



ABSTRACT

The purpose of this paper is to investigate and discuss historical active manager performance relative to the performance of an appropriate market benchmark. Although this subject has been written about extensively, much of the analysis has traditionally been plagued with data quality issues. Specifically, survivorship bias, selection bias, and classification noise within manager performance data are present in any analysis that does not account for them, potentially leading to misleading outcomes. Our analysis takes steps to correct these data quality issues where possible. Using this more representative and appropriate data, this paper uses a number of metrics to analyze the relative performance of active managers over an extensive time period. Our analysis evaluates whether managers have added value historically, if there is a difference across asset classes, and if past relative performance has been indicative of future relative performance (i.e., persistence).

MANAGER PERFORMANCE

The question of whether active management adds value has been raging within industry and academia since the inception of the first passive funds in the 1970s. The available literature, produced by the academic community, money managers, and others, does not always come to the same conclusions. This is partly because they evaluate different universes, over different time periods, and using different metrics. Still, the vast majority of the studies focus on U.S. domestic equity mutual funds, as this is the universe where the most data is available.

Kenneth French suggested in a 2008 paper that actively managed funds, in aggregate, are equal to the sum of the market, making active management a zero sum game, before trading costs and fees are applied. This implies that in aggregate, active managers will underperform the market by an amount equal to fees and expenses.¹ This paper seeks, among other things, to test the historical validity of this conclusion. This paper will also endeavor to answer whether the odds of outperformance are high, low, or purely random and whether the amount of value added from active management varies across asset classes, styles, and time.

Data

The dataset we chose for our analysis was MorningStar Direct's global investment database. This is an extensive database, which includes funds that are both 'live' (currently managed funds) and 'dead' (funds which reported at one time but have since ceased reporting, usually because the fund has gone out of business). The fact that the dataset contains both 'live' and 'dead' funds is an important feature that allows this paper to confront the first data quality issue – survivorship bias.

¹ A recent paper by Wermers and Yao (2010) has found large differences in the aggregate holdings of active managers versus the market, which would call this assumption into question. That said, a more recent review of the literature by Jones and Wermers (2011) finds the empirical fact of underperformance roughly amounting to the size of fees and expenses to be supported by the majority of the existing literature, whatever the causal reason.

Survivorship bias occurs when databases exclude the returns of funds that have closed or gone out of business. Longer-term time periods are generally preferable when analyzing performance, to account for market cycles, short-term trends, or other sources of end-point bias. However, the longer the time period, the greater risk there is of survivorship bias impacting the data.

Survivorship bias can cause results to be unrealistically positive, since it is those funds that underperformed their peers that are less likely to survive. Both intuition and past research² imply that the majority of managers who dropped out of a universe were underperforming. Inclusion of both currently available 'live' and 'dead' funds is essential for an accurate picture of manager performance. Survivorship bias is often found in manager return data because the majority of manager return databases do not include 'dead' funds. The existing literature that confronts this issue finds that this bias can be substantial.³ One specific example from Lo (2001) quantifies this bias as increasing the Sharpe ratio⁴ from 0 to 1.16 for a dataset of only 5 funds and this bias only increases as the size of the dataset grows. By including both 'live' and 'dead' funds in our analysis, our results should be free of significant survivorship bias.

Selection bias, which results from including of only those managers who chose to submit data, can be difficult to estimate. For example, a manager may create five products, run them for a few years, and subsequently report the returns of the outperforming ones while closing (and not reporting on) the underperforming products. There are several manager universes with which we are familiar (e.g., eVestment Alliance, PSN) that focus on the institutional investor universe. However, the composites in these universes are self-reporting, which means selection bias, in addition to survivorship bias, is likely present. This is another reason why we chose to use MorningStar Direct's data, as the funds in that universe must report (most even have to strike a daily NAV).

Finally, through the process of researching this paper, a large amount of "noise" could be observed which was caused by the mismatch between funds' strategies and their benchmarks.⁵ For example, within the foreign equity asset class, many managers owned emerging market stocks despite those securities' absence from the benchmark, usually the MSCI EAFE. This represents a mismatch versus the benchmark, as the long-term holding of such ex-benchmark securities results in a portfolio for the manager that possesses different risk-return characteristics than the benchmark. If a significant segment of managers in an asset class run portfolios that are meaningfully different from the benchmark, it can lead to erroneous conclusions.

² See "Mutual Fund Survivorship" by Cahart, et al (2001), and "Survivor Bias and Improper Measurement" by Barrett and Brodeski (2006).

³ For example, McEnally and Ravenscraft (1999); Brown, Goetzmann, and Ibotson (1999); Brown, Goetzmann, and Park (1997); Elton, Gruber, and Blake (1996); Fung and Hsieh (1997); and Schneeweis and Spurgin (1996); Lo (2001).

⁴ The Sharpe Ratio (SR) is defined as $SR = (\text{Return of Manager} - \text{Return of Risk Free Asset}) / \text{Standard Deviation of Manager}$. This is a measure of excess return which is standardized by risk and is a common performance metric.

⁵ We used the Analyst-chosen benchmark within this database as the reference benchmark, unless otherwise noted.

In order to compensate for this issue, a large amount of filtering was done on the dataset. The primary purpose of the filtering was to measure each manager against an appropriate benchmark, thus providing the clearest possible view of manager performance. A detailed explanation of the methodology can be found in Appendix A.

Fees and Expenses

Expenses, fees and trading costs can be a high hurdle for managers to overcome. All of the following results in this paper are before costs. The decision to compare before costs was made so that the benchmark index could be used directly for comparison. Furthermore, fees will vary for different investors. That is, larger institutions investing larger mandates will likely be able to negotiate lower fees than those available to smaller institutions.

When comparing active and passive management, it is important that investors consider the fees they would likely bear, both for active and passive management. Note that even index investing requires investors to bear some costs, albeit at a much lower level.

RESULTS AND METRICS

Broad Asset Classes

The table below shows median manager outperformance, before fees, by asset class using data from as long a period as possible.⁶ For this extended time period, in most asset classes, the median manager outperformed its benchmark, before fees. In several asset classes (high yield and U.S. large cap), the outperformance was quite close to zero. In domestic small cap and emerging market equities, the median outperformance was quite large. In contrast, it was negative for foreign (EAFE) large cap and core bonds.

