) because they were industry standard indices (i.e. the S&P 500 is a standard measure of the U.S. Large Blend category) and their correlation relative to the fund category was high. Table Three illustrates the index we assigned to each fund category and the R-Squared of each such index relative to the fund category.

Standard and Poor's (S&P), Morgan Stanley Capital Markets (MSCI) and Lehman Brothers (LB) are major providers of index data. Their index data is readily available. Furthermore, their index construction methodology results in return series that most industry participants accept as reasonable bogies for domestic and international asset class returns. These three families of indices minimize systematic measurement faults prevalent in certain other indices such as those promoted by Russell. For example, the Russell series is known for reconstitution idiosyncrasies that create an anomaly that some researchers purport to understate the returns measured by the Russell 2000 index relative to the actual domestic small stock universe. In 2001, Ananth Madhavan wrote, "The Russell Reconstitution Effect" in which he found that passive index funds suffer in their attempt to rebalance on the day of reconstitution.

We've also made every attempt to avoid data mining the universe of indices that we use for comparison to Morningstar biased and our reconstructed SBF series. Given the thousands of indices now available and their varying construction methodology, we could have make a stronger case for the existence of survivor bias and the failure of active management by creative index selection: intentionally choosing indices that supported our theories. We instead focused on mainstream index providers including S&P (for domestic equities), MSCI (for foreign equities) and LB (for domestic fixed income) to minimize this possibility. For some fund categories, the R-Squared of the index to the SBF data is less than desirable, and in fact, may have had a higher correlation with another index provider. Still, we choose to use these standard indices to remain consistent and objective.

Table 3 contains 3 blended benchmarks. The index design for the Aggressive Growth category involved an equally weighted average of the S&P/Barra Large Growth, Mid Growth and Small Growth. The methodology is similar for the Growth category. The final blended index represents the Balanced fund category. In order to create the index, we researched the stock to bond ratio of the balanced mutual funds at the end of each calendar year during the study period (1995-2004). We averaged the ratios and determined that the historical Balanced ratio is closest to a 60% stock & 40% bond ratio.

Table Three on page 9 identifies the study's 42 asset classes and the assigned index benchmarks. In most cases, the index benchmark's R-Squared with the data series is high–close to 1. However, in ten cases the R-Squared is less than 0.8. In 8 of the 10 cases, the R-Squared fell between .7 and .79. The two remaining cases, the R-Squares were in the mid-50 range.

Key Findings: Scope of Improper Measurement

In all studies that seek to measure investment performance relative to a benchmark, the outcome is directly dependent on the benchmark selection process. A valid study carefully selects benchmarks that reflect the risk and return nature of the asset classes from which the managers select securities. Many asset classes have standard, well-accepted benchmarks. However, there are exceptions to this general rule. This study confirms this. In the U.S. Large Equity categories (Growth, Blend and Value) standard S&P indices, closely correspond to the index. For example, the R-Squared of the Large Growth group relative to the S&P/Barra 500 Growth Index was .92 during the study period. Table Three lists the asset classes, assigned indices and respective R-Squares.

We identified 10 out of 42 cases where the benchmark index we selected was at least somewhat dissimilar from the corresponding fund categories. This is evidenced by such category's low R-Squares relative to their respective index. These categories included Morningstar U.S. Mid Growth and Small Growth, Morningstar and Prospectus Objective Utilities, Morningstar International Fixed Income, Morningstar and Prospectus Objective Technology, Prospectus Objective Aggressive Growth, Prospectus Objective U.S. Small Company and Morningstar Foreign. For these ten categories, it is important to understand the variances between the manager's strategies and the indices in order to draw accurate conclusions related to the level of manager generated alpha. Simple comparison of these ten fund categories to the indices risks improper measurement and/or drawing inaccurate or incomplete conclusions.

For example, the U.S. Mid Growth and Small Growth categories had R-Squares of .73 and .74 respectively relative to the S&P/Barra Mid Growth and Small Growth Indices. Though the R-Squares do not imply series returns are unrelated, the data patterns clearly vary in the short-term. This becomes obvious when comparing year-by-year index and manager data. During 1999, Mid Growth funds returned 61% whereas the index returned 28.7%: a 32.3% difference. Small Growth funds also returned 61% during 1999: 41.5% greater than the index. The magnitude of the difference implies that during 1999, mutual fund managers were not selecting the same assets contained within the indices, and/or were weighting them substantially different than they were represented in the corresponding indices. Continuing with our example, 1999 was the most extreme year for the technology "bubble".   Investors paid abnormally high prices for technology stocks in hopes of riding the wave of the telecom and internet craze. Many of the IPOs that fund managers purchased were not even yet included in the indices. And while this gave active managers a temporary positive boost during this extreme year, over the longer ten-year time frame, Mid and Small Growth category managers failed to outperform the indices. While managers briefly rode the technology wave by varying from the index holdings, this came back to haunt them as the bubble burst in 2000-2002 causing them to trail the index by similar amounts. For example, in 2000 Small Growth managers underperformed the index by 7%. This pattern continued in 2001 and 2002.

There were three noteworthy asset classes that fail to track a benchmark. The Morningstar Category Utilities, the Prospectus Objective Utilities and International Bond fund categories return patterns varied considerably from their corresponding indices. The Utilities fund category is composed of a small number of mutual funds, because the number of utility stocks is limited. Utility companies are large in size, because of the enormous cost of starting and running a utility company and related regulatory issues. Thus, utility mutual fund managers pick stocks outside the utility sector and accordingly hold a dissimilar set of assets than the S&P 500 Utilities Sector Index. For example, the S&P Index holds the major utilities companies such as American Electric Power and Peoples Energy. While utility funds frequently hold these securities, they typically do so in weightings that materially and systematically vary from the comparative index. Also, it is important to note that the average utility mutual fund, as of 12/31/04, held 42.9% of its assets outside the utility sector. These stocks are not part of the S&P 500 Utilities index. For example, some utility funds currently own stocks in the telecom sector and other non-utility stocks. Therefore, it is not surprising that the index and the

11

average utility mutual fund have dissimilar return patterns resulting in low correlation of the utility fund categories relative to the benchmark index. Thus, a comparison of utility fund categories to index returns, for purposes of determining manager skill (alpha), may lead to improper measurement.

The International Fixed Income category has a low R-Squared of .71. This category has relatively few funds focusing on the asset class. There is also not a very good index with which to benchmark this category. While we believe the index we chose, the Citigroup Non-$ World Government Index, is the best possible fit, even this index does not benchmark well with the category. The index, which began in 1985, is based on a country mix that continues to change over time. For example, Ireland was added in 1997 and Greece joined in 2000. Furthermore, the index is a market-capitalization weighted index that tracks sovereign debt issued in domestic markets in local currency. However, the average International Fixed Income mutual fund substantially deviates from the index. It varies by country exposure, security weightings, currency exposure, and ratio of corporate to sovereign debt. For these reasons, the International Fixed Income category does not closely resemble or benchmark against any index well. Thus, we dismiss the relevancy of any benchmark comparisons for this fund category.

Benchmark index returns were likewise not highly correlated with the Morningstar Category Technology, Prospective Objective Technology, Aggressive Growth, Small Company and Foreign fund categories. We explore these fund categories and provide further perspective as to why they are difficult to benchmark in the variances (anomalies) section.

