The Dow Theory: William Peter Hamilton's Track Record Reconsidered

Stephen J. Brown
New York University Stern School of Business

William N. Goetzmann
Yale School of Management

Alok Kumar*
Yale School of Management

March 13, 1998

Abstract: Alfred Cowles' (1934) test of the Dow Theory apparently provides strong evidence against the ability of Wall Street's most famous chartist to forecast the stock market. We review Cowles' evidence and find it supports the contrary conclusion. Cowles analyzed editorials published by the chief exponent of the Dow Theory, William Peter Hamilton. We find that Hamilton's timing strategies actually yield high Sharpe ratios and positive alphas for the period 1902 to 1929. Neural net modeling to replicate Hamilton's market calls provides interesting insight into the Dow Theory and allows us to examine the properties of the Theory itself out-of-sample.

Introduction

Alfred Cowles' (1934) test of the Dow Theory apparently provides strong evidence against the ability of Wall Street's most famous chartist to forecast the stock market. Cowles' analysis is a landmark in the development of empirical evidence about the informational efficiency of the market. He claims that market timing based upon the Dow Theory results in returns that lag the market. In this paper, we review Cowles' evidence and find that it supports the contrary conclusion - the Dow Theory, as applied by its major practitioner, William Peter Hamilton, over the period 1902 to 1929, yields positive risk adjusted returns. The difference in the results is apparently due to the lack of adjustment for risk. Cowles compares the returns obtained from Hamilton's market timing strategy to a benchmark of a fully invested stock portfolio. In fact, the Hamilton portfolio, as Cowles interprets it, is frequently out of the market. Adjustment for systematic risk appears to vindicate Hamilton as a market timer.

To estimate the risk adjusted returns that may have been obtained by following the Dow Theory over the Hamilton period, we classify the market forecasts he made over 255 editorials published in *The Wall Street Journal* during his tenure as editor. Using the riskless rate as a benchmark, we find that Hamilton's ratio of correct to incorrect calls was higher than would be expected by chance. Using total return data for the Cowles index of stock market returns and the S&P index over the 27-year period, we find that the systematic risk of a trading strategy proposed by Cowles based upon editorials published in *The Wall Street Journal* is relatively low. We apply market timing measures used to identify skill to the time-series of returns to the Hamilton strategy, and we find significant positive evidence. An event-study analysis of the Dow Industrial Index around Hamilton's editorials shows a significant difference in mean returns over a 40-day period following "Bull" vs. "Bear" market calls. The event study also shows that Hamilton's forecasts were based upon a momentum strategy. Our finding suggests a plain reason why the Dow Theory remains to this day a popular method for timing the market. During the first three decades of this century it appeared to work. Regardless of whether it has worked since then, this early success established a reputation that has endured for decades.

Although the Dow Theory has outlived most Wall Street analysts, and has been taught to generations of undergraduate investment students[1], it may come as some surprise to learn that the Dow Theory was never written down by Charles Dow. Hamilton claimed to base his market calls on the authority of his eminent predecessor as Editor, Charles Dow. It was left to others to infer the nature and content of the Dow Theory from an analysis of his editorials in *The Wall Street Journal* and his other writings. Does an analysis of Hamilton's calls justify our belief that there was indeed a coherent Theory that we have come to understand as Dow's? To investigate this issue, we develop predictive models for Hamilton's bull and bear market forecasts. A stepwise regression model provides a multivariate linear approximation to the Dow Theory. We also train a neural net on the Hamilton editorials. The results confirm the interpretation of the Dow Theory as a modified momentum strategy. This neural net produces an "automaton" which we then apply to out-of-sample forecasts from the period 1930 to the present. Preliminary results show that the Dow Theory may have continued to work after Hamilton's death.

This paper is organized as follows. The next section provides historical background on the Dow Theory and William Peter Hamilton. Section II describes the empirical test of the Dow Theory published by Alfred Cowles in 1934, and discusses its interpretation in light of current methods of risk adjustment. Section III describes our re-analysis of the Hamilton editorials, section IV reports the statistical modeling of Hamilton's editorials, and section V concludes.