Asset Class	Median Outperformance
Core Bonds	-13 bp
High Yield	6 bp
Domestic Large Cap	1 bp
Domestic Small Cap	90 bp
Foreign Large Cap	-30 bp
Emerging Markets	93 bp

As noted earlier, fees are variable, and depending on the size of the mandate, negotiable. The following table lists the median fee on different size mandates as provided by investment managers to eVestment Alliance.⁷ When comparing the median outperformance to the median fee for each asset class, the gross outperformance of the median manager has not justified the historical median fee. In other words, it seems that in the asset classes where

⁶ Outperformance is measured as the equal-weighted average of manager return minus the benchmark return for rolling 12-month periods.

⁷ Source: eVestment Alliance database as of December, 2012.

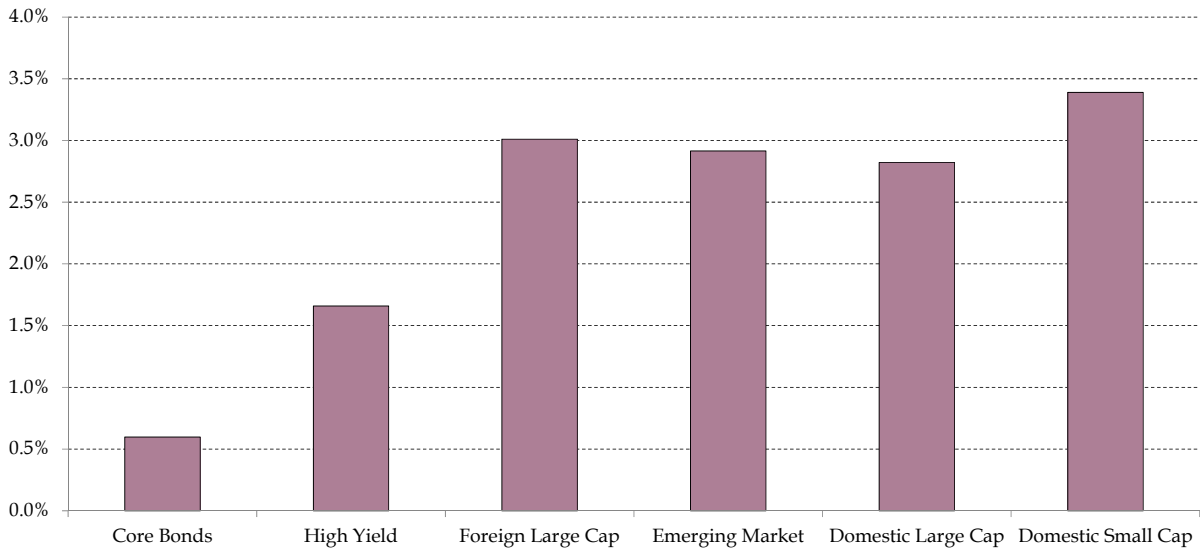
active managers have added value, the median level of fees negated any advantage. Furthermore, the fees tended to be highest in those asset classes that many investors consider to be the least efficient (e.g., small cap stocks and emerging markets).

Asset Class	Median Fee on \$10 mm	Median Fee on \$100 mm
Core Bonds	37 bp	31 bp
High Yield	57 bp	56 bp
Domestic Large Cap	72 bp	60 bp
Domestic Small Cap	103 bp	90 bp
Foreign Large Cap	81 bp	72 bp
Emerging Markets	104 bp	101 bp

Another important metric to consider is the dispersion of manager performance. We measure this dispersion by interquartile spreads which is the top quartile subtracted by the bottom quartile. For example, if 100 managers were ranked by performance and 1 was the highest rank, the interquartile spread would be the 25th manager minus the 75th. We do this by measuring the interquartile spreads over 10-year investment horizons within each asset class. As seen in the following chart, the range of the spreads can vary significantly. The size of this spread is a good indicator of how much value a “skilled” (or lucky) manager can add relative to an “unskilled” (or unlucky) manager. Another way to interpret these results is to think of the size of the spread as an indicator of how much potential value lies in selecting a superior active manager within these asset classes.

There was much more divergence in the returns of equity managers than there was for bond managers, perhaps reflecting the difference in volatility of the underlying asset classes or perhaps the amount of heterogeneity in the securities held by managers in these sectors. U.S. Small Cap managers exhibited the most divergence from each other historically, while the level of divergence was similar among U.S. Large Cap, Foreign Large Cap, and Emerging Market managers.

Interquartile Spreads: 1977 - 2012



Does Style Matter?

Another important dimension to examine is style. We wanted to evaluate whether growth or value managers had fared better than the broad market and if this was affected by capitalization. The following table displays the median manager outperformance based on size and style within U.S. equities.

Asset Class	Median Outperformance ⁸
Domestic Large Cap - Growth	13 bp
Domestic Large Cap - Value	10 bp
Domestic Small Cap - Growth	83 bp
Domestic Small Cap - Value	120 bp

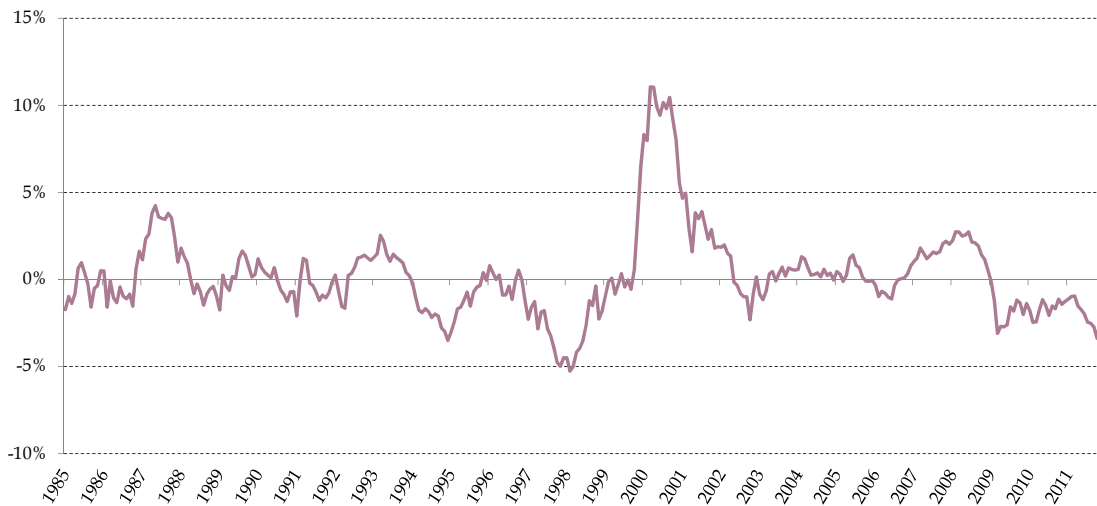
In the large cap space, both the median growth and value managers performed slightly better than their respective style benchmarks. In addition these style managers did slightly better against their benchmark than did the median large cap manager across the broader large cap market. The story is different within the small cap space, where the value managers outperformed the broader small cap market and the growth managers. The value-focused small cap manager median outperformance was also much higher. The non-trivial 37 basis points of outperformance compared to the small cap growth median suggests that style may make more of a difference within small cap than large cap.

