



"Statistics are like a drunk with a lamp post: used more for support than illumination." Winston Churchill

Backtesting - Myths Vs. Reality

In the highly competitive world of professional investment management, every manager strives to create or develop an investment strategy that will perform well during a variety of economic and market environments. The hope, of course, is to be able to market a new approach to investing that will attract new assets and generate untold wealth to the creators (and managers) of that strategy.

As the search goes on for the *Holy Grail*, researchers often resort to the time-honored tradition of backtesting a new strategy or investment discipline to demonstrate its superior performance relative to a widely followed benchmark. While admirable, logical and appealing to most observers, there are many pitfalls in backtesting that may not be obvious to most observers and can lead to unjustified and misleading conclusions. To thoroughly understand the pitfalls of this process, it is instructive to look at some of the common problems that confront any organization attempting to create a verifiable and repeatable discipline that will stand up under close scrutiny.

To begin, one of the most common errors is for a manager to produce a number of stocks that meet their investment criteria. In doing so, the manager suggests that these stocks are representative of their strategy and will be the basis for all future stock selections by the manager. In this instance, the weakness in this approach is that the manager ignores many other stocks that possessed the desired attributes but simply failed to perform up to the level of the stocks highlighted in their marketing materials. The error is simply in making a leap from the *particular stocks* selected to the *general stocks* available for purchase. To guard against this problem, an observer must look closely at the components of the overall strategy being implemented and avoid looking at the individual stocks. In this work, it is important to ascertain how often the strategy beat its benchmark, how consistently it beat its benchmark, how badly it performed in a worst case scenario, and how quickly the strategy recovered following a period of underperformance.

One of the pitfalls of any backtesting is the reliability of the database. The two most commonly used databases are the Standard & Poor's Compustat Active and Research Database and the Chicago Research in Security Prices (CRSP) data.

Generally, these two data sources are considered the gold standard for backtesting but both have issues that need to be recognized. Unfortunately, both contain errors such as unrecognized stock splits, inaccurate book value computations, misstated earnings, incorrect price data and a myriad of similar data mistakes. These errors will be present in all backtesting and need to be recognized when making judgements about the results. One way to compensate for this problem is to be certain that the investment results from a proposed strategy must show a clear and unmistakable superiority over a selected benchmark so that the results are not simply an accident of timing, and unlikely to be repeatable in the future.

Another significant issue is *timeframe limitations*. Although the CRSP starts in 1926, the data set does not cover more than 50% of the historical record of widespread, large scale stock trading in the United States that goes back almost 200 years. As for the Compustat data, its data is even more limited beginning only in 1963. Then, there is the lack of coverage for all traded stocks in the CRSP data as it excludes the majority of stocks trading in the U.S, particularly the smaller equities that are usually the more vulnerable and volatile companies. In addition, Compustat added many small company stocks in the late 1970's that may create an unintended bias to their results.

One of the major issues with all backtesting efforts is the issue of data-mining. With intense pressure to draw statistically significant results from any set of data, it is a known fact that if the data is tortured long enough, it will yield just about anything desired by the researcher. In brief, if there appears to be no sound theoretical, economic or intuitive, common sense reason for a relationship, it is most likely a completely random occurrence. Frankly, the best way to confirm that excess returns are genuine is to test them using different time periods or sub-periods.

Yet another pitfall in backtesting is the use of a limited time period. In brief, anything can look good for a five and, perhaps, even a ten year period. An interesting example is the choosing of stocks with ticket symbols that are the vowels: A, E, I, O U and Y in 1996. Those stocks beat the S & P 500 by 11% that year but it does not constitute a sound strategy for long-term investing. In most stock market literature, this is an example of a small sample bias. Obviously, a larger sample will generate a higher level of confidence than those derived from a smaller sample.

Unfortunately, many studies suffer from the fact that stocks that fail are not included in the raw statistics, creating a classic *survivorship bias* that totally distorts the results of the research work.

It is inevitable that many companies disappear from the database due to bankruptcy, takeovers or other similar corporate events.

And, lastly, there is the *look-ahead bias* in which researchers assume the investor had annual earnings data in January when the data might not have been available until March. This practice results in upwardly biased results.

In summary, backtesting is a useful way to explore a new investment discipline or strategy but it must be critically analyzed and viewed with a healthy amount of skepticism. Backtesting is susceptible to data manipulation, false assumptions and the use of misleading information to produce a desired results. While one should always assume that investment professionals operate with the highest ethical standards, there will always to those who attempt to skew the results in a desired fashion. In keeping with this thought, a recent conversation with a financial advisor resulted in his comment that “he’d never seen a backtest he didn’t like.” All of which reminds me of my college professor of statistics who often said.....given enough time and attention, any well-schooled statistician can produce just about any conclusion from a given set of data. Which, in turn, brings to mind one of my college textbook with the provocative title.....*How to Lie with Statistics*.

Stephen Kent, Jr.

**Stephen K. Kent, Jr., CFA
Founding Partner and CIO
Metis Value Partners, LLC
May 15, 2018**

DISCLOSURES:

Past performance is not indicative of future results, which may vary. The value of investments and the income derived from investments can go down as well as up. It shall not be assumed that recommendations made in the future will be profitable or will equal the performance of the securities mentioned here. While MVP seeks to design a portfolio which reflects appropriate risk and return features, portfolio characteristics may deviate from those of the benchmark.

Nothing herein should be construed as a solicitation or offer, or recommendation to buy or sell any security, or as an offer to provide advisory services in any jurisdiction in which such solicitation or offer would be unlawful under the securities laws of such jurisdiction. The material provided herein is for informational purposes only. Before engaging MVP, prospective clients are strongly urged to perform additional due diligence, to ask additional questions of MVP as they deem appropriate, and to discuss any prospective investment with their legal and tax advisers