I. William Peter Hamilton and the Dow Theory

Most of what we know of the Dow Theory of stock market movements comes not from the founding editor of *The Wall Street Journal,* Charles Henry Dow, but from his successor, William Peter Hamilton, who assumed the editorship of the paper upon Dow's death in 1902. Over the next 27 years until his own death in late 1929, Hamilton wrote a series of editorials in *The Wall Street Journal* and in *Barron's,* discussing and forecasting major trends in the U.S.

stock market. Hamilton cited his predecessor Charles Dow's theory of stock market movements as the explicit basis for market predictions. In his 1922 book *The Stock Market Barometer,* Hamilton further explains the basic outlines of the theory. The theory presupposes that the market moves in persistent "Bull" and "Bear" trends. While determination of these trends is hampered by short-term deviations, Hamilton asserts that "charting" past fluctuations in the industrial and transportation indices allows the analyst to identify the primary market movement.

An acute irony, given the current reputation Dow theorists enjoy among financial economists, is that Hamilton's book succinctly articulates and defends the concept we now term informational efficiency of the stock market. According to Hamilton, "The market movement reflects all the real knowledge available…" This assertion is interpreted by a later prominent Dow theorist, Robert Rhea, in 1932, to mean that:

> **The Averages Discount Everything:** The fluctuations of the daily closing prices of the Dow-Jones rail and industrial averages afford a composite index of all the hopes, disappointments, and knowledge of everyone who knows anything of financial matters, and for that reason the effects of coming events (excluding acts of God) are always properly anticipated in their movement. The average quickly appraise such calamities as fires and earthquakes[2].

How, then, could the theory be consistent with the notion that past market trends are predictive of future price movements? According to Hamilton, "...the pragmatic basis for the theory, a working hypothesis, if nothing more, lies in human nature itself. Prosperity will drive men to excess, and repentance for the consequence of those excesses will produce a corresponding depression." In other words, the bull and bear market cycles envisioned by the Dow Theory are due to "the irrational exuberance" of individual investors, which itself appears not to be rationally incorporated into prices. This assertion is one of the three main axioms[3] of the Dow Theory, as interpreted by Hamilton and Rhea.

While the basic outlines of the Dow Theory may be gleaned from Hamilton's book and editorials, Robert Rhea's reduction of the Dow Theory as "theorems"[4] is a useful guide. The main theorem states that the market movements may be decomposed into primary, secondary and tertiary trends, the most important of which is the primary trend. Primary trends are further classified as Bull and Bear markets, both of which are characterized by fundamental economic activity as well as market price changes. Bull markets have three stages: "first...[is]...revival of confidence in the future of business...second is the response of stock prices to the known improvement in corporation earnings, and the third is the period when speculation is rampant and inflation apparent." For primary bear markets, "the first represents the abandonment of the hopes on which the stocks were purchased at inflated prices; the second reflects selling due to decreased business and earnings, and the third is caused by distress selling of sound securities, regardless of their value[5]."

The Dow Theory is translated into a guide to market timing by Hamilton by identifying the primary trend through a few key signs. First, trends must be confirmed by both the industrials and the transportations, but the confirmation need not occur on the same day. In other words, market movements are unreliable unless evidenced across two different market sectors. Second, extended movements sideways, called "lines," presage the emergence of a definite trend. In other words, a big move following a period of quiescence is taken as the beginning of a primary trend in that direction.

These "theorems" are vague enough to admit a variety of statistical interpretations, Hamilton's fellowship in the Royal Statistical Association notwithstanding. Fortunately, we have a specific record of forecasts he made over his lifetime. These forecasts were compiled and published by Robert Rhea in 1932, and published by *Barron's*. While not cited in his references, this source is likely the one used by Alfred Cowles III in his analysis of the Dow Theory.

II. Alfred Cowles' Analysis of the Dow Theory

Alfred Cowles' article "Can stock market forecasters forecast?" was published in *Econometrica* in 1934, and is widely regarded as a landmark paper in the development of the efficient market theory. In the paper, Cowles tests the Dow Theory by coding each of Hamilton's editorials in *The Wall Street Journal* or *Barron's* as "Bullish," "Bearish" or "Neutral." Cowles then assumes that on a "Bullish" signal, an investor places all of his wealth in stocks (50 percent in the stocks comprising the Dow Industrial Index and 50 percent in those comprising the Dow Transportation Index). A "Bearish" signal is taken as a recommendation to short the market and a "Neutral" signal is taken as a recommendation to invest in a riskless asset. Cowles adjusts the Dow Index for splits and dividends and estimated transaction costs, to calculate total returns to the Dow timing strategy. For periods Hamilton is out of the market, Cowles assumes he earns a riskless rate of 5 percent. He then compares this strategy with the alternative of investing 100 percent in the stock market over the same period. He concludes that the Dow Theory would have yielded 12 percent per annum, while an all-stock portfolio would have yielded 15.5 percent per annum. He regards this as prima facie evidence that the Dow Theory does not work.