⁸ Domestic Large Cap Growth and Value as well as Domestic Small Growth all began in January 1979, while Domestic Small Value did not begin until January 1982.

Cyclicality and Efficiency

Up to this point this paper has only shown snapshot estimates of outperformance using all available data. Using this method gives the most robust estimates due to the high number of data points but it may be misleading because it implies a static level of outperformance. As the following charts indicate, this is not at all the case. For U.S. large and small cap managers, periods of over and under performances are highly cyclical and can be rather long lived.⁹

**Rolling Annual Median Outperformance¹⁰
Domestic Large Cap**



**Rolling Annual Median Outperformance
Domestic Small Cap**



⁹ Note that the other asset classes exhibit similar behavior.

¹⁰ Reflects rolling one-year performance minus the respective benchmark performance over that same period.

One very interesting aspect of both charts is that outperformance tended to occur during bear markets. This is in agreement with the industry arguments in favor of active management and the existing academic literature.¹¹ For example, during the bursting of the technology bubble from 2000 to 2002 there was a large amount of persistent manager outperformance. This result continued during the Great Recession in 2008, although the impact was smaller in scale.

In general, it appears that recently (roughly since 2002) manager performance has been closer to their benchmarks and the magnitude of the cyclicalities has decreased. This observation leads to another interesting question that we can probe: “Have markets become more efficient through time?”¹²

The supporting argument for this thesis is that, as time passes, successful investment strategies become more widely known. As more managers adopt and execute the strategy, the informational advantages of the strategy decrease as more information is reflected in market prices, thus reducing arbitrage opportunities and mispricings.

The following table again examines median manager outperformance. However, it now compares the median outperformance over the past ten years to the median outperformance for entire preceding period. It shows that, across the board, median manager alpha declined over the past ten years relative to the period that came before. This supports the theory that markets have continued to become more efficient through time within every major asset class.

Asset Class	Median Alpha: 2003 - 2012	Median Alpha: pre-2003	Change at the Median from Prior 25 Years
Core Bonds	-59 bp	-6 bp	-53 bp
High Yield	-23 bp	119 bp	-142 bp
Domestic Large Cap	-15 bp	57 bp	-72 bp
Domestic Small Cap	20 bp	427 bp	-407 bp
Foreign Large Cap	-89 bp	282 bp	-372 bp
Emerging Markets	38 bp	300 bp	-263 bp

While we cannot know for sure why this has happened, several possible theories stand out. First, the advent of the internet and the adoption of Regulation FD¹³ made security analysis more of a commodity than it was in the 1980s and 1990s. This development likely reduced

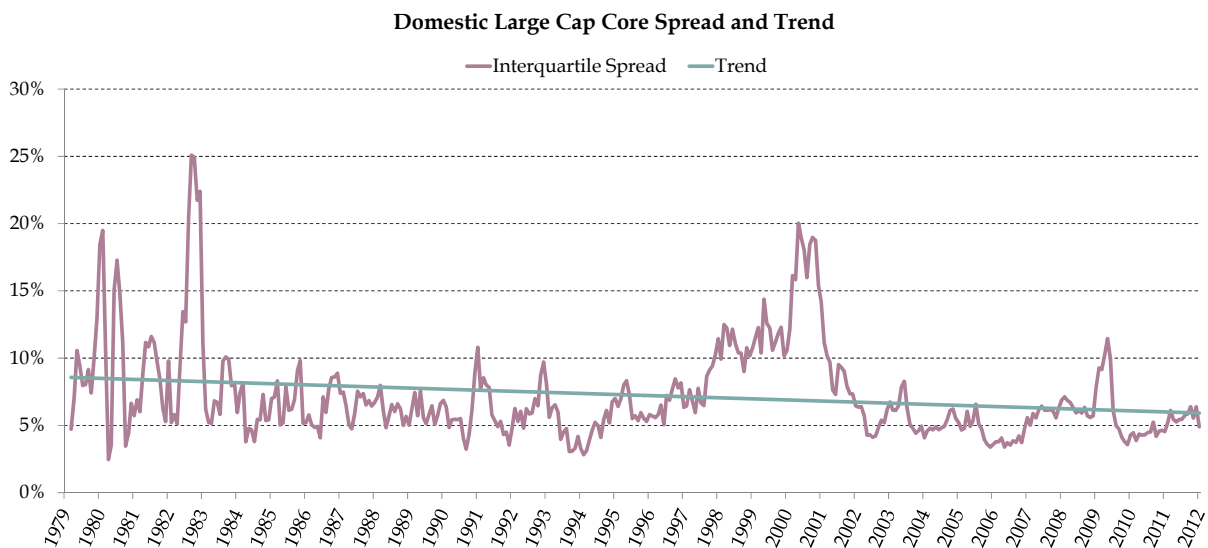
¹¹ Specifically Moskowitz (2000) and Kosowski (2006) found that funds of several different styles produced significantly more alpha during recession periods.

¹² For our purposes, we define an “efficient market” in the manner that is somewhat different than the traditional academic definition. Specifically, we define an efficient market as one in which it is difficult for a manager to consistently outperform their peers and the market as a whole.

