SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

Introduction

Boston Harbor's mission is to preserve and grow our clients' capital. We achieve that mission by creating investment strategies that meet the goals of preservation of capital, total return, and liquidity. We are differentiated from other asset managers in that the cornerstone of our investment philosophy is based on a warning from William F. Sharpe.

William F. Sharpe, 1990 Nobel Prize winner for his theory of risk and return, precisely states our belief about the work to be done before offering any investment strategy to our clients:

"Although it is always perilous to assume that the future will be like the past, it is at least instructive to find out what the past was like."

We are about using our proprietary micro economic theory of the firm to create validated investment strategies. All of our strategies will show our clients historical performance over at least 20 years of market cycles using a disciplined scientific methodology to analyze the historical data.

Unfortunately, not many investment strategies meet Boston Harbor's back-test standard of covering bull and bear market cycles to show the investor how a strategy performed over market cycles. Moreover, those that do are too often receiving the formula response that "back-tested strategies are not a substitute for live returns."

We explore the "live returns" vs. "back-tested returns" controversy and conclude that the Prudent Investor will find out what the past was like before investing and that there is no statistical evidence that live returns persist into the future.

Regulatory Guidance

Investment adviser advertising, including performance advertising, is principally regulated at the federal level under the general antifraud provision of the Investment Adviser Act of 1940, Section 206 and Rule 206(4)-1. The SEC has articulated guidelines for advertising through rule, a series of "no-action" and interpretive letters under the rule and numerous enforcement actions that together provide the regulatory guidelines for scientific back-testing which is not fraudulent.[1] Absent specific guidance from the SEC, the key test remains whether the data is presented in a manner that is not misleading.

Back-testing, like any other disciplined scientific methodology to analyze the data, can be misleading if applicable robust protocols are not followed to prove the hypothesis. The SEC guidelines that have been developed over time are designed to provide evidence that a hypothesis is valid and not misleading. This is accomplished with some commonsense guidelines, which we will summarize after discussing the classification of back-tests.

Classification of Back-Tests and Risk

Today, a common business model for the top institutional asset managers is to sponsor a back-test with leading financial academics or Index providers and to use the results of the test (performed with strict protocols) to attract assets to manage from institutional investors.[2] Consequently, top money managers and institutional investors are not likely to invest in any investment strategy that does not demonstrate historical returns over bull and a bear market cycles.

Back-testing covers several different methods of which the scientific method will produce the most valuable predictors of the future. Back-testing can either be performed by creating a hypothesis of causal relationships and then testing the hypothesis with historical data to obtain returns (scientific method) or by finding variables in historical data that correlate to returns and developing a hypothesis from the historical data (data mining) or applying any hypothesis to different time periods until favorable returns are discovered (also data mining).

[1] Liklass Jennifer L, "SEC Regulation of Performance Advertising by Investment Advisors," Morgan, Lewis & Brockius LLP, October 23, 2002.
[2] Baker Malcolm, Brendan Bradley and Jeffrey Wurgler, "Benchmarks Limits to Arbitrage: Understanding the Low Volatility Anomaly," Financial Analyst Journal, Volume 67, Number 1, 2011 CFA Institute.

SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

Data mining back-tested models are developed with the benefit of hindsight, or ***in-sample***, because the data is used to define the correlations for the variables. These correlations may not have foresight of the future, because they may not define any causation.

Scientific back-tested models are developed from a theory or hypothesis developed without the use of the data. The testing is done to verify the hypothesis. This type of testing is defined as out-of-sample testing. ***Out-of-sample*** means that the model makes the decisions before the answers are available in the data. In technical terms, the protocol is free of look-ahead bias.

Hypothetical returns do not reflect the macroeconomic risks of using the strategy in a different time period or the market impact risk of executing trades in a live portfolio . Market impact risk is the potential market impact on stock prices, caused by buying or selling that could cause the model's buy or sell prices to differ from the frictionless trades of the back-tested model.

Macroeconomic and market impact risks are very real risks that must be considered for any strategy. The SEC guidelines provide that a test should be over a bull and bear market cycle to minimize the macroeconomic risks. As a result, any test that provides returns for 20 years or more provides historical return data that is not likely to get any better to support investment strategy decisions. The investor should beware of any back-test returns that are not over two different market cycles. For example, a new ETF marketed with historic performance statistics that begin at the bottom of the 2008-2009 decline through calendar 2012!

