Elusive return predictability: Discussion

International Journal of Forecasting 2007 (forthcoming)

Stephen J. Brown
David S. Loeb Professor of Finance
NYU Stern School of Business
44 West 4th Street
New York NY 10012

sbrown@stern.nyu.edu

Two major conclusions follow from this very careful study. First, sophisticated prediction tools do not fare well relative to naive models predicting return based on past sample means. Second, there appear to be short-lived episodes of quite limited return predictability. These conclusions are consistent with all we know from the theoretical developments in financial economics over the past thirty five years and more. Yet how do we reconcile these facts with the widespread perception that market returns are in fact predictable, and that hedge funds in particular are adept at exploiting this predictability?

The increasingly sophisticated apparatus presented in the financial asset pricing literature has as its objective deriving models that might most accurately describe the cross section of expected security returns. The sample mean return can be considered a reasonable estimator of expected return, and so it is not surprising that it is a very robust straw man when comparing alternative predictive models. The apparent disconnect between the financial markets and real economic activity has encouraged a behavioral view of the markets and the hope and promise of limited predictability. However, recent work by Lettau and Ludvigson (2001) explains that at least part of this tension arises from the strong assertion of stationarity. Allowing for conditional expectations which can change through time as economic conditions fluctuate explains much of

this disconnect. This suggests that only sample means based on the most recent historical data are relevant for predicting future returns, consistent with the results in this paper. Further, it suggests some benefit might be drawn from conditioning the predictions on current state variables characterizing the state of the economy. However, while this is of some scientific interest, it provides limited guidance for practitioners, as presumably the alternative investments necessary to capitalize on this perception of predictability depend on the same state variables.

More recently work by Lo (2004) argues that there is no innate mechanism that guarantees that the market will adjust instantaneously to new and changing economic circumstances. Arguing for a more evolutionary approach to the market, he proposes an Adaptive Expectations Hypothesis according to which change leads to profit opportunities which erode through time as market participants learn to take advantage of them. One prediction of this model is that these profit opportunities will give rise to predictability in the markets that will be localized and limited in scope and duration. This elusive return predictability appears to be confirmed in the empirical results of this paper.

How can we reconcile this evidence with the widespread perception among professional investors that market returns are in fact predictable and that there are considerable fortunes to be made predicting the market? Much of the perception of predictability comes from the undue faith many practitioners have in backtesting. Overfitting the past history of returns is one of the occupational hazards of the business, and brings to mind the aphorism of G. K. Chesterton that "ten false philosophies will fit the universe"[1] *Ex post* conditioning can also give rise to the

[1]"The Honour of Israel Gow" in *The Innocence of Father Brown* (1911)

perception of predictability as the practitioner is necessarily a prisoner of history[2]. In each case as this paper shows, the comparison of models out-of-sample provides a healthy antidote to the false perception of predictability.

This paper shows that simple models based on the past history of equity returns and (principal components of) macroeconomic innovations find it difficult predict returns, at least according to standard and well established statistical criteria. Indeed, the paper considers at different points at least four different criteria of predictability, including the Bayes Information criterion, RMSE out of sample, backward looking out of sample R^2, and the sign criterion proposed by Peseran and Timmermann (1992). However, predictability is neither necessary nor sufficient to establish the profitability of a trading strategy based on the same information. Advances in neural net technology cannot exclude the possibility that there is some arcane pattern recognition algorithm used by a successful seat-of-the-pants trader that might dominate an overparameterized neural net procedure in an out-of-sample exercise. Predictability if detected may not be profitable if the

[2]Brown, Goetzmann and Ross (1995) discuss the role of *ex post* conditioning in giving rise to a false perception of predictability. The success of many well-known hedge fund managers (see for example "Make Less Than $240 Million? You're Off Top Hedge Fund List"*New York Times* April 24, 2007) can be attributed to many factors, including the willingness to take great risk on behalf of their wealthy clients providing liquidity to the markets. In another example of *ex post* conditioning, the great wealth of successful hedge fund managers is widely reported. Not so widely reported are those who take great risk and lose.

range of variation in the predictable component of returns matches the predictable rise and fall of the cost of funds used to exploit this predictability.