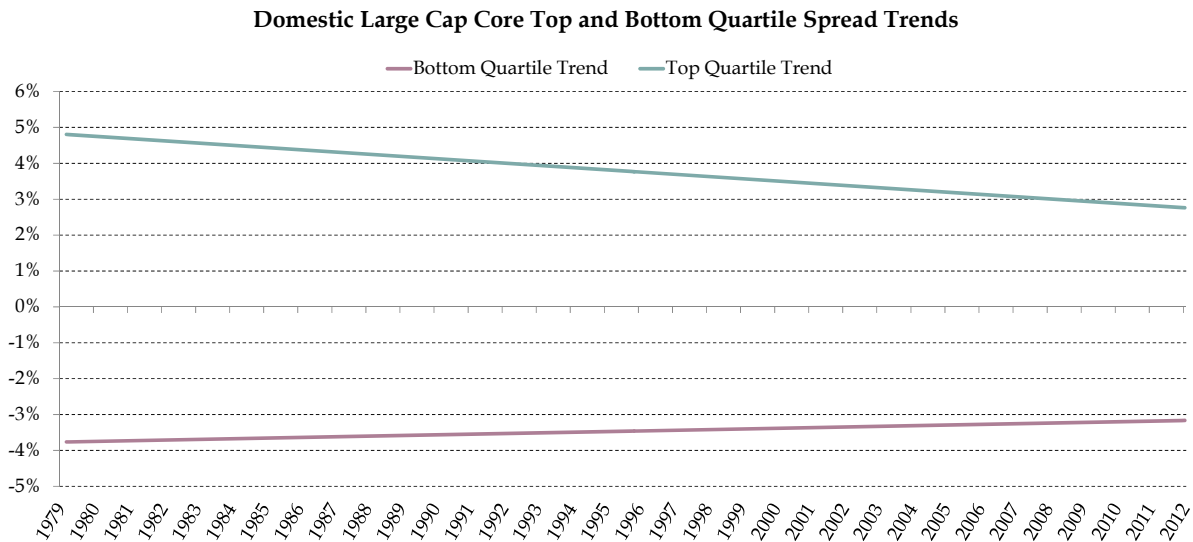
¹³ On August 15, 2000, the SEC adopted Regulation FD to address the selective disclosure of information by publicly traded companies and other issuers. Regulation FD provides that when an issuer discloses material nonpublic information to certain individuals or entities—generally, securities market professionals, such as stock analysts, or holders of the issuer's securities who may well trade on the basis of the information—the issuer must make public disclosure of that information.

the information advantage that some managers possessed. The fact that the reduction in the magnitude of outperformance occurred at roughly the same time as these events lends some credence to this theory. In addition, as mentioned earlier, the strategies used by managers have become more widely known and adopted, resulting in portfolios more closely resembling each other (and the market) than they did in the 1980s and 1990s.

Examining the possibility of increased market efficiency further, the following chart shows the spread between the top and bottom quartile of Domestic Large Cap Core managers through time. Although the interquartile spread did increase during recession periods like 2001 and 2008, the general trend has been one of decline. In other words, the trend since 1979 has been a decreasing amount of difference between the best and worst performing U.S. equity managers.



To further clarify this point, the following chart shows the trend of the top and bottom quartile managers in the large cap core asset class. As before, we see that the distance between these two lines has been decreasing (i.e., the spread has shrunk), but perhaps more interestingly, the chart shows that this decrease has not been perfectly symmetrical. The trend in the bottom quartile managers has increased by 2 bps per year, but the trend of the top quartile managers has decreased by 6 bps per year. This indicates that the majority of the decrease in interquartile spreads came from a reduction in the potential upside rather than from a decrease in potential downside.



Looking across a variety of asset classes, there is a clear and persistent trend that shows interquartile spreads decreasing over time (see the following table). This implies that the potential for outperformance has decreased. The table also shows that while this decrease has been symmetric for the fixed income asset classes, the potential for outperformance declined asymmetrically and to a larger degree for equity managers (with the exception of small cap managers).

Asset Class	Change in Bottom Quartile Trend per Year	Change in Top Quartile Trend per Year
Core Bonds	+8 bp	-8 bp
High Yield	+4 bp	-4 bp
Domestic Large Cap	+1 bp	-10 bp
Domestic Small Cap	+14 bp	-2 bp
Foreign Large Cap	+37 bp	-49 bp
Emerging Markets	+29 bp	-32 bp

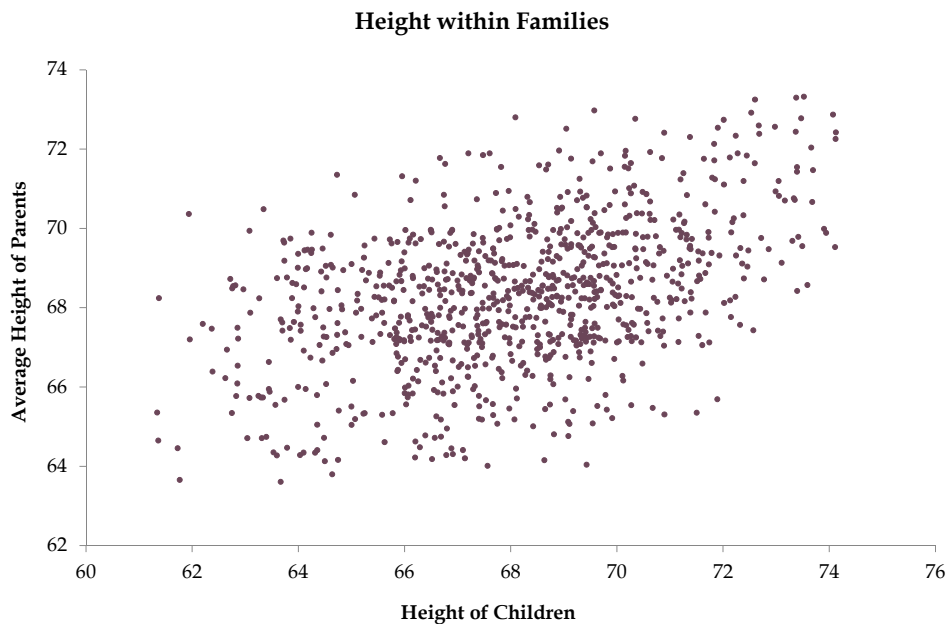
This further indicates the increasingly competitive nature of active investing and the declining reward for investors who select superior managers.