Despite Cowles' careful work at calculating total returns for the two strategies, he neglects to adjust for differences in relative risk. These differences in fact appear to have been substantial. According to Cowles, "Hamilton was long of stocks 55 per cent, short 16 per cent, and out of the market 29 per cent, out of the 26 years under review." These numbers suggest that the systematic risk of the strategy is very different from 100 percent. Indeed, using the crude approximation for the average beta of $0.55 - 0.16 = 0.39$, it seems that the Dow strategy earned a risk adjusted return of $0.12 - [0.05 + 0.39(0.155 - 0.05)] = 0.029$. In other words, Cowles' interpretation of Hamilton's strategy seems to earn 290 basis points per year on a risk adjusted basis.

Cowles also performs a nonparametric analysis of the Hamilton recommendations, reporting the frequency of correct "Bull" and "Bear" market calls. Out of the 255 forecasts, he takes only the *changes* in recommendations as data. Thus he analyzes 29 "Bullish" forecasts, 23 "Bearish" forecasts and 38 "Neutral" forecasts. He concludes from this that half the changes in position were profitable, and half were unprofitable. The inescapable conclusion of this analysis is that an investor might just as well have flipped a coin. Or would he? Note that Cowles neglects to consider the efficacy of repeated bull forecasts in a rising market and repeated bear forecasts in a falling market. Any sequence of positive calls confirmed by a rising market would be reduced to a single datum. Given that the Dow Theory is essentially a momentum strategy, this possibility is not remote. Consider an extreme example. Suppose that Hamilton had made 100 forecasts : 49 bull forecasts in a row that proved correct, and then an incorrect bull forecast, then 49 correct bear forecasts in a row, then an incorrect bear forecast. Cowles would have scored this as two correct forecasts and two incorrect forecasts. However, an investor following that advice might have done quite well. The very fact that Cowles analyzes only 90 changes in position out of 255 forecasts in a momentum-based strategy suggests that some significant percentage of the remaining 165 forecasts may have been correct.

Of course, we cannot blame Cowles for not knowing in 1934 how to calculate Jensen's alpha, nor should we have expected him fully to appreciate the subtleties of conditioning in nonparametric tests. Nevertheless, a close look at the Cowles evidence suggests that the Dow Theory, as practiced by William Peter Hamilton, merits reconsideration.

III. Analysis of the Hamilton Editorials

To evaluate Hamilton as a market timer, we code the 255 Hamilton editorials as "Bullish", "Bearish", "Neutral" or indeterminate. We then collect total return information on the U.S. stock market over that period, and perform parametric and nonparametric tests of trading strategies analogous to those evaluated by Cowles. Finally we examine the price dynamics of the Dow Industrials around editorial publication dates.

A. Hamilton's Editorials

Unfortunately, the recommendations in the editorials are not always clear. Cowles' solution is to have five subjects score the editorials and then take the majority opinion on each. We use only one subject to score the editorials and find eleven indeterminate cases out of the 255, and eliminate them from the study. We calculate that the portfolio is in stocks 46 percent of the time, in bills 38 percent of the time and short 16 percent of the time. These percentages are based upon the number of months in each asset. When we count the number of "Bull", "Bear" or "Neutral" calls, the ratios are much closer to Cowles': long 54 percent, neutral 24 percent and short 22 percent. Our scoring therefore appears slightly different from the Cowles analysis, which has the portfolio long more frequently. As we show in the following analysis, it is unlikely that the minor differences in interpretation of the editorials are the basis for the divergence in our results.

B. Nonparametric Tests

To address the basic question of Hamilton's timing skill, we examine how often the Dow beats the riskless rate over the interval following an editorial, conditional upon a "Bull" or "Bear" call. The interval following the editorial is defined by the day following the editorial to the day of the next Hamilton editorial. Our analysis of the frequency of successful calls differs substantially from Cowles. Table I shows the relationship between market calls and subsequent performance. The proportion of successful "up" calls is greater than failed "up" calls and the proportion of successful "down" calls is much higher than failed "down" calls. In fact, Hamilton appears to have been extremely successful in his “Bear” market calls - he was right twice as often as he was wrong. In total, Hamilton was right 110 times and wrong 74 times, by our count. The neutral scores are not included in this analysis, since they are interpreted as stock returns equaling bill returns. The first panel of the table reports the results of the contingency table analysis. It shows strong evidence of association between Hamilton’s calls and subsequent market performance. The Fisher’s test[6] is statistically significant at the 1 percent level.