Market impact is also an important risk that could impact returns. This risk is a function of liquidity. For example, returns for a low turnover portfolio of S&P 500, liquid, large-cap stocks are estimated to be decreased by about 0.32%,[3] but that risk might increase to over 1.00% for a portfolio of illiquid small-cap stocks or by high turnover of portfolios. Buying and selling can cause large deviations from expected returns ('tracking error"), because of lack of liquidity. Those deviations could make the expected returns from frictionless trades disappear.

OUT-OF-SAMPLE vs. IN-SAMPLE BACK-TEST

Stephen H. Penman, Chair Accounting Division, Graduate School of Business, Columbia University, an authority for financial statement analysis and security valuation (author of approximately 50 academic papers using scientific back-testing methodology) and a member of Boston Harbor Board of Advisors comments on the Boston Harbor's SELECT 40 back-test:

"One must maintain the reservation that a "back testing" result might just come from data dredging. Even so, there is a difference between an in-sample finding and an out-of-sample confirmation. But the process of taking a hypothesis, that has been developed independently of the data, to the data for testing is the scientific method. There is a problem if the "experimenter" then tweaks the strategy after seeing what works in the data. There is also the problem of "stationarity": the conditions in the back testing period do not apply going forward. That is mitigated somewhat by testing the robustness under different conditions in the back test and out-of-sample.

These reservations are diminished by the strength of the idea. A back test supporting SELECT 40 is different from one supporting a sun-spot theory of stock prices (in my mind). You will have difficulty persuading an investor whose inclination is not of a fundamentalist nature. But, someone with some commitment to fundamental analysis would certainly have reservations about a strategy if he or she did not see confirmation in a back test!"

[3] Ferraris Andrew Dr., "Equity Market Impact Models," Deutsche Bank AG, December 4, 2008.

SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

SEC BACK-TEST PROTOCOL FROM REGULATORY DECISIONS:[4-5]

- Scientific method used to develop a hypothesis rather than data mining
- Test performed out-of-sample
- Data supplied by a recognized data vendor to the financial industry
- Back-test performed by a qualified independent third party
- 20 year test period, but if fewer years, the test must include bull and bear market cycles
- Stock selection rules to be the same for the initial test and all subsequent tests
- Systematic execution timing of stock purchase decisions.
- Back-test results reproducible by a third party that has access to the data and the rules
- Back-test results and data available for audit by the SEC
- Disclosure of back-test methodology and appropriate risk factors and reporting the difference, if any, from historical data of actual performance

SELECT 40 SCIENTIFIC OUT-OF-SAMPLE BACK-TEST:

- A hypothesis (theory of the firm) was developed and a memorandum prepared defining the economic drivers of a well-managed, healthy, company that had a high probability of sustainable earnings.
- A respected global math and statistics firm was engaged to convert the memorandum into rules and embed those rules into proprietary software. The software would then execute the rules to select stocks that met all the criteria for selection based on fundamental factors. Those rules are the same today as the first time the math firm pushed the go button to execute the rules to select stocks from the historical database.
- A global data vendor was engaged to create a custom database of SEC Form 10-K fundamental data containing twelve factors for all the S&P 500 companies for each year from 1991 through 2007. (Subsequently, databases were compiled each year for 2008, 2009, 2010, 2011, and 2012. Those databases included the bull market of the 1990s and the bear markets of the 2000s.)
- The fundamentals database was delivered to the math firm and used in conjunction with the developed software to select the Strategy portfolios for each year.
- For each of the years 1990 through 2012 the software applied the exact same rules to each stock using the same SEC Form 10-K data to select a portfolio. The average size of the portfolios is 85 S&P 500 large-cap stocks. The stocks were selected in March following the calendar year end. March month end closing prices were used for all initial buy and one year later sell transactions. Monthly close prices were used to compute monthly returns. Consequently, all returns are out-of- sample and none of the rules were derived in-sample from the data. Because these portfolios are comprised of large-cap, liquid, securities, the only difference between the study returns and live returns is logically market impact and trading costs, which together for large-cap, liquid, low turnover, portfolios are estimated at 0.32%[6]. (SELECT 40 market capitalization is estimated as of April 1, 2012 at $2.5 trillion.)
- All information is time stamped and held by the math firm and Boston Harbor. This information is available for audit by the SEC.
- Portfolios averaging 85 large-cap S&P 500 stocks are sub-divided into strategies based on rules. Large-Cap Strategy (average 85 stocks), SELECT 40 (40 of the average 85 stocks with the highest probability of sustainable earnings), and SELECT DIVIDEND (10-20 of the average 85 stocks paying a consecutive and increasing annual dividend for at least 20 years).
- SELECT DIVIDEND portfolio has been live since 2010 and the SELECT 40 portfolio since April 1, 2013. The live portfolio performance has been as expected based on the back-tested performance.
- See SELECT 40 brochure, page 7, for comprehensive SELECT 40 performance statistics.