Persistence

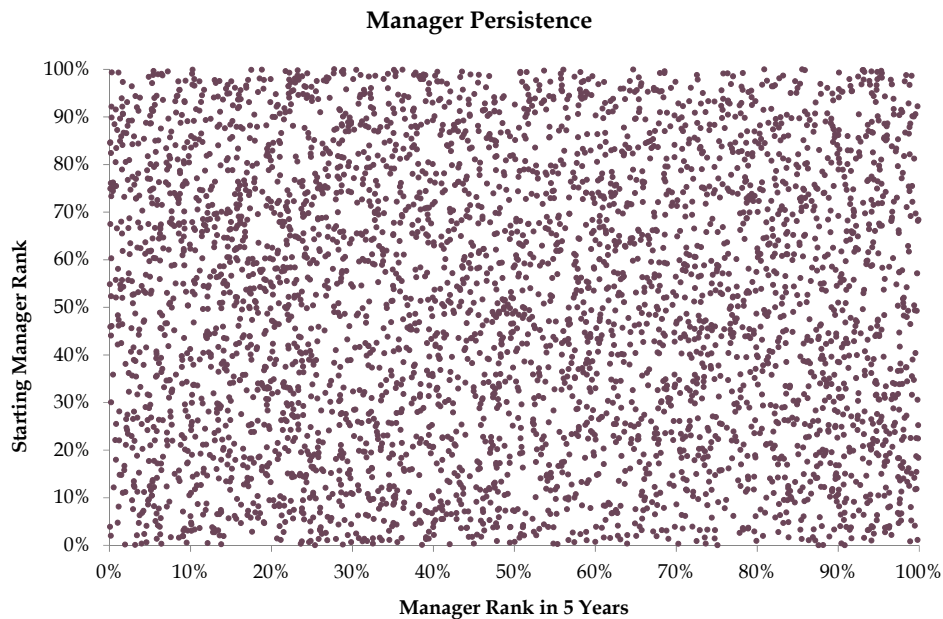
At the beginning of this paper, we briefly discussed the theory of active management as a zero sum game. Among the implications of this are that some managers will outperform and some underperform. Put differently, the majority cannot be significantly above (or below) average, before fees. An obvious next question to ask is whether this out/under performance is purely random or if superior managers persistently outperform. From a practical perspective, the question can be re-phrased as “Is past performance a good predictor of future returns?”

If a manager did possess superior information or analytical capabilities, then that advantage should persist over at least some limited period of time. However, it is quite possible that any advantage will be offset by the competitive forces addressed above. Specifically, the competitive spread of informational advantages is likely to decrease the magnitude and duration of a manager’s advantage, presuming it exists in the first place.

To better understand the idea of persistence, we illustrate a clear case of persistence from outside the realm of economics and finance. The following chart shows the average height of parents on the Y-axis (vertical) and the height of their children on the X-axis (horizontal). Although there is variation of the height of the children relative to the height of their parents, the fact that there is a distinct upward slope of the points is a clear example of persistence -- in this example, that tall parents tend to have tall children.



To evaluate persistence in money managers, we compared their rank from one five-year period to the subsequent five-year period. The following chart shows the results of this analysis for the Domestic Large Cap sector. In this chart, no persistence is evident. In fact, the data points in the Manager Persistence chart appear to be as close to a random distribution as possible.



This analysis indicates that managers who outperformed over a five-year period were no more likely to outperform over the next five-year period than any other manager selected at random. The implications of this are clear and in line with the broader current academic literature,¹⁴ which shows that there is no evident persistence in manager rankings. Put plainly, past performance is not a good predictor of future performance.

¹⁴ As we will discuss in a subsequent paragraph, some studies have found persistence after adjusting for risk or other factors, but even this evidence is mixed.

To further illustrate the point, the following table shows the subsequent ranking of managers in the top decile (i.e., top 10%) and bottom decile (i.e., bottom 10%) over a five-year period. If there were persistence in manager ranking, then the top decile managers would stay near the top (or at least above average) and the bottom decile would stay near the bottom (or at least below average).

Persistence in Manager Ranks
(1% is Best, 100% is Worst)

Asset Class	Rank of Top 10% Managers over Subsequent 5 Years	Rank of Bottom 10% Managers over Subsequent 5 Years	# of Managers
Core Bonds	42%	43%	244
High Yield	53%	48%	326
Domestic Large Cap	51%	46%	2,690
Domestic Small Cap	55%	42%	474
Foreign Large Cap	49%	48%	463
Emerging Markets	49%	53%	283

As the table shows, the average rank over the next five years for both the top decile and bottom decile managers is right around the median, or 50%, rank. It is also worth noting that as the number of managers in the sample increases, the closer the rank tends to be to the median over the subsequent five years. This indicates that as the sample size grows (and the estimate improves), the closer the odds are to a fair coin flip. It is also worth mentioning that, although the evidence is mixed at a statistically significant level, there does seem to be a trend in the subsequent rank of the bottom decile manager tending to be a bit better over the subsequent five years, suggesting that there may be some mean reversion in performance.

Although this paper has provided strong evidence against persistence within manager performance, it is important to note that we have not controlled for any other factors and therefore do not make any statement about manager skill. These factors include but are not limited to Macroeconomic Timing, Style, Sector, or Industry concentrations and variation, and overall active risk.¹⁵ It is possible that, after controlling for these factors, an investor may be able to identify managers who possess skill but exhibit a lack of persistent outperformance due to the cyclicity of these factors or the market rather than any cyclicity of the manager's skill.

¹⁵ The amount of literature on manager performance and contributing factors is dense. For a good summary of the academic literature and factors that may help identify superior managers, see "Active Management in Mostly Efficient Markets," FAJ November/December 2011.

IMPLICATIONS AND CONCLUSION

The results of our analysis show how difficult it is for active managers to consistently add value. Our analysis indicated that median manager outperformance is near zero, depending on the asset class, and has shrunk or turned negative over the past decade. Similarly, the difference in returns between top quartile and bottom quartile managers has shrunk over time. This idea is supported by the narrative of the large competitive forces to which active managers are subject.