A natural test of the Dow Theory is the nonparametric Henriksson-Merton (HM) test. Developed for tests of timing ability, given a market forecast, the HM test effectively determines whether the manager provided a “put” on the market when it was valuable. The HM test is particularly appropriate in the Hamilton case because it uses the frequency of correct “Bear” market calls as the basis for determining timing success. This is important, because it explicitly conditions upon the frequency of down markets - down markets provide the only opportunity for a timer to manifest skill. The HM test gives compelling evidence that Hamilton was particularly effective in bear markets, and thus the proportion of correct “Bear” calls is much higher than would be anticipated by chance. The second panel of the table reports the expected number of correct calls under the null, and HM’s parametric approximation to the distribution of this value. The actual number lies more than three standard deviations above the benchmark.

One important issue is the implicit “I told you so” option that Hamilton had. Since we define the interval from editorial to editorial, Hamilton could simply have waited until the market confirmed his previous call, and then written an editorial claiming success. To address this issue, a different trading test in necessary.

C. Testing a Trading Strategy

Following Cowles, we simulate a trading strategy that moves from long stocks to short stocks to T-bills, depending upon the Hamilton editorial. While Cowles apparently uses a 50/50 portfolio mixture of the Dow industrials and the Dow railroads, we use the Cowles market index: a value-weighted index of U.S. stocks, including income return. Since this index ultimately became the basis for the S&P index, we will call it the S&P. This is widely considered the highest-quality monthly return series available, and mimics a passive strategy of holding stocks. As the alternative investment we use the prevailing short-term commercial paper rates. We further assume that the portfolio could only be rebalanced monthly, which allows us to use the monthly Cowles indices. Accordingly, we take the last recommendation that appeared in a month, and then assume that this is used as a guide to rebalancing at the end of the month. As a consequence, we do not pick up intramonth returns to the Hamilton strategy. The advantage is that we avoid any benefits that might have accrued to trading on daily trends and reversals that might have been possible.

Figure 1 shows the relative performance of the Hamilton portfolio compared with a portfolio invested entirely in the market over the 27 years. Notice that, for most of the period, the stock market drifts sideways, until a major bull market begins in 1924. The Dow Theory actually beats a full market investment until 1926, at which point the fully invested portfolio advances beyond the timing portfolio. Hamilton's major success occurs in 1907, when he avoids the worst of the panic of that year. He also does well in 1917 and 1920, when the portfolio is out of the market during both bear runs. Overall, the figure shows that the Hamilton portfolio is less volatile than the fully invested strategy.

The first column of Table II reports the results of the simulated investment strategy over the 27-year period. The annual arithmetic return to the Hamilton portfolio is 10.73 percent (10.71 percent geometric), almost indistinguishable from the annual average return obtained by holding the S&P all-stock portfolio, which yields an annual arithmetic average of 10.75 percent (10.38 percent geometric). On a risk adjusted basis, however, the Hamilton portfolio has a higher Sharpe ratio (0.559 compared with 0.456) and a positive Jensen measure of 4.04 percent -- more than 400 basis points per year. This high Jensen measure is due to a beta of 0.326 with respect to the S&P index.

D. Bootstrapping Tests

The rest of Table II reports the results of significance tests generated by bootstrapping the Hamilton strategy. The bootstrap is done in two different ways. In the first panel, test statistic distributions are generated by bootstrapping in the space of returns. We generate stock return series by drawing monthly returns with replacement from the S&P total return series over the sample period. Thus, we construct a null hypothesis that Hamilton has no forecasting ability, that the market follows a random walk, and that mean and variance for the market are constant. We report the mean, median, standard deviation, *t*-test and extreme simulated values (5 percentile values for standard deviations and 95 percentile values for other statistics). The Hamilton portfolio yields an unusually high annual return compared with the null. The expected return from such a strategy may be around 5 percent. The actual return of 9.95 percent ranks above the 99th percentile of the bootstrap distribution. While the standard deviation of the strategy is also low, it appears that the full-investment strategy also resulted in an unusually low standard deviation[7]. This appears to provide evidence against the random walk assumption of the bootstrap. The Sharpe measure of the Hamilton portfolio exceeds all of the bootstrapped values, and the Jensen measure of the Hamilton portfolio exceeds the 99 percent level. Neither the mean return nor the Sharpe ratio for the all-stock portfolio are unusual, although the low standard deviation puts the Sharpe ratio at the 63 percent level. Note that the standard deviation of the Hamilton Jensen measure is 1.97 percent. This means we cannot reject the joint null hypothesis that the that Jensen measure is zero and returns follow a random walk.