Survivor Bias Results

In all but 1 of the 42 categories, survivor bias existed over the ten-year period. For those 41 categories, the <u>cumulative</u> return difference–return of the reconstructed SBF data minus the return of the biased data ranged—from -0. 4 % to -116% (see Table 3). *A negative difference implies that the cumulative return for the improperly measured biased Morningstar data exceeds the reconstructed SBF data which includes all funds that existed over the entire ten-year period).* The Corporate High Quality Bond category demonstrated the least survivor bias with -0.4%. The largest evidence of survivor bias exists in the Aggressive Growth Fund category. Table 4 illustrates the annualized and cumulative difference between the biased results and the SBF results for the 9 "style boxes." (Table 3 illustrates the survivor bias for all fund categories.)

Table 4A: Annualized Survivor Bias-Free returns minus Survivor Biased Returns
(1/1/95 - 12/31/04). *The results below represent the annualized ten-year reconstructed SBF data minus the annualized ten-year biased Morningstar data.*

	VALUE	BLEND	GROWTH
US LARGE	-1.0%	-1.9%	-0.6%
US MID	-2.0%	-2.4%	-1.4%
US SMALL	-1.1%	0.4%	-1.5%

Table 4B: Cumulative Survivor Bias-Free returns minus Survivor Biased Returns
(1/1/95 - 12/31/04) *The results below represent the cumulative ten-year reconstructed SBF data minus the cumulative ten-year biased Morningstar data.*

	VALUE	BLEND	GROWTH
US LARGE	-24.9%	-43.9%	-13.8%
US MID	-63.3%	-72.2%	-37.7%
US SMALL	-37.4%	13.8%	-42.43%

 (A negative number implies that the biased series return is higher than that of the survivor bias free series.)

The 1 of 42 fund categories that provides no evidence of survivor bias (or possibly even negative survivor bias) is the Small Blend Category. The cumulative survivor bias return was +13.8% (+0.4% Annually) which indicates that the funds disappearing from the Morningstar database had returns higher than the Small Blend funds still in existence on 12/31/04. We discuss this apparent anomaly later.

Survivor Bias and Time Periods Results

Table 5 demonstrates the effect of time on survivor bias by illustrating the style box annualized returns for the first five-year period (1/95 - 12/99), the second-five period (1/00 - 12/04) and the entire period. During the first five-year period, in 3 out of 9 style boxes the bias returns beat the SBF returns. Only the Small Blend category showed no evidence of survivor bias during the study period. A negative number implies that fund category return was overstated by purging liquidated or merged funds from the database. For example, the reconstructed SBF Small Growth fund category returned 2.6% less per year than biased Morningstar returns during the first five year period. The significance of analyzing cumulative survivor bias over various time periods is critical. Table 5A shows that the level of survivor bias is affected by market cycles and more consistent patterns develop over time.

Table 5A: Survivor Bias Free Returns minus Biased Returns

First five-year period (1/95 to 12/99)

	VALUE	BLEND	GROWTH
US LARGE	0.1%	-2.4%	0.1%
US MID	-1.4%	-1.7%	-0.7%
US SMALL	-1.5%	0.6 %	-2.6%

Second-five period (1/00 - 12/04)

	VALUE	BLEND	GROWTH
US LARGE	-1.9%	-1.5%	-1.1%
US MID	-2.6%	-3.0%	-2.0%
US SMALL	-0.7%	0.3%	-0.7%

Entire period (1/95 - 12/04)

	VALUE	BLEND	GROWTH
US LARGE	-1.0%	-1.9%	-0.6%
US MID	-2.0%	-2.4%	-1.4%
US SMALL	-1.1%	0.4%	-1.5%

Survivor bias is cumulative in nature—becoming more pronounced as time passes. Our study further confirmed this. As illustrated in Table Five B, Morningstar data published as of 12/99 (January 2000 Principia disk) indicated that the Small Growth category had out-performed its index by 4.9% per year for the previous five years. However, reviewing the same time period with Morningstar data as of 12/04 from the January 2005 disk (after the Morningstar data had been further "purified" via five additional years of mergers and liquidations), Morningstar then reported that the Small Growth category beat the index by 11% per year—same identical period but materially different results (see Table Five C). As illustrated by this extreme example, survivor bias data gets retroactively more pronounced year-after-year. There are several exceptions such as Large Value and Large Blend in which the overstatement is not higher as time passes. However, the overstatement is small (i.e. Large Value: 3.3% - 4.0% = - 0.7%) as compared with the Small Growth's return difference of 6.1% (11% -

4.9 %.) While it was beyond the scope of this study, we suspect that the compression of the survivor bias effect for Large Value and Large Blend, as the time frame expands to ten years is actually a function of underlying funds migrating between style box categories (as time passes) as opposed to fund management choosing to purge the better performing funds during the five years between 2000-2004.

Table Five B: 1995 - 1999 Index minus Bias Morningstar Style Box
Data source: Principia disk - Jan. **2000** with returns as of 12/99

	VALUE	BLEND	GROWTH
US LARGE	3.3%	3.0%	4.3%
US MID	2.2%	2.8%	1.4%
US SMALL	3.7%	0.6%	-4.9%

Table Five C: 1995 - 1999 Index minus Bias Morningstar Style Box
Data source: Principia disk - Jan. **2005** with returns as of 12/04

	VALUE	BLEND	GROWTH
US LARGE	4.0%	3.9%	3.8%
US MID	1.7%	2.3%	-0.6%
US SMALL	2.5%	0.8%	-11.0%

(A negative number implies that the biased series return is higher than that of the survivor bias free series.)

Survivor Bias: Effect on historical returns

Survivor bias exists and is <u>not</u> insignificant as Morningstar would like us to believe. As illustrated in Table 3, in all but one of the 42 categories studied, survivor bias inflated actual returns. Adding back the track records of dead funds substantially reduces category returns making the true picture of fund returns less rosy.

The level of bias is also important. Table 3 demonstrates the magnitude. The fund category that holds the title for the most survivor bias is the Aggressive Growth category. Cumulative biased returns are overstated by a cumulative 116%. Over the ten-year period studied, the Mid Blend fund category returned a whopping cumulative 72% less than Morningstar would have you believe. Annually, this equates to 2.4% annualized.

Survivor bias had the least distorting effect in fixed income categories. During the 10 year period, the Morningstar Category Municipal Bond Long, Intermediate and Short categories resulted in an upward distortion of return by less than 5% each on a cumulative basis. This was also the case in the Prospectus Objective Corporate High Quality and Treasury bond categories. This makes sense. In fact, the minimal existence of survivor bias in fixed income categories (relative to equity categories) is further evidence that the fund industry deliberately purges equity funds with the worst track records. Clearly, fixed income categories <u>should</u> experience minimal impact when poorly performing funds are merged into other funds or disappear from the database. This conclusion is supported by the fact that the fixed income asset class is less volatile than equities. Furthermore, the cross section of returns within various fixed income funds is fairly narrow, as fixed income returns are determined more by duration and credit quality than security selection. As such, poor performing fixed income funds' returns will vary less from average fixed income category returns than poor performing equity fund returns vary from their respective fund category returns. This would suggest that there would be little motivation for fund companies to purge fixed income funds since few deliver embarrassing returns that vary significantly from their peers. For example, during the study period, the largest annual biased Large Cap (fund category) return was 31.7% and the lowest annual return was -28.0% (range of

59.7%) whereas the largest PO Treasury annual return was 13.7% and the lowest annual return was 0.0% (range of 13.7%). Bonds returns were much less volatile.