The results do show that an investor could improve their odds by seeking active management in certain asset classes (e.g., small cap and emerging markets). The results have also shown that the potential reward for identifying superior managers can vary widely by asset class.

Importantly, our analysis indicates that identifying managers that will perform ex-ante by relying on past performance alone will prove to be a fool's errand. Thus, investors who are seeking superior active managers should primarily rely on other elements, including qualitative analysis.¹⁶

¹⁶ For insight into what qualitative analysis we believe is valuable, please see Meketa Investment Group's white paper titled "The Art of Selecting Investment Managers."

APPENDIX A

DATA FILTERING AND CLEANING

The dataset used in this paper is from Morningstar Direct. There are a total of 6,962 managers included in the dataset with 39 unique benchmarks. The first step in the data cleaning process was to eliminate managers whose benchmark was obviously not appropriate for the asset classes we were trying to analyze. For instance, a benchmark that differs in terms of size, style, or region relative to our target asset class would cause that manager to be removed from the sample. The following table displays how many managers were removed from the sample for each asset class at this step.

Asset Class	Managers Removed Due to Mismatch
Core Bonds	0
High Yield	40
Domestic Large Cap	905
Domestic Small Cap	172
Foreign Large Cap	13
Emerging Markets	0

The next step was to remove managers who indicated in their name that they focus on a sector, size, style, or region which was not the target group. For instance, several frontier market managers were filtered out of the emerging market manager class as well as SMID from small. A small list of filter words was customized to each asset class, and all were similar to the examples above.

The last step was to eliminate any double-counting by eliminating managers who were running the same strategy but in a different vehicle. This was accomplished by examining the correlation of all manager strategies. For any manager strategy with a correlation above 99.8%, one of the correlated pair was dropped from the sample. The following table displays how many managers were removed from the sample for each asset class during these two final steps.

Asset Class	Managers Removed Due to Filtering
Core Bonds	6
High Yield	14
Domestic Large Cap	117
Domestic Small Cap	246
Foreign Large Cap	53
Emerging Markets	17

APPENDIX B

BENCHMARKS AND TIME PERIODS

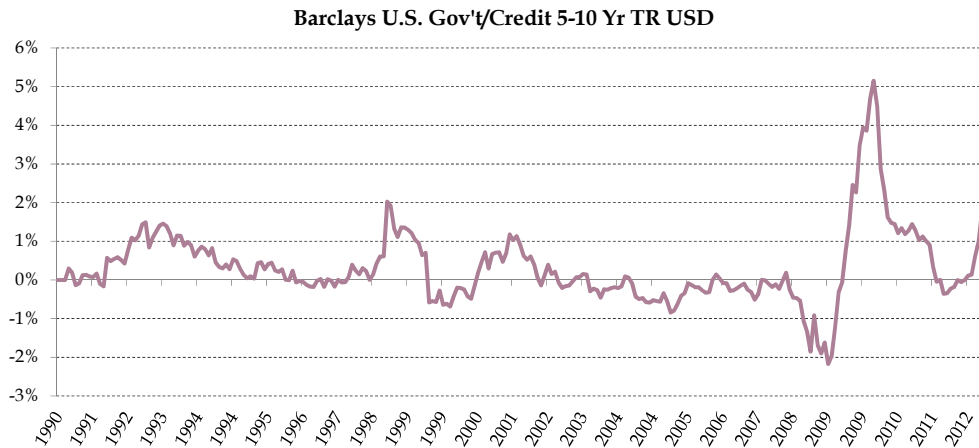
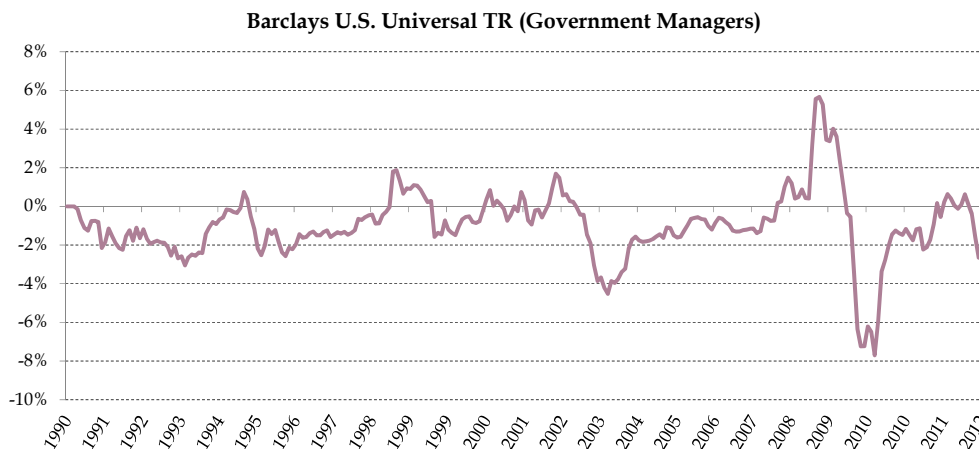
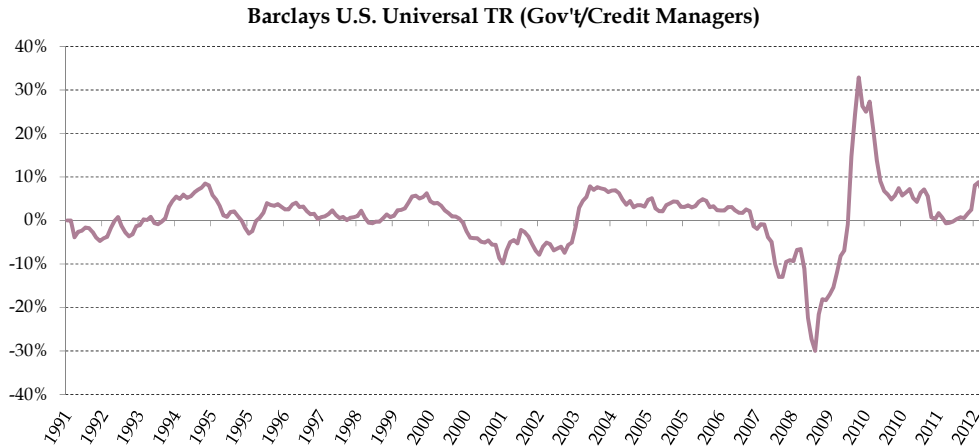
Asset Class	Benchmarks	Starting Time Period
Core Bonds	Barclays U.S. Universal ¹⁷	June 1981
High Yield	Bank of America/Merrill Lynch High Yield Master II	September 1986
Domestic Large Cap	Russell 1000	January 1979
Domestic Small Cap	Russell 2000	January 1979
Foreign Large Cap	MSCI ACWI Ex U.S. ¹⁸	August 1987
Emerging Markets	MSCI EM	January 1988