The second panel in Table II reports the results of a different form of bootstrap. Rather than destroying the time-series structure of stock returns over the period to construct a null, we randomize in the space of strategies, holding the market realization constant. The methodology was pioneered by Cowles himself, in another part of the landmark 1934 paper. In order to test whether a sample of investment newsletters had forecasting ability, he simulates a null of random stock selection (using a deck of cards!) and then compares the distribution of actual analyst performance records to those generated under a null that forecasts are simply random.

Inability to reject this null led Cowles to the conclusion that stock market forecasters could not forecast.

We apply this same procedure to the Hamilton forecasts to generate our null. We draw "Bull", "Bear", and "Neutral" forecasts with replacement from the actual Hamilton editorial series. We thus generate 500 simulated track records under a null that the editor was, in effect, flipping a coin, properly weighted so as to give the same expected proportions "Bull", "Bear", and "Neutral" forecasts as in the original series. The advantage of this is that we do not break the actual time-series characteristics of the market history itself. Our bootstrap in the space of strategies now conditions upon the true market realization. We do, however, alter the time-series characteristics of Hamilton's calls. While they no longer forecast future returns by construction, they also bear no relationship to past returns. They are no longer conditioned upon the time-series behavior of the market.

The result of bootstrapping in the space of strategies yields essentially the same result as bootstrapping in the space of returns. The alpha and Sharpe ratio are in the extreme tails of the bootstrapped distributions. We can clearly reject the null that Hamilton could have done as well by flipping (an appropriately weighted) coin.

E. Editorials as Events

Another measure of Hamilton's skill at market timing is to treat each editorial as an event, and examine whether "Bull" market calls are followed by positive market moves and "Bear" market calls are followed by negative market moves. We use event-study methods applied to the daily Dow Industrial Average data to examine the index dynamics around Hamilton's calls[8]. Figure 2 shows the price path for "Bull", "Bear" and "Neutral" calls. The paths represent the cumulated sum of the equal-weighted average appreciation return of the Dow Industrial Index over a window of 81 trading days: 40 days before publication date and 40 days following publication of the editorial. “Bull” calls are followed by a 1.5 percent price increase over the next 40 days on average, while “Bear” calls are followed by 1.74 percent price decrease over the next 40 days. The difference between these two, as measured by a two-tailed *t*-test allowing for unequal variance is significant at the 95 percent level (0.034 prob.value). The “Neutral” calls have a 0.21 percent return over the next 40 days.

The figure also indicates the basis for Hamilton’s calls. “Bear” calls follow steep recent declines in the Dow, while “Bull” calls follow recent positive trends. This is consistent with a theory of market trends. The result is clearly a momentum strategy, in which steep recent declines or advances are taken as signals of future trends in that direction.

IV. Recovering the Dow Theory

A. Step-wise Regression

Hamilton’s editorials give us a rare opportunity to recover the rules used by a successful Dow theorist. The issue of what exactly is the Dow theory has challenged market analysts virtually since the beginning of the century. With the series of Hamilton calls, and the technology of modern nonlinear statistical methods, we can attempt to understand the basis for the theory. Was Hamilton simply lucky or did he stick to basic rules? Were these rules consistent with his writings and the writings of others about the Dow theory? Can these rules help time the market today? To address these questions we develop a linear and a nonlinear model of Hamilton’s behavior over the 27-year period. We then see how the rules have performed in the period since Hamilton’s death. Table III reports the results of a stepwise regression for which the standard AIC criterion has been used to prune variables. The dependent variable in the regression is a Hamilton “Bear” call. In particular, we use continuous “Bear” calls rather than the handful of bear editorials. The reason for this is our presumption that the failure to make a “Bull” or “Neutral” call is equivalent to a continued forecast of a down market. Of course, this effectively treats each day of following a “Bear” call as an independent event for statistical purposes. Nevertheless, we throw away much potentially valuable information that enters the decision to be a bear when we concentrate on editorials alone.

