ABSTRACT

With few exceptions, capital markets do not provide predictable short-term investment returns. Careful examination of capital market history, however, may help investors to make estimates of expected long-term returns. These estimates are critical to setting strategic allocations to various asset classes.

In generating expected returns for asset classes using historical data, investors should incorporate the longest period possible. However, investors should also examine whether the period includes a variety of market and economic environments and should test multiple sub-periods to mitigate the bias that may result from arbitrary starting and ending points (known as endpoint bias).

Even historical returns for periods as long as twenty years may not prove a useful guide for generating expected returns. To complement a historical analysis of asset class returns, investors may benefit from forward-looking scenario analysis, based on an understanding of the fundamental drivers of historical returns.

DEFINITION AND DISCUSSION

To avoid huge disappointments and strategic errors, it is crucial that investors understand how much patience and farsightedness the capital markets require.

Periods of twenty years may not be long enough to provide predictable returns. Unfortunately, for many market indices (e.g., emerging markets, high yield bonds), return data do not even extend prior to the mid 1970s. The fewer market environments experienced by a benchmark, the less useful it is as an indicator of long-term investment expectations. Therefore, it is reasonable to assume that investors should focus on the longest period of data available. However, this is only partly true.

Examinations of data from only the longest period available, from inception to the present day, may suffer from *endpoint bias*. Statistically, endpoint bias refers to the inclusion or exclusion of data that significantly influence results. Practically speaking, endpoint bias refers to investors' tendency to place undue significance on results for measurement periods ending in the present. If the recent past witnessed unusually high or low returns, then long-term results can change considerably.

Endpoint Bias Caused by Changing Markets

One might assume that twenty years of data are sufficient to stabilize short-term volatility. However, as illustrated in the following example, twenty-year trailing returns can change dramatically, even when as little as twelve months of return experience are added or deleted from the record. In this first example, data are presented for two different endpoints: March 2000 and March 2001, separated by just one year.

	20 Years As of March 2000	20 Years As of March 2001
Russell 1000 Growth	18.5%	13.2%
Russell 1000 Value	17.2%	15.3%

As of March 2000, the Russell 1000 Growth index had outperformed its Value counterpart by a 130-basis-point annual margin over the trailing twenty years. Using this data to form expectations, investors might conclude that growth stocks offer a long-term premium relative to value stocks. However, when the twenty-year trailing return is measured just one year later, the premium had reversed itself, favoring value stocks by an annualized margin of 210 basis points. Hence, the value-minus-growth gap changed by 340 basis points per year, from –130 to +210 basis points, just by shifting the timeframe by twelve months.

As another example, the twenty-year period ending February 2008 indicated that the S&P 500 index had earned 3.4% more than the Barclays Aggregate, annually. This was fairly consistent with the long-term premium observed for stocks over bonds in the US. However, when measured one year later, investment grade bonds exhibited an annualized twenty-year outperformance of 0.2%. Note that this relationship (of bonds outperforming stocks) lasted for only that single month.

	20 Years As of February 2008	20 Years As of February 2009
S&P 500	10.8%	7.1%
Barclays Aggregate	7.4%	7.3%

Measured over shorter periods, such as five or ten years, the changes in average annual returns can be even more extraordinary. The following table compares the trailing five-year performance for the Russell 2000 Growth and Value indices as of March 2000 and March 2001. As of March 2000, small growth stocks had beaten small value stocks by an average of 10.8% per year. Yet, twelve months later, the outperformance of small growth stocks over small value stocks had reversed wholesale.

	5 Years As of March 2000	5 Years As of March 2001
Russell 2000 Growth	31.8%	11.6%
Russell 2000 Value	21.0%	14.2%

Clearly, making strategic investment decisions based on five years of data can be disastrous. An investor shifting assets from value to growth early in 2000 in response to the prior five years was punished by a 40% loss for small cap growth stocks and a 43% loss for large capitalization growth stocks during the year beginning April 1, 2000. For both the five- and twenty-year periods examined for growth and value stocks, endpoint bias was significant due to the extraordinary rise and fall of technology stocks.

The following example emphasizes that reversals of the data (and the conclusions investors are likely to draw) are not uncommon. International investing looked especially enticing based on the experience of the 1980s (see the following table). As a result, many investors in the 1990s shifted heavily into non-US stocks. Unfortunately, foreign stocks significantly underperformed US equities over the subsequent decade, mainly as a result of a disastrous decline for Japanese stocks. However, the relative performance switched again during the ten years ending in December 2009.