¹⁷ Core Bonds represent an exception from using the MorningStar Analyst defined benchmark. In this case, the analyst defined benchmarks included the Barclays U.S. Government Total Return, Barclays U.S. Gov't/Credit 5-10 Year, and the Barclays Universal. These were all replaced by the Barclays Universal because it has become the industry standard.

¹⁸ In other cases when growth or value is cited it is the sub-index of the larger index that is used (i.e., Russell 1000 and the Russell 1000 Growth). In the case of Foreign Large Cap, the MSCI ACWI Ex U.S. is used for the broad index but MSCI EAFE Growth and Value are used for the style indices.

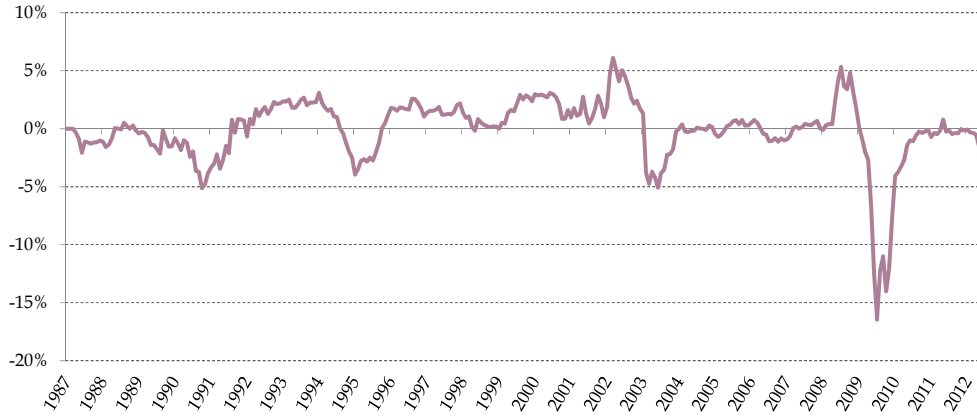
APPENDIX C

TIME SERIES CHARTS OF AGGREGATE MANAGER OUTPERFORMANCE BY BENCHMARK¹⁹



¹⁹ Some data may be skewed due to low number of observations, particularly in earlier years.