Notice in Table III that the coefficients confirm the hypothesis that the Dow theory is a momentum theory. Decreases in the 60 day trends for the Industrials and Transportation indices forecast “Bear” calls. The indicator variable capturing whether the past 30 day returns have been the same sign for the two indices is positive, but not significant. The interaction between this variable and 60 day returns is significant, as are interactions between 60 day and 30 day lagged returns for the two indices. In fact, the significance of these inter-temporal interactions strongly suggests a nonlinear response of the decision to call a bear market with respect to past price dynamics. This result is consistent with the conventional wisdom regarding the content of the Dow theory in general and the axioms set forth in the Rhea(1932) volume in particular.

However, the significance of these nonlinear intertemporal interactions suggests that there may be possible interpretations of the Dow Theory yet to receive attention in the academic or practitioner literature. These potential interpretations are necessarily excluded from the stepwise regression procedure. Recent developments in artificial intelligence-based neural net procedures allow us to search over all possible patterns in the data that could conceivably be input to the Theory. In addition, these procedures allow us to construct a Hamilton automaton that can be used to examine the properties of the Dow Theory both during the Hamilton period and after he stopped writing his editorials.

B. Artificial Intelligence Methods for Detecting Patterns

Artificial intelligence-based procedures have become quite popular among practitioners who seek to identify recurring patterns in time series data generated in financial markets. The belief is that these patterns reflect the dynamics of the marketplace and that they allow us to predict future market movements. Such patterns generate trading rules. For example, a simple moving average (MA) trading rule states that, when the short-term (usually 1-5 days) moving average is greater than the long-term moving average (usually greater than 50 days) a rising market is indicated. The trading rule would then generate a buy signal. Such a trading rule would identify any pattern with a shape similar to the pattern shown in Figure 3.

Several studies provide support for the idea that these patterns can predict market movements. In a study by Brock, Lakonishok, and LeBaron (1992), the ability of two simple trading rules to predict the movements of Dow Jones Industrial Average (DJIA) were investigated using bootstrapping techniques. Allen and Karjalainen (1995) and Neely, Weller, and Dittmar (1997) use genetic programming to search for optimal trading rules in the foreign exchange market. Our research objectives and the modeling approach is different, though it is similar in spirit to Brock, Lakonishok, and LeBaron (1992) and to Neely, Weller, and Dittmar (1997).

Our goal in this paper is neither to identify a set of optimal trading rules nor to propose an autonomous trading system[9]. Instead, we test whether Hamilton's interpretation of the Dow Theory (based on 255 editorials that he wrote for the *The Wall Street Journal* during his lifetime) can predict stock market movements. In more general terms we want to uncover the rules of the Dow Theory (as interpreted by Hamilton) and understand its implications for the efficient market hypothesis.

In contrast to other studies, we have developed a prediction model that operates directly in the domain of patterns. We use the Feature Vector Analysis (FEVA) methodology developed in Kumar and McGee (1996) to forecast the state of Hamilton's recommendation. Feature vector analysis is particularly appropriate for modeling Hamilton's decision-making process, because it reduces the dynamics of past price series to trend shapes called features. These features amplify the "topological" characteristics of the data set and are typically cited by timers and technical traders as indicators of future market activity. These features include rising trends, falling trends, "head and shoulders," resistance levels and so on. A recurrent neural net algorithm uses these shapes as inputs and, through "training" on the 1902 through 1929 period, identifies a set of characteristic features related to the state of Hamilton's recommendation at any given time. The algorithm then develops a nonlinear function mapping of these features into a recommendation. Clustering of the preprocessed data allows us to identify dominant patterns in the data. In general, interpreting the model identified by the neural net is difficult. However, interpreting the prediction function learned by the neural net is possible. It is the identification of trend shapes through FEVA analysis that distinguishes our methodology from other modeling techniques.

The FEVA methodology used in our study can be extended in an obvious way for identifying "optimal" trading rules. Following the FEVA approach, there is no need to pre-specify the trading rules (or rule templates). A search is done directly in the domain of patterns. Such issues are currently being investigated and the results will be reported in subsequent work.

This methodology is closely related to artificial intelligence-based methods such as neural networks[10] (NN) and evolutionary computation[11] that have been the subject of much recent interest in finance. Neural network algorithms are statistical procedures to fit a reduced functional form to data. They are similar in intent to stepwise regressions in that they try a wide range of model specifications to reduce in-sample residual variation. But unlike stepwise regressions, their specification is not necessarily linear and the model form need not be explicitly specified. In fact, the innovation of NN models is that they offer a parsimonious but flexible nonlinear specification. Campbell, Lo, and MacKinlay (1997) provide a general overview of the applicability of neural net modeling to financial time series prediction problems[12]. Recent applications of neural net methods to financial markets suggest that nonlinear dynamics are potentially important characteristics of markets.