The one category that failed to show survivor bias is Small Blend. The year-by-year data shows that the survivor bias free data had lower returns than the biased data in 1999, 2000 and 2001. For example, in 1999 the return for the survivor bias free data was 20.8% versus 19.4% for the biased return. At face value, this would suggest that the Small Blend funds that were purged actually performed better than those funds that survived. This is an obvious anomaly that we suspect is merely an improper measurement issue.

One plausible explanation is that the anomaly is related to the technology bubble of the late 1990s. Table 4 illustrates the annualized and cumulative difference between the biased results and the SBF results. Survivor bias is most prevalent in the Mid categories (Growth, Blend and Value). One explanation is that many Small companies grew during the bubble period into mid-cap companies. In a sense, funds that invested in such companies may have "graduated" up a level to Mid in the style box categories. The implication is that formerly small stock funds would have posted outsize returns in the SBF database in the small fund categories during their ascent. As the "bubble" burst, many of the, then mid-sized companies owned by the now Mid funds, dropped dramatically in value. As result, many Mid funds would have fallen back into the Small Blend universe or altogether disappeared. In the SBF data series, the positive returns from the initial appreciation would have been credited to the Small Blend category while the losses related to the bursting of the technology bubble would have been credited to the Mid category. Furthermore, technology and telecom sectors took years to regroup. As such, many Small Growth companies likely moved laterally to the Small Blend category as markets struggled. As the market began to recover again in 2002, these troubled mutual funds that then resided in Small Blend (both those that fell from the Mid to Small and moved laterally from the Small Growth to Small Blend) experienced the wrath of the investors who didn't have the patience to hold the slowly improving funds. The Small Blend funds that then disappeared might have had better returns than average over the entire 10 year period (thus causing the anomaly), but investors didn't care. The funds' recent bad experience, and related redemptions, may have caused fund companies to purge such funds despite reasonable long-term returns.

A future study that involves tracking fund level returns between Mid and Small fund categories in the late 1990s and early 2000s would be required to lend support to our Small Blend survivor bias theory. To the extent that such a study were to confirm that the lack of survivor bias in the Small Blend category is actually a function of improper measurement (as opposed to manager skill possessed by purged funds), it would confirm that, when properly measured, survivor bias existed in all 42 Morningstar and prospectus fund categories during the ten-year period studied.

Impact of Survivor Bias

Who is impacted when data providers publish biased returns? Everyone. Mutual Fund companies often compare their funds' returns to the class average or use peer group rankings (of only surviving peers). Investors often buy actively managed funds on the premise that managers in certain categories appear to add value (prior to adjusting for survivor bias). Investors assume that the average manager in such biased categories must deliver better than index returns based on inaccurate peer (category) group returns. Of course, they buy such active managers under the presumption that such funds will do better than average in the future. When the class average is overstated, the public is makes buying decisions on a false premise. Investors are duped. Mutual Fund companies promoting active management gain assets (and income) by remaining silent on the matter. Data providers publish incomplete data. No one is left unscathed.

Active vs. Passive Management: Where's the Alpha?

As shown in Table 8, over the ten-year period, the annualized return of the mutual funds in 37 of the 42 categories studied underperformed the corresponding index return. The 5 categories that outperformed are Morningstar Category (MC) Utilities, MC World Bonds, Prospectus Objective (PO) Asia, PO Emerging Markets, and PO Utilities. Table 6 illustrates the style box return difference between the index and the survivor bias free data. A **positive** number indicates that the mutual funds underperformed the index. In all categories below, active managers underperformed the indices. For example, the Large Blend managers underperformed the S&P 500 by 3.2% per year.

Table 6: Annualized Index returns minus Survivor Bias Free Returns (1/1/95 - 12/31/04)

	VALUE	BLEND	GROWTH
US LARGE	1.5%	3.2%	2.2%
US MID	4.0%	4.3%	4.7%
US SMALL	1.6%	0.3%	0.2%

Active vs. Passive and Sub-Periods Results

In addition to the entire ten-year period, we further broke the study period down into 2 sub-periods: first five years (1/95 - 12/99) and the second five years (1/00 - 12/04). There were sub-periods in which the active managers outperformed the indices. Using the style box as an example, the data shows that Small Growth managers outperformed the index during the first five year period from 1995-99. However, during the second period, the Small Growth category underperformed the index by 7.6%. For the entire period, all fund groups (including Small Growth) underperformed their indices.

Table 7: Annualized Sub-Period Index returns minus Survivor Bias Free Returns

First Five-Year Period (1/1/95 - 12/31/99)

	VALUE	BLEND	GROWTH
US LARGE	4.0%	6.2%	3.8%
US MID	3.0%	4.1%	0.0%
US SMALL	3.9%	0.2%	-8.4%

Second Five-Year Period (1/1/00 - 12/31/04)

	VALUE	BLEND	GROWTH
US LARGE	-0.6%	0.8%	1.0%
US MID	5.0%	4.4%	8.3%
US SMALL	-0.7%	0.4%	7.6%

Table 8: Index and SBF Data (1995 - 2004)