It is understandable why statisticians prefer predictability to profitability. It is dangerous to draw statistical inferences from the pattern of trading returns where the distribution of those returns depends heavily on the strategy used to exploit apparent predictability. The interpretation that an extreme tail event brought down Long Term Capital Management was based on a parametric representation of the process generating trading profits (Lowenstein 2000). Indeed Goetzmann *et al*. (2004) show that traders have a positive incentive to vary the statistical distribution of trading payoffs to influence the performance metrics by which they are judged.

In a quite remarkable early paper Cowles (1934) addressed both the issue of how to characterize the pattern recognition algorithm of seat-of-the-pants traders as well as the technology which might be used to examine the statistical significance of resulting trading profits[3]. William Peter Hamilton, the editor of the *Wall Street Journal* from 1902 to 1929 had a reputation for successful forecasting established over a long period of years. Experts analyzed his editorials on the state of the markets and concluded that he had essentially recommended buying into the market 140 times, selling the market 41 times and holding the market 74 times. While this strategy earned 12 percent per annum, this strategy lost 3.5 percent per annum relative to simply buying and holding stocks comprising the Dow Jones Industrial Average over the period. The same analysis applied

[3]These results are discussed in Brown Goetzmann and Kumar (1998)

to the recommendations of 24 financial publications yielded similar results. But were these losses statistically significant? Cowles proposed a novel bootstrap in strategy space to examine this question. In one instance, a forecaster had 240 weeks of experience. 239 cards were made up which recorded the trades implied by their forecasts, and were drawn at random to create a hypothetical trading record with the same frequency of buy and sell operations. The best forecasters had a realized performance insignificantly different from the hypothetical bootstrap strategy, while the worst performance was significantly worse than the random trading strategy. Cowles therefore concludes with this paper that predictability is indeed elusive[4].

[4]The use of the Peseran and Timmermann (1992) sign test comes closest to an analysis of trading rule profitability. While we did not attempt to replicate precisely the experiments reported in this paper, we found similar results over the same period of time using a VAR based on past returns augmented by Fama and French (1992) factors where the model was estimated for each month based on prior month data and lags chosen each month according to a Bayes Information Criterion. For the period January 1970 to December 2005 the resulting Peseran and Timmermann (1992) p-value was 4.83% suggestive of predictability, and the performance was superior to that of the S&P500 over the same period with a Sharpe ratio (Sharpe (1966)) of .063 relative to the S&P500 over the same period of 0.048. The Cowles bootstrap procedure with one million replications using the same frequency of buy and sell decisions, yielded a p-value of 6.32% for this performance, very similar to the Peseran and Timmermann number.

References:

Stephen Brown, William Goetzmann and Stephen Ross (1995), “Survival” *Journal of Finance* 50, 853-873.

Stephen Brown, William Goetzmann and Alok Kumar (1998), “The Dow Theory: William Peter Hamilton's Track Record Re-Considered” *Journal of Finance* 53 1311-1333

Eugene Fama and Kenneth French (1992), “The Cross-Section of Expected Stock Returns” *Journal of Finance* 47, 427-465.

William Goetzmann, Jonathan Ingersoll, Matthew Spiegel and Ivo Welch (2004), "Portfolio Performance Manipulation and Manipulation-Proof Performance Measures" Yale ICF Working Paper No. 02-08 Available at SSRN: http://ssrn.com/abstract=302815

Martin Lettau and Sydney Ludvigson (2001) “Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying” *Journal of Political Economy* 109 , 1238-1287

Andrew Lo (2004) “The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective”, *Journal of Portfolio Management* 30, 15-29.

Roger Lowenstein (2000), *When genius failed: The rise and fall of Long Term Capital Management*. New York: Random House.

Hashem Pesaran and Allan Timmermann (1992), “A Simple Nonparametric Test of Predictive Performance” *Journal of Business and Economic Statistics* 10, 461-465.

William Sharpe (1966). “Mutual Fund Performance” *Journal of Business*, 39, 119-138.