While current research shows that neural net models have the potential to uncover sophisticated nonlinear processes that lead to price changes in financial markets, they have limitations. They are simply good tools to fit in-sample -- indeed, given enough computer time and enough hidden layers, they can fit in-sample perfectly[13]. The resulting functions may have no economic interpretation. Unlike the stepwise regression above, it is difficult to check how the rules conform to the general understanding of the Dow Theory. Not only are there no "standard errors" about coefficients, there are no identifiable coefficients, since input data are recombined as intermediate variables in the hidden layers. Finally, as with all curve-fitting, there is no guarantee of out-of-sample performance. Though the neural net can recover sophisticated rules, no such consistent rules may be underlying the observed data. In the case of the Dow theory, for example, Hamilton's calls may be full of "sound and fury, signifying nothing." The FEVA methodology employed in our work (described in detail in the Appendix) attempts to eliminate these drawbacks associated with standard NN modeling procedure.

C. Feature Vectors

To identify the distinct shapes or visual cues that can be used to formulate market timing strategies in the DJIA time series, we classify features according to basic patterns of the DJIA time series. The means of half of the most influential 24 clusters are pictured in Figures 4 and 5. Each cell represents the daily DJIA trends used to predict a Hamilton call. Figure 4 shows the 12 most influential sell signals used by the algorithm. Figure 5 shows the 12 most influential buy signals used by the algorithm. Together, they represent a mapping of a visual sign into a recommendation -- presumably Hamilton used the stock level movements in this way, rather than applying a complex formula estimated with "post-modern" statistical techniques like neural nets.

The sell indicators are what we expected from the linear model. Recent down trends are signs to sell. We also find, however, that falls from recent peaks are strong sell indicators. One can also find in these patterns classic technical indicators such as "head and shoulders" forms. The buy indicators show forms that differ dramatically: besides upward slopes, it appears that recovery from recent declines is an indicator to get into the market.

The variety of influential clusters is instructive. Notice that there are not large differences across many of the basic shapes. Could Hamilton have discriminated among these different types? Can we reject the hypothesis that the first, say, five clusters are actually shapes drawn from the same distribution? It is doubtful on both counts. Not only does the clustering stage of our algorithm appear to split cases too finely, it also appears to identify unique cases too often, rather than searching for general rules. While many shapes represent basic forms like positive and negative trends, “U” shapes and “hump” shapes, some feature clusters appear noisy, or structureless. These must result from overfitting within sample. The structureless clusters are likely to be limited in applications out of sample and their presence in the set is an indication that the model is overfit to some degree. Without training the net on out-of-sample data, however it is difficult to fine tune the FEVA approach to focus only on general shapes.

D. In-Sample Performance

In the 1902 through 1929 period there were 3,599 “buy” calls, 1143 “sell” calls and 2,912 “neutral” calls. Compared with this, the NN trained over this interval predicts 4,464 sells, 469 buys and 2,721 neutrals for the in-sample data. One issue that is not clear from analysis of the Hamilton editorials is why there were long periods over which no forecasts were made. Notice that there are two long neutral periods. Are we to presume that he felt no updating of his previous predictions were necessary, or was he temporarily out of the business of forecasting the market? If the long hiatus in 1917 and 1918 were due to the latter, then the neural net would be training on periods in the sample with no information content. While it might fit these periods well in sample, this would introduce errors in the model.

E. Out-of-Sample Performance

We have an ideal holdout sample for evaluating whether the Dow Theory (or more properly, the Hamilton strategy) works to forecast trends in the market. We apply the NN model to predict calls for the period of 15 September 1930 - 1 December 1997, a total of 17,457 trading days. For the out-of-sample data, the NN predicts 10,004 buys, 6,131 sells and 1,322 neutral calls. Table IV shows the result of a timing strategy based upon the Hamilton model[14]. Instead of shorting the market on sell calls we assume the investor holds the riskless asset conditional upon a sell call. We do this under two assumptions. First we assume that the investor can act immediately on the call. This implies that it is possible to buy or sell at the opening stock prices and that these are the same as yesterday's closing prices. This strategy is the "Next Day Hamilton Strategy." As an alternative, we also consider what would happen if the investor could only trade at the close of the day that the forecast comes out. This is the "Second Day Hamilton Strategy". We consider this alternative because on days with big drops, like the 1987 crash, opening prices differ considerably from their previous close. Thus an investor who bought the paper, or saw the signal before the opening of the market, probably cannot take full advantage of the implied signal even absent trading costs or other frictions.