Morningstar Category	Index Data	A Index Cum.	B SBF Cum.	C A-B Cum.	D A-B Annl.	E 95-99 Annl.	F 2000-04 Annl.
					A-B	**A-B**	**A-B**
U.S. Equities							
Large Growth	S&P/Barra LG	195.3%	142.1%	53.3%	**2.2%**	3.8%	1.0%
Large Blend	S&P 500	212.5%	134.5%	78.0%	**3.2%**	6.2%	0.8%
Large Value	S&P/Barra LV	217.4%	177.5%	39.9%	**1.5%**	4.0%	-0.6%
Mid Growth	S&P/Barra MG	313.3%	172.5%	140.8%	**4.7%**	0.1%	8.3%
Mid Blend	S&P 400	344.9%	206.4%	138.6%	**4.3%**	4.1%	4.4%
Mid Value	S&P/Barra MV	372.8%	233.1%	139.7%	**4.0%**	3.0%	5.0%
Small Growth	S&P/Barra SG	198.5%	193.3%	5.2%	**0.2%**	-8.4%	7.6%
Small Blend	S&P 600	280.2%	270.2%	10.0%	**0.3%**	0.2%	0.4%
Small Value	S&P/Barra SG	343.6%	285.7%	57.9%	**1.6%**	3.9%	-0.7%
Large	S&P 500	212.5%	159.7%	52.8%	**2.1%**	4.1%	0.5%
Mid	S&P 400	344.9%	190.8%	154.1%	**4.8%**	0.9%	8.2%
Small	S&P 600	280.2%	233.9%	46.3%	**1.5%**	-2.5%	5.1%
Technology	S&P 500 Information Technology	240.0%	139.1%	100.8%	**3.9%**	7.2%	1.8%
Utilities	S&P 500 Utilities	119.1%	129.9%	-10.8%	**-0.5%**	-6.1%	4.3%
International Equities							
Foreign	MSCI EAFE	78.1%	77.6%	0.6%	**0.0%**	-1.8%	1.6%
Europe	MSCI Europe	182.2%	137.9%	44.3%	**1.9%**	3.0%	0.9%
Diversified Pacific	MSCI Pacific	-4.2%	-4.4%	0.2%	**0.0%**	-3.5%	3.3%
Diversified Emerging**	S&P/IFCI Emerging Markets**	61.2%	60.8%	0.4%	**0.0%**	-2.6%	2.4%
U.S. Fixed Income							
U.S. Treasury Long	Lehman Brothers Long Treasury	151.5%	131.6%	20.0%	**0.9%**	-0.2%	2.0%
U.S. Treasury Intermediate	Lehman Brothers Intermediate Treasury	89.7%	85.5%	4.3%	**0.2%**	0.4%	0.0%
U.S. Treasury Short	Lehman Brothers 1-3 Year Government	75.6%	66.4%	9.1%	**0.6%**	0.8%	0.3%
Municipal National Long	Lehman Brothers Municipal 15 Year	112.4%	75.9%	36.5%	**2.0%**	2.0%	2.0%
Municipal National Intermediate	Lehman Brothers Municipal 7 Year	87.4%	69.1%	18.3%	**1.1%**	1.1%	1.1%
Municipal National Short	Lehman Brothers Municipal 3 Year	62.5%	47.8%	14.7%	**1.0%**	1.0%	1.0%
International Fixed Income							
International Fixed Income	Citigroup Non-dollar World Government	102.8%	106.0%	-3.2%	**-0.2%**	-1.2%	0.9%
Prospective Objective							
U.S. Equities							
Aggressive Growth	Blend (S&P/Barra LG, MG, SG)	241.3%	74.1%	167.2%	**7.4%**	4.8%	9.3%
Growth	Blend (S&P/Barra LG, MG, SG)	241.3%	168.1%	73.2%	**2.7%**	1.8%	3.4%
Growth & Income	S&P 500	212.5%	172.0%	40.5%	**1.5%**	6.5%	-2.4%
Equity Income	S&P/Barra LV	217.4%	171.4%	46.0%	**1.7%**	5.1%	-1.1%
Small Company	S&P 600	280.2%	255.9%	24.3%	**0.8%**	-5.1%	6.1%
U.S. Equities excluding small*	S&P 500	212.5%	167.7%	44.8%	**1.7%**	5.2%	-1.0%
Technology	S&P 500 Information Technology	240.0%	141.5%	98.5%	**3.8%**	7.3%	1.6%
Utilities	S&P 500 Utilities	119.1%	127.8%	-8.7%	**-0.4%**	-6.1%	4.5%
International Equities							
Foreign	MSCI EAFE	78.1%	74.1%	4.0%	**0.2%**	-1.4%	1.6%
Europe	MSCI Europe	182.2%	136.3%	45.9%	**2.0%**	3.0%	1.0%
Diversified Pacific	MSCI Pacific	-4.2%	3.3%	-7.4%	**-0.7%**	-2.6%	1.0%
Diversified Emerging**	S&P/IFCI Emerging Markets**	61.2%	64.4%	-3.2%	**-0.2%**	-4.3%	3.3%
Balanced Fund							
Balanced Fund	Blend (60% S&P/B 500, 40% LB Agg.)	180.3%	126.1%	54.2%	**2.4%**	4.9%	0.2%
U.S. Fixed Income							
U.S. Treasury/ High Quality Corporate	Lehman Brothers Government/Credit	112.0%	81.5%	30.5%	**1.7%**	1.3%	2.1%
Corporate High Quality	Lehman Brothers Credit	124.2%	79.6%	44.6%	**2.4%**	1.9%	2.8%
Treasury	Lehman Brothers Government	105.2%	88.3%	16.9%	**0.9%**	0.9%	1.0%
Municipal National	Lehman Brothers Municipal	97.7%	72.0%	25.7%	**1.5%**	1.5%	1.5%

* Average of Aggressive Growth, Growth, Growth & Income and Equity Income

** The Diversified Emerging Fund and the S&P/IFCI EM Index data covers the period 6/1/95 - 12/31/04.

Average   **1.6%**

Was Sharpe Right? Do Managers Defy Arithmetic?

Sharpe points out that "properly measured, the actively managed dollar must underperform the average passively managed dollar, net of costs. To lend support to Sharpe's theory, each fund category in the study should trail its proper index by approximately the average fund cost (contained within the category) including expense ratios and other trading costs.

Most Fund Managers Failed to Add Value

Active managers did, in fact, fail to add value when: 1) measured versus appropriate benchmarks, 2) controlled for survivor bias, and 3) measured over a long time period (i.e. ten years) encompassing an entire market cycle. Table 8 lists the annualized underperformance of 42 fund categories vs. their respective benchmarks. In 37 of the 42 asset categories, the funds failed. (Note: A negative number indicates that the active managers failed to keep up with the benchmark.) In the five fund categories where category returns did exceed their benchmark, reasonable explanations, other than manager's skill, exist. Each explanation involves measurement related challenges.

Though the majority of asset categories confirmed that active managers failed to beat their index, this study does not explore the reasons for the underperformance. The likely reasons are active managers' costs including: timing mistakes, cash drag, trading costs, and management fees. To illustrate, we highlight the failure in several fund categories. Within Morningstar Category U.S. Large Equities, the indices beat the fund managers between 1/95 - 12/04 across all styles (Growth, Blend, and Value) by an amount reasonably close to the funds' estimated expenses. The regression results further suggest that the average active mutual fund manager in the Large "style box" categories does not generate excess return.

The European equity funds for both Morningstar Category and Prospective Objective failed to beat the MSCI EAFE Europe index during the ten-year period by approximately 2.0% annually. They failed by an amount consistent with their approximate fund expenses. In order for foreign fund managers to "beat" the market, they must be able to add value through either currency and/or stock selection. Unlike domestic fund strategies, international funds are affected by the exchange rate between the U.S. dollar and the basket of foreign currencies imbedded in the benchmark. Thus, the changing value of the U.S. dollar potentially adds or subtracts return. During the study period, a U.S. dollar investor in the MSCI EAFE Europe index returned 10.9% per year whereas a non-U.S. investor earned 9.6% per year. Thus, during the study period, the U.S. dollar fell relative to the basket of foreign currencies by 1.3%. Still, managers were unable to add value by either managing the weakening of the dollar or via superior stock selection.

All 10 domestic fixed income funds and the Balanced fund category funds failed to beat their indices. Balanced and Corporate High Quality categories performed worst trailing their benchmarks by 2.4% per year.

Potential Variances (Anomalies) from Sharpe's Expected Performance

Sharpe's simple logic suggests that active managers (fund categories), on average, should fail by approximately their expenses. As such, material variance, either too positive (indicating that a given fund category outperformed its benchmark) or too negative (indicating that a fund category failed by a margin larger than the fund category's average total expenses) could cast doubt on Sharpe's theory. That is unless it is demonstrated that excess variance is actually a function of improper measurement. We believe each such variance (anomalies) discussed below offers an explanation not inconsistent with Sharpe's premise.

U.S. Mid Categories

Mid category active funds actually performed worse than one would have expected (failed by more than their estimated costs). During the 10-year period, the annual underperformance for the combined Mid category is almost 5%. Mid Value, Blend and Growth relative manager returns were similarly bad. Though the reason for the large underperformance has yet to be proven, Mid category funds routinely compliment their core mid-cap holdings with both small and large small stocks. Very few funds are 100% pure mid-cap. And since mid-cap stock indices delivered stronger returns than either large or small stocks over the ten-year period, we suspect that Mid category funds' variance from owning pure mid-cap stocks systematically penalized the average Mid category fund (i.e. small and large stocks dumbed down Mid category returns vs. mid-cap indices) resulting in excessive negative alpha. As such, comparing the Mid fund category to pure mid-cap indices may result in improper measurement making the managers appear worse than they really are.