LIVE PERFORMANCE IS NOT AN INDICATOR OF FUTURE RETURNS.

Historically our SELECT 40 Large-Cap Low-Volatility Low-Beta Investment Strategy, based on systematic (disciplined) quantitative traditional fundamental financial statement analysis, produces absolute and high-risk adjusted returns over

[4] Ibid: Liklass
[5] In the Matter of Raymond J. Lucia Companies Inc, and Raymond J Lucia, SR, Securities Exchange Commission, Administrative Proceeding File No. 3-15008, July 8, 2013.
[6] Ibid Ferraris

SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

market cycles and long term. The classic question when evaluating quantitative managers is "How often do you tinker with your model?" If the investment strategy becomes an index, the answer is "never.[7]"

SELECT 40 is an index based on Wilshire's insight that differentiates an index from a quantitative systematic (disciplined) strategy where rules can change to "fine tune" the model, or when an investment committee applies a "qualitative overlay" to the model's decision. Traditionally, indexes are classified as passive investment. Strategies that pick stocks or use judgment to override systematic strategies are classified as active investment.

We consider the S&P 500 Index to be an investment strategy that actively changes its large-cap constituents using rules, but slowly (about 5% annually[8]). SELECT 40 is an index that changes its constituents slowly (60%-70% annually). Its constituents are a sub-set of the S&P 500 companies. The S&P 500 Index is recognized as the most difficult index in the world to outperform. SELECT 40's annualized return for the 20 years ending December 31, 2012 was 11.39% as compared to the S&P 500 annualized return of 8.22%. Let's take a look how other asset managers performed as compared to the S&P 500.

S&P Indices versus Active Funds [SPIVA®] Scorecard Year-End 2012:

PERCENT OF EQUITY MUTUAL FUNDS OUTPERFORMED BY S&P INDICES
JANUARY 1, 2003 to Year-End 2012

Exhibit 1: Annual League Tables

Fund Category	Benchmark Index	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Average
All Domestic Funds	S&P Composite 1500	47.7	51.4	44.0	67.8	48.8	64.23	41.67	57.63	84.07	66.08	57.3
All Large-Cap Funds	S&P 500	64.6	61.6	44.5	69.1	44.8	54.34	50.75	61.83	81.28	63.25	59.6
All Mid-Cap Funds	S&P MidCap 400	56.4	61.8	76.0	46.7	46.4	74.74	57.60	78.19	67.36	80.45	64.6
All Small-Cap Funds	S&P SmallCap 600	38.8	85.0	60.5	63.6	45.0	83.77	32.22	63.02	85.78	66.50	62.4
Large-Cap Growth Fund	S&P 500 Growth	44.7	39.5	31.6	76.1	31.6	89.95	39.15	82.00	95.63	46.08	57.6
Large-Cap Core Funds	S&P 500	66.0	66.9	44.6	71.3	44.0	52.03	52.06	63.20	81.31	66.29	60.8
Large-Cap Value Funds	S&P 500 Value	78.5	83.2	58.8	87.7	46.3	22.17	46.24	34.67	54.26	85.06	59.7
Mid-Cap Growth Funds	S&P MidCap 400 Growth	31.7	59.6	78.5	34.8	39.3	88.95	59.60	82.14	75.39	87.22	63.7
Mid-Cap Core Funds	S&P MidCap 400	50.0	51.8	72.4	35.9	64.6	62.28	68.60	82.00	64.07	79.66	63.1
Mid-Cap Value Funds	S&P MidCap 400 Value	81.9	63.6	71.8	38.4	56.1	67.06	47.83	71.76	64.86	76.24	64.0
Small-Cap Growth Fund	S&P SmallCap 600 Growt	35.3	93.6	72.2	52.1	39.4	95.50	33.49	72.68	93.75	63.72	65.2
Small-Cap Core Funds	S&P SmallCap 600	33.3	82.9	61.4	62.8	51.9	82.46	34.45	60.21	86.10	68.44	62.4
Small-Cap Value Funds	S&PSmallCap 600 Value	49.3	77.5	46.0	76.7	39.8	72.55	26.27	51.81	83.00	61.83	58.5

Source: Standard & Poor's 2003-2006; S&P Indices, CRSP 2007-2012

For the 10 year period, 2003-2012 the Passive S&P 500 Index outperformed 59.6% of all large-cap mutual funds. The SELECT 40 outperformed the S&P 500 50% of the time. In 2008, S&P 500 declined -37.00% (outperformed 54.34% of funds) as compared to SELECT 40's -13.81%. SELECT 40 annualized return for the 10 years period was 11.84% as compared to the S&P 500 7.10%.