B of A ML U.S. HY Master II TR USD



MSCI EAFE Growth NR USD



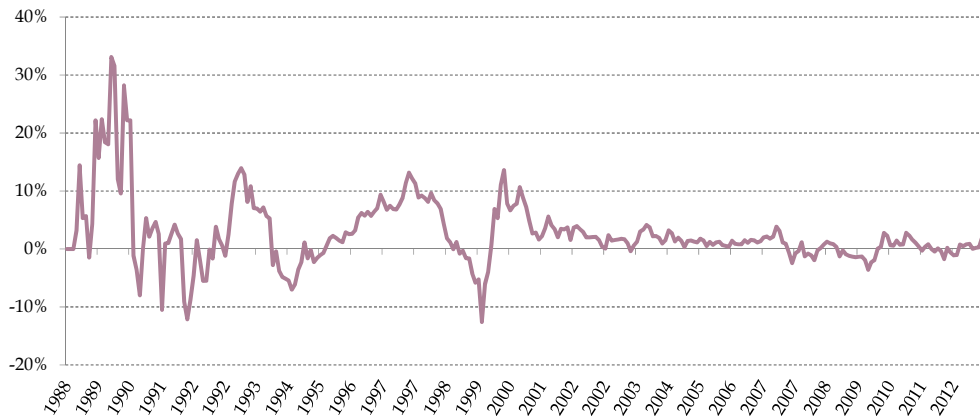
MSCI ACWI Ex USA NR USD



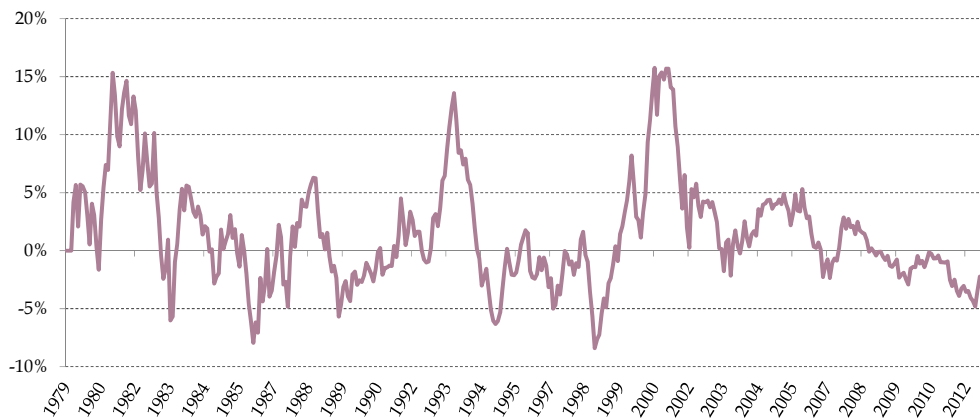
MSCI EAFE Value NR USD



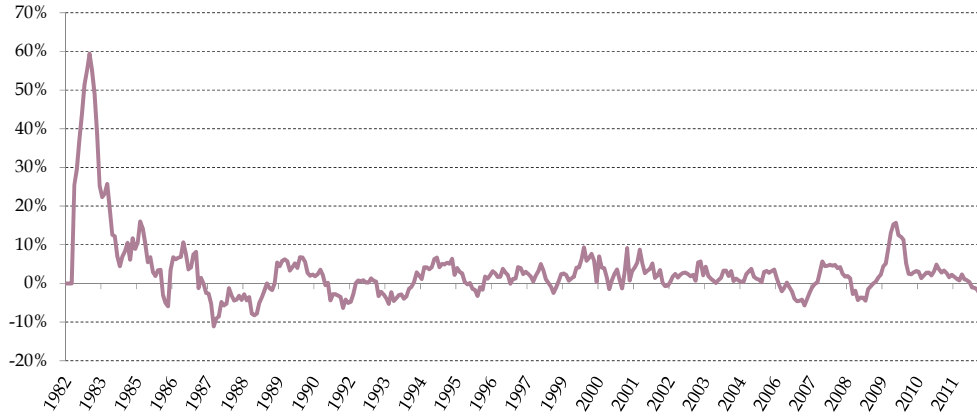
MSCI EM NR USD



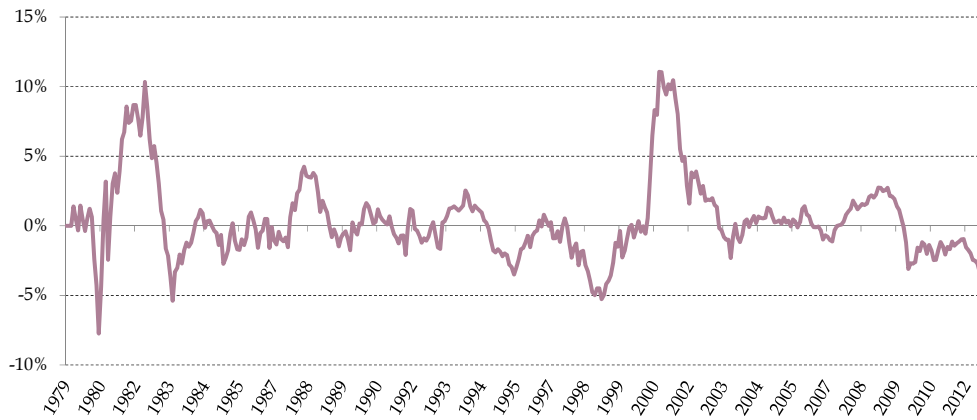
Russell 1000 Growth TR USD



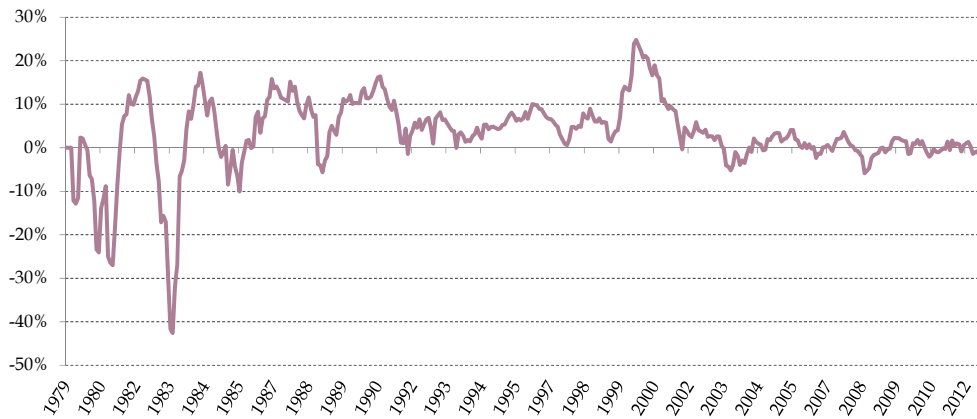
Russell 2000 Value TR USD



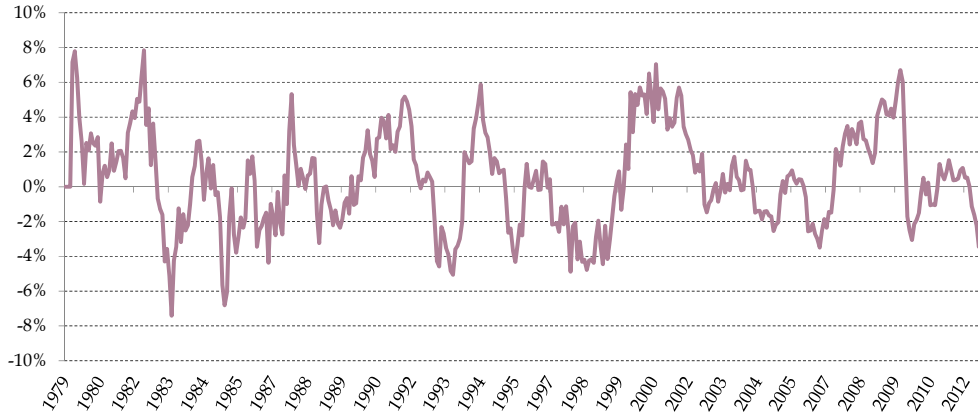
Russell 1000 TR USD



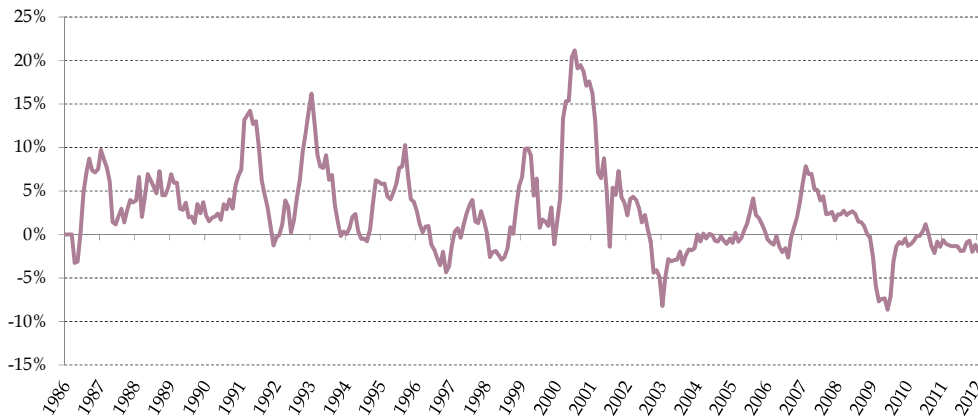
Russell 2000 Growth TR USD



Russell 1000 Value TR USD



Russell Mid Cap Growth TR USD



Russell Mid Cap Value TR USD