Technology Categories

The Morningstar Category and Prospective Objective Technology sector funds underperformed the index by more than expected. While over the entire ten-year period technology funds were down 3.9% and 3.8% respectively—not too far from expectations–sub-periods reveal a different picture. During the first sub-period (1995-99), both technology categories underperformed the technology index by a very large margin (approximately 7% per year). This was far worse than one would have anticipated given the fund category's average expenses. However, in the latter sub-period, funds trailed index but by a more reasonable 1.7% (approximately in line with average fund costs).

There are two factors that appear to have affected the sub-period variance. First, the set of stocks in the index was not the same as those held by fund managers. The index (S&P 500 Technology index) contained approximately fifty stocks during 2000 (1999's data was not available from Standard & Poor's web site). In contrast, on 12/31/99, there were 104 technology funds holding an average of 82 stocks. The R-Squared quantifies this revelation. The two fund categories had R-Squares of just under 0.80. Though the series are not highly correlated (perfect correlation = 1.0) with the index, they are also not uncorrelated. The less than perfect correlation was caused by the fact that the funds held a different number and different types of securities than those contained in the index.

What type of securities did the fund managers purchase? On 12/31/99, the average technology fund contained only 82% technology stock. Some of the fund managers held as little as 42% of fund assets in technology stocks. The remaining stock was in other sectors (such as services, health and industrial cyclical). In addition, the fund managers often held assets outside the market capitalization range represented in the S&P 500 technology index (which includes large stocks). Of the 104 technology funds, Morningstar categorized the funds as 71 Large Cap, 23 Mid, 2 Small Cap and 8 not fitting one style. However, even the funds that Morningstar categorized as large capitalization (on average) generally also owned a limited number of mid and small capitalization stocks. For example, as the technology bubble grew new companies were offering stock via IPOs on a frequent basis. Many of

these newly formed entities were initially small stocks. At the apex of the technology bubble, technology funds (including those that Morningstar identified as, on average, large capitalization funds) frequently bought such newly minted small and mid-cap technology stocks. As such, fund managers' selection of securities (during a period when the cross section of returns within the market—and even more so in technology stocks) —was abnormally large. This caused fund managers to vary from the index during this abnormal time for technology stocks. It is worth noting that markets cannot be considered "normal" during a period when Morningstar Category Technology funds earned 50% and 144% in 1998 and 1999 respectively, yet the category still substantially trailed the index over the entire period of 1995-99.

The second factor affecting sub-period performance was the prevalent (and in retrospect irrational) "market sentiment" that technology stock prices could only go up. This encouraged fund managers to hoard technology stocks during their initial offerings (IPO). Fund managers paid extraordinary prices for assets within and outside their sectors. In contrast, the index did not include IPO stock offerings immediately upon issue. There is inherently a "lag" between when a stock goes public and market indices incorporate the company. As such, during a "hot" period for IPOs, it would be expected that the index may trail the managers who can immediately purchase the IPOs (and benefit from the frenzy). For example, in 1999, the index's annual return was 79%. However, the average Morningstar Category Technology fund return was 144%. The difference (65%) confirms that the average fund managers bought securities included in the index as well as additional securities (i.e. IPOs and securities from other sectors) that temporarily performed extraordinarily well. Though the managers were rewarded for their speculative (and possibly irrational) behavior during 1999, many suffered the consequences over the next three years as the technology "bubble" burst leaving many of the revered IPOs worthless.

In contrast, during the second half of the study (2000-04), the average technology fund trailed the index by 1.8% per year. Yet, the technology funds' 7% first half underperformance weighed on the overall ten-year period performance. Thus, for the entire ten years, the funds' underperformance was approximately 4%.

Intermediate Treasury Category

The smallest outperformance was in the Intermediate Treasury category (0.2%). We suspect managers failed for the same reason that active equity mutual funds fail: bad timing, higher cost, and so on. Overall, bond fund costs were lower. As such, this generally resulted in less underperformance for bonds relative to indices than in equity categories. Another factor that might contribute to the relatively low fixed income fund underperformance (versus the indices) could be a duration mismatch between the average fund and the indices. The U.S. Treasury Intermediate category would have been expected to fail by its costs. This was typically around 1%. However, this category failed by just 0.2%. The difference may be due to a systematic duration decision by managers to keep the duration of their fund longer than that of the index during a period of declining rates. Over the past ten years, yields fell (prices rose) from 7.1% on 1/95 to 4.3% on 12/04. Fund managers would have been rewarded by maintaining duration longer than that of the index. Of course, in a rising rate environment, the opposite is true.

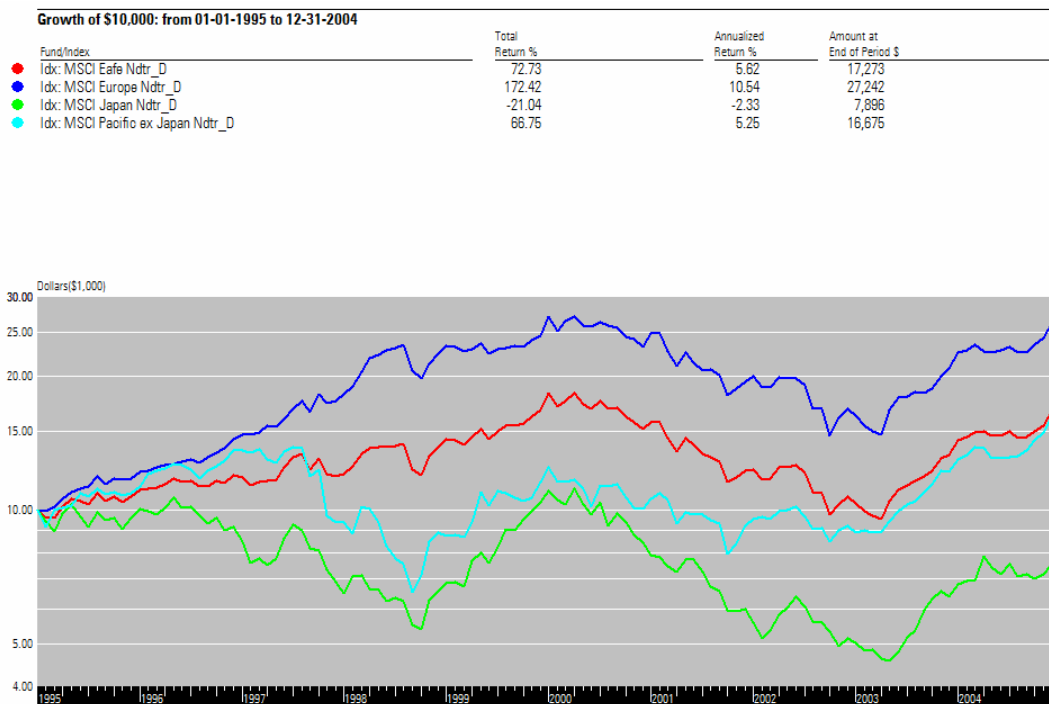### Small Growth and Small Blend Categories

The U.S. Small Growth and Blend funds underperformed their corresponding index but by less than their costs would suggest (0.2% and 0.3% annually, respectively.) The results may imply that assets held by the fund managers were not completely similar to the securities contained in small stock indices. For example, in 1999 the Small Growth managers had an average return of 61% while the index returned 19.6%. Managers seemingly added alpha during this extreme period by looking different from the index. With this wind at their back, during the first five years of the study, these

managers outperformed the index by 8.4%. But, they gave it all back and then some in the five-year period that followed.