Importantly, the equity funds that outperformed their benchmark are a changing universe of funds making the index performance that much more impressive. The bottom line is that the data proves that it is very difficult for active managers to consistently outperform the passive S&P 500 or other S&P indexes, and to add value that exceeds active management fees in an efficient market. For the 10 years ending March 1, 2013, SELECT 40 outperformed 99.2% of all low risk domestic equity funds and 99.8% of all low-risk large-cap domestic equity mutual funds.[9]

[7] Waid Robert J., "Fundamentally Active," Wilshire Associates Incorporated, 2007.
[8] Asem Ebenezer, Shamsul Alam, "Changes in the Constituents of the S&P 500 Index and the Performance of the Index," University of Lethbridge, Canada, AB, September 24, 2011.
[9] Morningstar Mutual Fund Screener March 1, 2013.

SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

S&P Persistence Scorecard July 2013:

The phrase "past performance is not an indicator of future outcome" can be found in the fine print of most mutual fund and other financial industry literature. However, notwithstanding the admonition, some investors and advisors consider past performance to be important factors in fund selection. So does past performance really matter? To answer this question on a continuous basis, the S&P Persistence Scorecard, released twice a year, tracks the consistency of top performers over yearly consecutive periods and measures performance persistence through transition matrices.

Summary of Results:

- Very few funds can consistently stay at the top. Out of 703 funds that were in the top quartile as of March 2011, only 4.69% managed to stay in the top quartile over three consecutive 12-month periods at the end of March 2013. Further, 3.35% of the large-cap funds and 6.08% of the small-cap funds remain in the top quartile. It is worth noting that no mid-cap funds managed to remain in the top quartile.
- For the three years ended March 2013, 16.57% of large-cap funds, 14.22% of mid-cap funds and 23.05% of small-cap funds maintained a top-half ranking over three consecutive 12-month periods. Random expectations would suggest a rate of 25%.
- Looking at longer-term performance, only 2.41% of large-cap funds, 3.21% of mid-cap funds and 4.65% of small-cap funds maintained a top-half performance over five consecutive 12-month periods. Random expectations would suggest a repeat rate of 6.25%.
- While top-quartile and top-half repeat rates have been at or below the levels one expects based on chance, there is consistency in the death rate of bottom-quartile funds. Across all market cap categories and all periods studied, fourth-quartile funds had a much higher rate of being merged and liquidated.

The S&P Persistence Scorecard July 2013 shows that only 3.35% of the large-cap funds and 6.08% of the small-cap funds consistently stay at the top quartile over three consecutive 12 month periods. However, the worst performing fourth-quartile funds show persistence of poor performance. The lack of persistence in positive performance is also being found in the performance of institutional managers other than mutual funds.

A recent study examined the persistence in performance of 4,617 active domestic equity institutional products managed by 1,448 investment management firms between 1991 and 2008. Even though there is considerable heterogeneity in performance, there is only modest evidence of persistence.[10]

William H. Miller

The active vs. passive management debate and the scientific ***out-of-sample*** back-test vs. live returns debate are linked in the history of William H. Miller who is a famous active value investment manager that picked stocks without offering the investor a scientific back-test of his value investment strategy to demonstrate historical performance and risk.

Bill Miller spent nearly two decades building his reputation as the era's greatest mutual-fund manager. Then in 2008, he destroyed it. Fueled by winning bets on stocks other investors feared, Mr. Miller's Legg Mason Value Trust outperformed the broad market every year from 1991 to 2005. It's a streak no other fund manager has come close to matching.

Mr. Miller was in his element when troubles in the housing market began infecting financial markets. Working from his well-worn playbook, he snapped up American International Group Inc., Wachovia Corp., Bear Stearns Cos. and Freddie Mac. As the shares continued to fall, he argued that investors were overreacting. He kept buying.

What he saw as an opportunity turned into the biggest market crash since the Great Depression. Many Value Trust holdings were more or less wiped out. After 15 years of placing savvy bets against the herd, Mr. Miller had been trampled by it.[11] In 2008, Value Trust was down -54.61% as compared to -37.00% for the S&P 500 and -13.81% for the SELECT 40.