	10 Years As of December 1989	10 Years As of December 1999	10 Years As of December 2009
MSCI EAFE	22.0%	7.0%	1.2%
S&P 500	17.5%	18.2%	-1.0%

Endpoint Bias Caused by Insufficient Data

Often, the time period being measured may be particularly favorable (or unfavorable) for a certain investment style. For example, growth stocks were strongly supported by the dot-com bubble of the late 1990's. Similarly, the time period being measured simply may not have included a financial or economic environment that highlights the true nature of an asset class. Consider bank loans as an example. Bank loans had never experienced more than a 2% loss over a twelve-month period until the arrival of the Global Financial Crisis. Yet during 2008, bank loans declined by -28.8% (see the following chart).[1]

One-Year Rolling Returns for Bank Loans, 1992 to 2011

On a similar theme, consider the case of commodities. An investor looking at commodities in the late 1990's would likely have found them very unattractive relative to stocks based on the performance of the two asset classes since 1980. However, this time period encompasses an extended bull market for US stocks and ignores two bull markets for commodities, one on either end of the original time period (note that "starting point bias" is as significant as endpoint bias when dramatic investment results are at the beginning of the period). Hence, when US stocks are compared to commodities for the full history of the S&P GSCI, the annualized returns were nearly equal (see the table below).

	1980-1998	1972-2010
S&P GSCI	5.7%	10.0%
S&P 500	17.7%	10.2%

---

[1] Bank loans are proxied by the CSFB Leveraged Loan index.

Endpoint Bias in Active Management

The presence of endpoint bias is not restricted to the returns of asset classes. It can also be found in volatility and correlation data, as well as the returns of active managers. The following two tables compare how the median active manager in a particular segment of the market performed versus its benchmark over the previous five years. Unsurprisingly, whether and by how much the median manager outperformed or underperformed changes when viewed at different time periods.

Performance versus Benchmark for Large Cap Value and Growth Managers[2]

	Median for 5 Years Ending December 2006	Median for 5 Years Ending December 2008	Median for 5 Years Ending December 2010
Large Cap Value	-124 bp	35 bp	118 bp
Large Cap Growth	87 bp	103 bp	5 bp

The table above shows that the median value manager was underperforming their benchmark by more than 100 basis points as of 2006, while the median growth manager was outperforming. Just two short years later, both the median growth and value manager was outperforming. And an additional two years led to the median value manager performing much better than the median growth manager relative to their respective benchmarks.

The following table shows similar patterns, with the median large cap manager swinging from laggard to stalwart to mediocrity. Such changes are more likely due to market factors (e.g., cap or sector bias) than they are due to a surge (or decline) in manager skill. This is important to realize, as investors often use this data to inform their decision on whether to use active or passive managers.

Performance versus Benchmark for Large Cap and Small Cap Managers[3]

	Median for 5 Years Ending December 2006	Median for 5 Years Ending December 2008	Median for 5 Years Ending December 2010
Large Cap	-15 bp	42 bp	2 bp
Small Cap	179 bp	39 bp	57 bp

---

[2] Represents annualized manager returns, gross of fees, minus the annualized benchmark return. The benchmarks used were the Russell 1000 Value and Russell 1000 Growth indices. The universes are composed of Morningstar mutual fund returns.

[3] Represents annualized manager returns, gross of fees, minus the annualized benchmark return. The benchmarks used were the Russell 1000 and Russell 2000 indices. The universes are composed of Morningstar mutual fund returns.

DEALING WITH ENDPOINT BIAS

Meketa Investment Group recommends four analytical approaches to both gauge and mitigate the effects of endpoint bias:

1. Examine the longest time period available.

2. Examine periods that contain a variety of market and economic conditions.

3. Examine sub-periods or calculate trimmed means.

4. Examine the underlying drivers of asset class returns.

The first recommended approach acknowledges the simple statistical fact that more data is always better when making inferences based on a sample. The second approach addresses the need to be certain that one's sample is drawn from a representative long-term distribution of returns. That is, if one only has historical data from a bull market, one cannot reasonably suppose that any analysis based on that data will represent performance during a full business cycle (e.g., both a bull and bear market). The third approach acknowledges that full history endpoints may include periods of extreme volatility, and that observing behavior during sub-samples or discarding anomalous data points may help to form better estimates of asset class behavior.

The example of feast and famine in the foreign equity markets illustrates the importance of understanding the actual conditions that drove the returns, our fourth recommended approach. For example, foreign equity returns in the 1980s were dominated by a dramatic increase in the Japanese equity market that elevated the ten-year returns. An investor considering a large allocation to non-US stocks in 1990 or 1991 needed to know that much of the decade's gains were due to a huge expansion of price-earnings ratios in Japan late in the decade. The primary lesson we derive from the return history of the MSCI EAFE index is to avoid dramatic shifts in strategy based on results that may appear to be conclusively long.

Finally, the same considerations hold true for manager evaluation and selection. Managers are routinely hired and fired due to the effects of recent results on long-term performance. More often than not, investors hire managers with very strong recent results, as those managers have almost always produced three-year, five-year, and ten-year returns that are among the highest of their peer group. Similarly, most fired managers have experienced weak short-term results that depress their three-, five-, and ten-year returns.

Our philosophy is based on rigorous investment analysis and a long-term view, accompanied by skepticism of recent fads and trends. Sound investment strategies and concepts will stand the test of time.