The Small Blend category did not have as dramatic year-to-year variance from their index. However, this may be misleading. The survivor bias findings imply an alternative anomaly. Funds in the Small Blend group did not show survivor bias. This indicates that funds that disappeared from the database had better returns than funds that stayed in the database. (See possible explanation in the Survivor Bias Discussion of the Results.) Clearly there appears to be a mis-measurement issue related to small blend. We believe that Small Growth and Blend funds likely performed less poorly than anticipated (failed by less than their expenses) given that most small stock funds benefited from owning a limited amount of mid-cap stocks. During the study's ten-year period, mid-cap indices performed substantially better than small stocks indices. As such, Small fund categories would have benefited from a systematic inclusion of higher returning mid-cap stocks. Another theory relates to the technology bubble. After the technology bubble popped, the Small Blend category may have been the depository for distressed Small Growth stocks that fell out of style but made a strong comeback during the five years following the pop. This would have buoyed up the small blend category relative to the index. Further study is needed to confirm this theory.

Foreign Categories

Morningstar and Prospectus Objective Foreign Fund category returns barely underperformed the MSCI EAFE Index return. The Index includes countries in Europe, Australia and the Far East. The majority of the Index is currently composed of Europe (currently approximately 64%). Japan, is likewise a large contributor to the index (currently approximately 26%). However, Japan's and Europe's allocation in the index changed over time as each region's market capitalization expanded and contracted relative to the other country components. Until recently, the Japanese equity market was in a long-term bear market. At the beginning of the study, Japan had constituted a much larger portion of the index. Its MSCI EAFE weight declined as result of lackluster performance of Japan equities—especially during the first five years of the study. The chart below illustrates the pattern of Japanese vs. EAFE returns.

Growth of $10,000: from 01-01-1995 to 12-31-2004

Fund/Index	Total Return %	Annualized Return %	Amount at End of Period $
Idx: MSCI Eafe Ndtr_D	72.73	5.62	17,273
Idx: MSCI Europe Ndtr_D	172.42	10.54	27,242
Idx: MSCI Japan Ndtr_D	-21.04	-2.33	7,896
Idx: MSCI Pacific ex Japan Ndtr_D	66.75	5.25	16,675

Source: Morningstar Principia

Our study identifies that foreign funds performed less poorly than expected (failed by less than their expenses). While this fact could be used by managers to demonstrate limited skill, we suspect the real explanation is that foreign fund manager's had systematically under-weighted Japanese stocks relative to the EAFE index. This was clearly the case early in the study period when Japan dominated the MSCI EAFE index. Early in the ten-year study period, Japanese stocks struggled as their economy sank. As evidenced by the sub-period returns from 1995-99, foreign funds appear to have been buoyed by the fact that they were systematically underexposed to Japan (relative to the EAFE). As such, during this period, the average foreign fund out-performed the index by 1.8%. However, as time passed, and Japan recovered, foreign funds began to underperform the index. Thus, by the end of the ten-year period, fund categories had underperformed the index—though by less than anticipated

Aggressive Growth Category

The Aggressive Growth categories underperformed the index by a huge 7.4% per year. This would clearly point to a flaw in Sharpe's logic if it were not for an important factor at play: a misfit index. The Aggressive Growth category (like the Small Company categories) have a low R-Squared (0.7) to their index. The index does not fit the data. Yet, we were unable to find a better fit index. Principia software identifies the "best fit" index by measuring a fund's return performance verses all available indices return performances. The index with the highest R-Squared is matched as the best fit. A quick review of the funds in the Aggressive Growth fund category shows that no single index dominates the fund category. In other words, some of the funds track the Russell 1000 while other funds track other indices. We attempted to find an index, or blend of indices that had a higher correlation with the Aggressive Growth fund category but were unsuccessful. The composite index we selected equally weights the S&P 500 Growth, S&P Mid-cap 400 Growth and S&P Small 600 Growth indices.

International Fixed Income Category

The International Fixed Income category outperformed the index, but had a low R-Squared of .71. This category has relatively few funds focusing on the asset class. On again, there is no ideal index with which to benchmark this category. While we believe the index we have chosen, the Citigroup Non-$ World Government Index, is the best possible fit, even this index does not benchmark well with the category. The index, which began in 1985, has a mix of countries that continues to change over time. For example, Ireland was added in 1997 and Greece joined in 2000. Furthermore, the index is a market-capitalization weighted index that tracks sovereign debt issued in domestic markets in local currency. However, International Fixed Income mutual funds substantially deviate from the index. They vary by country exposure, security weightings, and currency exposure, and ratio of corporate to sovereign debt. For these reasons, the average International Fixed Income category does not closely resemble or benchmark against any index well.

Utility Categories

The two Utility fund categories (Morningstar and Prospectus Objective) outperformed their indices. As discussed previously in the Benchmark Section, there is no valid benchmark for either Utility fund category. This is evidenced by the categories low R-Squares relative to the index. This is likely caused by the fund managers within the utility categories investing in assets substantially outside their benchmark. In order to accurately judge whether or not these managers are adding alpha, one would have to construct a custom index that would include non-utility assets. For example, it would be unfair to claim that a balanced fund was superior to an all-stock index during a five-year period in which stocks fell and bonds rose. Balanced fund category outperformance would be due to strong bond market returns and not manager skill. Similarly, Utility category (like International Bond funds) outperformance should likewise be dismissed due to poor benchmark fit.

Emerging Markets Categories

Prospectus Objective Diversified Emerging Market category managers outperformed their index. Over the entire period they outperformed the index by 0.2% per year. We believe the outperformance is once again a function of mis-measurement. First, the asset class was newly developed when the study began. The S&P/IFCI Emerging Market index only began in 1989. As of the start of 1995, the emerging class was not even covered by Morningstar. Emerging market data was first available in mid-1995. And, the initial number of funds in the category was minimal. Even as of 12/99–five-years later, there were only 168 emerging market funds. Furthermore, the true sample size is even fewer because the fund families often have multiple share classes. Dreyfus had a generic, A, B, C and R class of their emerging market fund. Since the sample size (number of funds) was small, the managers had little pressure to conform to an index. Once additional competitors enter an asset class, the industry begins comparing funds to an independent measure or index. Though it appears emerging managers add a slight amount of alpha, a truer measure will be evident as additional competitors enter that asset class and time passes. In addition, there is a mis-match between the index and the investment method commonly used by emerging market fund managers. Because the S&P/IFCI index is dominated by 4 countries (Brazil, Korea, Taiwan, and S. Africa), emerging market funds tend to systematically deviate from the country weightings (based on market capitalization) methodology used by the index. Over the ten-year study period, such a variance tilted emerging market managers towards: 1) smaller frontier countries, and incorporated a bias towards 2) emerging small and 3) emerging value stocks. All three factors (geographic, size and style) added material return in the emerging market asset class during this period. This becomes apparent when reviewing passive (structured) mutual fund strategies available during portions of this period that focused on equal weighting schemes as opposed to market capitalization weights (Dimensional Fund Advisors Emerging Markets I, Dimensional Fund Advisors Emerging Markets Small, Dimensional Fund Advisors Emerging Markets Value and Eaton Vance Tax-Managed Emerging Markets Institutional) . Such passive funds substantially outperformed capitalization weighted indices. To the extent fund managers likewise tended towards equal country weighting, it would be expected that they would outperform the index—despite high expense ratios and trading costs. While it might be argued that the typical fund manager consciously designed their portfolio in this fashion to add value, a more plausible explanation is that they merely got lucky. In the later case, it would be expected that such funds would trail the market capitalization weighted index by a large amount if the larger countries that dominate the benchmark perform best in a future extended period. A proper method to evaluate manager skill would be to reconstruct the index in a similar weighting to the average manager.