When the 2008 financial crisis began, Merrill Lynch, Bear Stearns, Lehman Brothers, American International Group, Washington Mutual, and others were absent from our portfolios because the Strategy rated those firms no buys for 2008.

[10] Busse Jeffrey A., Amit Goyal, Sunil Wahal, "Performance and Persistence in Institutional Investment Management," The Journal Of Finance, Vol LXV, No 2, April 2010

[11] Lauricella Tom, "The Stock Picker's Defeat," WSJ.com, December 10, 2008.

SELECT 40 BACK-TEST PROTOCOL AND TEST RETURNS VS. LIVE RETURNS AS PREDICTORS OF FUTURE PERFORMANCE

Legg Mason Value Trust (LMNVX) annualized return for the 20 years ended December 31, 2012, was 9.14% as compared to SELECT 40's back-tested annualized return of 11.39% and the S&P 500 annualized return of 8.22%.

Importantly, to obtain a complete comparative profile of the three strategies for the 20 year period we need to compute the return per unit of risk: Value Trust 0.35, SELECT 40 1.08 and S&P 500 0.54. For the 20 years, Value Trust annualized volatility-risk was 26.01% as compared to SELECT 40 10.52% and S&P 500 15.12%.

The spectacular Value Trust 15 year annualized return of 15.00% as compared to SELECT 40 12.16% to December 31, 2007, was fueled by 64.87% greater volatility-risk than the market. (SELECT 40 return per unit of risk for the 15 years was 1.27 as compared to Value Trust 0.67). In the 2008 decline Value Trust investors' loss -54.61% or 47.59% more than the S&P 500 decline of -37.00%. If you invest in high volatility-risk strategies to make money long term you have to know when to get off the train. SELECT 40 2008 loss was -13.81% without the investor having to time when to get off the train!

SUMMARY:

S&P SPIVA Scorecard Year-End 2012 reports that the passive S&P 500 Index outperforms 62.29% of large-cap equity mutual funds for the five years ended calendar 2012. For this period, a changing universe of 37.71% of large-cap equity mutual funds outperformed the S&P 500. For the 5 year period, SELECT 40 performed as expected given market conditions and outperformed the S&P 500 40% of the time with a back-tested annualized return of 9.11% as compared to an annualized return of 1.66% for the S&P 500.

S&P Persistence Scorecard July 2013 reports that only 2.41% of large-cap mutual funds maintain top-half performance over the previous five consecutive 12 month periods. The lack of persistence of performance is confirmed in the Busse 2010 paper for institutional investment management. And, the history of Legg Mason's flagship active stock picking fund, Legg Mason Value Trust, and it's horrific -54.61% loss in 2008 is additional evidence that live returns are not a predictor of future performance.

If we draw one conclusion from this discussion of back-tested vs. live returns it is that for investors to make an investment manager decision the optimum is for him or her to have both scientific back-tested and live returns available to make a judgment before investing in an ETF, mutual fund, institutional money manager, or fund managed by a financial advisor. The odds are that if one invests based only on live returns the outcome will not be good, because there is no evidence live returns predict the future.

In contrast to the facts that the S&P 500 consistently outperforms about 60% of all actively managed mutual funds and that mutual fund and institutional investment manager live returns do not persist, we have the SELECT 40 scientific back-test for 20 years over bull and bear market cycles: SELECT 40 annualized return for 20 years of 11.39% as compared to S&P 500 8.22%. Moreover, the back-test is presented for the investor with an explanation of the systematic (disciplined) rules based Strategy and supporting performance and risk statistics.

The Prudent Investor will seek evidence of live long-term outperformance for a recommended investment strategy as compared to the S&P 500 and the performance statistics for that strategy to evaluate risk. The Prudent Investor will also seek scientific back-tested strategies rather than rely solely on the less than random chances, that untested live strategies or stock picking returns that outperform the benchmark, will persist. Given easy access to computing power and available commercial databases, it is irrational and dangerous to invest if the investment manager does not demonstrate, as is the practice of top institutional money managers, how its strategy performed in the past over bull and bear market cycles.

"Although it is always perilous to assume that the future will be like the past, it is at least instructive to find out what the past was like." - William F. Sharpe, 1990 Nobel Prize winner.

PAST HYPOTHETICAL OR LIVE PERFORMANCE IS NOT A GUARANTEE OR A RELIABLE INDICATOR OF FUTURE RETURNS. All investments contain risks and may lose value

Dennis N. Caulfield | CIO | Executive Vice Chairman
978.369.8035 | dennis.n.caulfield@bostonharborim.com

September 15, 2013