Pacific Categories

The largest fund category outperformance relative to the index was for the Prospective Objective Diversified Pacific category. At first glance, it appears that fund managers beat the index by 0.74% per year. However, we believe the outperformance of the Diversified Pacific fund category is due to a mis-measurement phenomenon. As in the discussion above regarding Foreign funds, the Morningstar Category Diversified Pacific category slightly underperformed the index. These two categories– Morningstar Category and the Prospective Objective–are very similar. Morningstar uses it own definitions to place a fund in a category according to its actual holdings. In contrast, Prospective Objective categories include funds based on the fund's prospectus. When Morningstar defined the Pacific fund's category, the Diversified Pacific category underperformed the index. However, when the fund managers defined their objective, the resulting category appears to outperform the index. However, review of the funds contained in the Prospective Objective Pacific category show that such funds do not have a standard asset management philosophy. For example, the 12/99 data (January 2000 Principia disk) shows that the Prospective Objective mutual fund category contains funds that include Japan and other funds that exclude Japan. Japan is the dominate Pacific economy. Funds that exclude Japan can have vastly different returns than funds that include Japan. For example, the five-year annualized return (ending 12/99) for the AIM **Japan** Growth A fund was 14.8%. The AIM Pacific **ex Japan** Growth A fund was -3.9%. During this time frame, the Pacific Rim Countries

(which include China, Australia, S. Korea, Singapore, Taiwan and several smaller countries) performed poorly as compared with Japan. Both these funds are in the Diversified Pacific category. The fact that the Prospective Objective Diversified Pacific category outperforms the index is thus most likely a mis-measurement phenomena rather than an indication of manager alpha.

As one recognizes the unique measurement challenges for each fund category whose performance varies from that of its benchmark minus estimated costs, it becomes apparent that active managers do not defy the laws of mathematics—for any fund category. Rather, most fund categories fail by an anticipated and reasonable amount—especially when using properly aligned benchmarks and after adjusting for survivor bias.

Time Period: Why sufficiently long measurement periods are critical?

This study could not have been conducted effectively any earlier because the breadth of the fund category data was not previously available. As of 12/31/04, there was ten years worth of data that covered a complete economic cycle that consisted of a boom, bubble, bust, and recovery with which to study fund category returns. It is important to note that, given the stage of the market cycle, had more limited data been analyzed to assess manager skill during the previous five years, it is likely that inaccurate conclusions might have been drawn (such as the then apparent anomaly that small growth fund managers seemed to substantially outpace their passive benchmark).

Active Managers often suggest they can beat the market. And, for short time periods, they are often right. As illustrated in Table 8, for a while, it appeared that Small Growth fund managers added value over their index. Table 8 illustrates that for the 1/1/95 to 12/31/99 period, Small Growth fund managers beat the index by 8.4% annually. But, after five additional years passed, Small Growth managers underperformed their benchmark. The same table points out that during the five years between 1/1/2000 - 12/31/2004, Large Value and Small Value fund categories slightly out-performed their corresponding indices. This was after failing miserably to add value during the first five years. The real question is: How were Small Growth managers originally able to beat the index by 8.4% annually? By drilling down into the numbers, we learn that outperformance was mainly attributable to excess returns in one year: 1999. The Index return was 19.6% where the class average was 60.8%: a huge 41.2% difference. Clearly, Small Growth managers strayed from the index holdings. To better understand, recall this year represents the height of the technology bubble. Market returns were dismal after the bubble popped. The S&P 500 index had negative returns for the following three years. During the five-year period after the bubble popped, Small Growth fund managers underperformed the index by 7.6%. Over the entire 10-year period, the category's return underperformed the index return by 0.2% annually. Had an investor jumped into the market in 1999 on the newfound assumption that Small Growth managers were uniquely capable of adding value (based on 1999 excess returns and the previous five-year track record), the investor would have been sorely disappointed.

During 1995 - 1999, there were several other categories in which managers appeared to "beat" a passive strategy. Foreign, Diversified Pacific and Emerging Markets outperformed the index by 1.8%, 3.5% and 2.6% respectively. Why did these categories perform well during this sub-period?

We observe that sub-period fund category performance varies significantly from appropriate benchmarks due to several factors. We can point to at least six possible explanations. First, a five-year time period is far too short to draw conclusion regarding manager's skill—even when analyzing a broad fund category that includes many managers. Markets are affected by short-term economic influences...not long-term trends. For example, an asset class can be affected by demand driven speculation for certain subsets of the securities contained within a fund category's benchmark index (such as technology and telecom IPOs in the late 1990s). In most circumstances, such speculation

ebbs after a period of time. Still, active investors may briefly benefit from their short-term variance from benchmarks favoring such securities with significant momentum.

Secondly, investment styles go in and out of favor. During the technology boom, investors globally clamored for growth investments. After the bust, growth became less appealing. These trends can last quite a while. Investment sectors and equity capitalization also cause sub-period variance. As such, given that the funds included in a given fund category typically do not invest 100% of their assets in the exact securities represented in their benchmark, fund category returns vary from their benchmark during periods when certain investment styles, sectors and sizes sustain a significant advantage (or disadvantage) relative to other asset classes. For example, we observe that Small fund categories would have benefited during the study period from the systematic inclusion of mid-cap stocks since mid-cap stocks outperformed small stocks. However, demand cycles go up and down over longer time frames. Thus, over full market cycles, improper measurement due to short-term style, sector and size biases tend to wash out of fund category data.

The fifth sub-period bias factor is geographic trends. This affects foreign and emerging market fund categories. Foreign benchmarks are typically constructed based on the total market capitalization of the countries contained within the index. Thus, in the late 1980s and early 1990s, Japan dominated foreign benchmarks such as the MSCI EAFE and the MSCI Pacific. Likewise, emerging market indices like the S&P/IFCI are heavily biased towards just few larger emerging markets. To the extent that the average active foreign or emerging market fund manager systematically under or overweight their funds to heavily weighted countries during periods when such countries do materially better or worse than their benchmark, fund category returns can significantly vary from their benchmark during sub-periods (i.e. Foreign and Pacific fund categories led their benchmarks during the late 1990s due to miserable returns in Japan). Lastly, changes in currency valuations (value of the U.S. dollar relative to foreign currencies) tend to add little to no value to foreign investments over the long run since the dollar rises and falls. However, in shorter time horizons, currency exposures often alter fund manager returns relative to un-hedged benchmarks. Given that many foreign fund managers partially hedge their funds, their returns often suffer relative to foreign benchmarks when the dollar is weak yet outperform when the dollar is strong.

In the case of Foreign, Diversified Pacific and Emerging Markets fund categories, the five-year overperformance between 1/1995–12/1999 is likely due to the short sub-period and the biases described above. For example, each of the fund categories soared in 1999 versus the index. In that year alone, foreign funds beat the MSCI EAFE benchmark index 43.6% vs. 27.3%.

The discussion of sub-periods would not be complete without an example illustrating size biases in U.S. Equity fund categories over the study period. During the first five-year period (1995-99), large stocks were "in favor" while small stocks struggled (relative to large stocks). Thus, investors poured money into U.S. Large fund categories. Despite their popularity, large stock fund managers lagged their indices by an unusually large amount over this five year period. For example, the Large Blend category fell behind the S&P 500 index by 6.2% per year. The Growth and Income category (basically large equity funds) underperformed the S&P 500 index by 6.5% annually. For the period, large stock managers failed by more than anticipated.

In perfect contrast, during the first five years, small stocks struggled (relative to large stocks). Still, despite small stocks lagging performance, Small category fund managers beat the small stock indices by huge margins. For example, the Small Growth fund category outperformed their index by 8.4%. Likewise, the Prospectus Objective Small Company category beat its index by 5.1%. As result, many investors began to conclude that, unlike other fund manager categories, small stock managers were able to add value.

Then things reversed during the second sub-period (2000-2004). As time passed, demand for large stocks subsided and small stocks became investors' new "darling." But, even though small stocks

indices began to take the lead (doing better than large stocks), small stock managers substantially failed versus their indices. Small stock managers lost their poster child status as the Small Growth category underperformed its index by 7.6%. The blended Small category trailed the index by 5.1% annually. Prospectus Objective Small Company category fell behind the index by 6.1%. Each failure was by a greater margin than simple arithmetic would have suggested. And, as large stocks struggled, the Large Value fund category actually beat its index by 0.6%—again reversing the trend in the first period.

The key to understanding size biases is to realize that fund managers universally drift from their "target" style. For example, from 1995 - 1999 the average large stock fund manager "snuck" an allocation of mid-cap stocks into their large stock funds. During the first five years, this dragged down Large category returns (i.e. the Mid Growth index trailed the Large Growth index by almost 6% per year). During the same five years, on average, small cap funds owned some mid-cap stocks. This, however, enhanced Small fund category returns since mid-cap stocks did better than small stocks during the first period. For example, the Mid Growth index exceeded the Small Growth index from 1995-99 by 11.6% annually.

The tide turned after 1999. Managers' tendency towards style drift now helped large cap managers (since mid-cap stocks did better than large stocks from 2000-2004). In contrast, it hurt Small fund categories. To put this in perspective, consider that the Small Growth index was up almost 7% per year while the Large Growth index was DOWN 7% per year during the same time. Thus, when the average Large category fund manager added some mid-cap or small stocks, their returns inevitably improved. In contrast, Small fund categories suffered from style drift during the second period as their returns lagged versus their index from the negative influence of most small fund manager's tendencies to include some mid-cap and large stocks in their portfolio.

Although style drift (i.e. size bias) may help managers during sub-periods, over the long run style drift tends to wash out fund category returns. This become evident as you observe the fact that both Large and Small fund categories failed to keep up with their index by 2.0% and 1.5% per year over the entire ten year study. Incidentally, this is approximately the amount of underperformance that Sharpe might have predicted.

Implications to Investors:

1) It is critical to utilize proper benchmarks in evaluating mutual funds and fund categories. Many investors err in their assumption that select managers and fund categories generate alpha by comparing such managers and fund categories to improper benchmarks.

2) For several fund categories, no ideal benchmark, or composite of benchmarks, exist to properly measure manager returns.

3) It is critical that the time period selected to evaluate managers and fund categories is sufficiently long to incorporate enough data to eliminate biases that often lead to improper measurement. Such biases are generally related to style, size, geography, market cycles, currency and sector factors.

4) Survivor bias is pervasive in industry mutual fund data—particularly for equity funds.

5) Given the extent of survivor bias in commonly published data, investors should be wary of the value of peer group, fund objective and category data.

6) Morningstar, the leading source of fund category data to financial professionals and investors, systematically distorts and overstates historical performance of fund categories.

7) Investors should beware of purported anomalies that cast doubt on market efficiency. Such anomalies generally result from improper measurement as opposed to true manager skill.

8) When properly measured, active managers fail in nearly every fund category. This reinforces that investors should have a strong preference for low-cost passive investment strategies.

Appendix: Limits on Data and Recommendations for Further Inquiry

There are six limitations that have the potential of affecting the study outcomes.

1) We developed reconstructed SBF data using quarterly returns. Though monthly returns were available, we chose quarterly data. We studied one year: 2001. In 2001, the 10-year SBF returns for the categories studied were not significantly different we used monthly or quarterly data. It is possible that monthly returns could have had a different outcome than quarterly returns if we had tested a year other than 2001. We looked at only 16 of the 42 categories.

2) The mutual fund category returns reflected average as opposed to dollar weighted returns. For example if there were 100 funds in the technology fund category, Principia averages the return of all 100 funds by weighting them equally despite the fact that certain funds hold far more assets. This equal weighting may produce very different category returns than would dollar weighted returns. An illustration will help explain the point. If in the technology category, 10 funds were very large–90% of the assets in the group—and had a return of 50%, and the remaining 90 funds had a return of 1%, the average return would be 6%. However, the dollar weighted return would be 45%. This is an extreme example but is useful in illustrating that the method of calculating return will influence the category return outcome. Despite this potential limitation, we feel average returns are the appropriate measure for this study given that the largest funds are typically those that historically did well. Equal weighting essentially increases the sample size in that it provides an equal opportunity for each fund, and its manager or team, to contribute alpha to the fund category returns.

3) The biased data is taken from one Morningstar disk: 12/04 data. The biased data certainly would have been different had the study ended on a day other than 12/31/04. The data from 12/03 would not be the same as 12/04.

4) Though we made every attempt to select the most appropriate benchmark index and believe our rationale for selecting families of index providers is reasonable, it is possible that other indices could have caused the study to have different results. The task of picking a benchmark is as much a science as it is an art. The benchmarks were chosen after the period ended and after the fund category results were available. This was unavoidable but should be noted.

5) The study ignores the impact of taxes on mutual fund returns. It is safe to assume that taxes would have lowered actual returns—generally more so for the overall fund categories that includes primarily actively managed funds than for funds that mimic the indices.

6) Other costs associated with active management are ignored. The study did not attempt to estimate loads, surrender fees, brokerage fees assessed on the purchase of funds at fund supermarkets, wrap fees or any other fees that might have been incurred by investors.

Sources and References

Board of Governor of Federal Reserve System - Historical Data.

Carhart, Mark, 1997, On Persistence in Mutual Fund Performance, The Journal of Finance, Vol. LII, No. 1.

Damato, Karen, 1997, Dead-and-Buried Funds Would Alter "Averages" - If Included, WSJ Interactive.

Davis, Jim, 2000, Mutual Fund Performance and Manager Style, Working Paper, Dimensional Fund Advisors, Santa Monica, Ca.

Jensen, Michael, 1968, The Performance of Mutual Funds in the Period 1945 - 1964, Journal of Business.

Ibbotson Associates - Index data

Madhavan, Ananth, 2001, "The Russell Reconstitution Effect", Working Paper, ITG Inc., New York, NY.

Morningstar Inc. - Principia® Mutual Fund data disks.

Morgan Stanley Capital Markets - Index data

The Vanguard Group, 2006, Total Return by Asset Class, Market Review, Valley Forge, Pa.

Van Kampen Funds Inc., Van Kampen Utility Fund, 2005, Oakbrook Terrace, Il.

Sharpe, William, 1991, The Arithmetic of Active Management, The Financial Analyst's Journal, Charlottesville, Va.

Sinquefield, Rex, 1995, Asset Management: Active vs. Passive Management, Lecture, Dimensional Fund Advisors, Santa Monica, Ca.

Swedrowe, Larry, 2001, "Warning: Returns Shown Contain Biases We Are Not Required to Report", St. Louis, Mo.