Note that this change makes a big difference. The Next Day Hamilton Strategy has a much higher return than the Second Day Hamilton Strategy. The Second Day Strategy has a return over the entire history almost exactly equal to the buy and hold return, and would be less than that after transaction costs. However, the strategy generates returns with less variance and lower systematic risk. Whether the substantial returns of the Next Day Strategy represent genuine daily persistence, or whether they are simply artefacts of nontrading (which is likely to be more important in the early years of the sample, when volume was lower for all stocks) is a subject for further investigation.

Table IV reports summary statistics for the three strategies over the entire out-of-sample period as well as decade by decade. It is clear that the strategy works best in periods of sharp market decline. The Second Day Strategy dominates the buy and hold even on an arithmetic return basis in the 1930s, the 1940s and the 1970s. Even the Same Day Strategy does not dominate the buy and hold in the 1980's. These results indicate that while the Dow Theory appears to have some power to predict returns in the post-sample period, normal trading frictions would preclude using the Theory to generate large excess returns, particularly in the most recent period. However, it is interesting to note that under the assumption of trading at the close of the editorial day, the results are quite comparable to those obtained during Hamilton's lifetime: returns close to the buy and hold with lower levels of risk. A more precise estimate of trading profits would depend on a realistic estimate of trading costs[15].

In sum, the results from the application of neural net estimation to William Peter Hamilton's 1902 to 1929 market forecasts suggest that the Dow Theory was more than a random decisionmaking process on the part of one editorialist. Our cluster analysis of the feature vectors used to successfully fit a model of Hamilton's forecasts in-sample suggests that he based his decisions on structures that resemble both persistent positive and negative trends, as well as positive and negative reversals. The market forecasts that our neural net algorithm developed appear to have some predictive power out of sample. Lack of reliable daily return data and trading cost data over the period since 1930 prevents us from precisely calculating return earned by following the Hamilton model vs. return to the buy and hold strategy. However, it appears that the strategy does reduce portfolio volatility and depending upon whether immediate execution of a sell signal is possible, it may enhance returns in some periods.

V. Conclusion

A review of the evidence against William Peter Hamilton's timing abilities suggests just the opposite -- his application of the Dow Theory appears to have yielded positive risk adjusted returns over a 27-year period at the beginning of the century. The basis of this track record seems to have been his ability to forecast bull and bear market moves. Whether this means that his interpretation of the Dow Theory is correct, or whether it simply means that Hamilton was one lucky forecaster among many market analysts, is still an open question. Given all of the financial periodicals published at the beginning of the century, it may not be surprising that one turned out to have been correct in calling market moves. Regardless of the issue of luck vs. skill however, it appears that Hamilton followed rules based upon observation of recent market trends that are recoverable by nonlinear estimation methods.

The contribution of this paper is not simply to show that Hamilton was a successful market timer. Alfred Cowles' (1934) analysis of the Hamilton record is a watershed study which led to the random walk hypothesis, and thus played a key role in the development of the efficient market theory. Ever since Cowles' article, "chartists" in general, and Dow theorists in particular, have been regarded by financial economists with skepticism. Our replication of Cowles' analysis yields results contrary to Cowles' conclusions. At the very least, it suggests that more detailed analysis of the Hamilton version of the Dow Theory is warranted. In broader terms it also suggests that the empirical foundations of the efficient market theory may not be as firm as long believed.

References

Allen, Franklin and Risto Karjalainen, 1998, Using genetic algorithms to find technical trading rules, *Journal of Financial Economics* Forthcoming.

Angeline, Peter, George Sanders and Jordan Pollack, 1994, An evolutionary algorithm that constructs recurrent neural networks, *IEEE Transactions on Neural Networks* 5, 54-65.

Arthur, Brian, John Holland, Blake LeBaron, Richard Palmer, and P. Tayler, 1997, Asset pricing under exogeneous expectations in an artificial stock market, in Brian Arthur, Steven Durlauf and David Lane ed: *The Economy as an Evolving Complex System II, , A Proceedings Volume in the Santa Fe Studies in the Sciences of Complexity* (Addison-Wesley Publishing Company, Reading, MA.).

Bodie, Zvi, Alex Kane and Alan Marcus, 1997, *Essentials of Investments Third Edition* (Irwin McGraw Hill, Boston, MA).

Brock, William, Josef Lakonishok and Blake LeBaron, 1992, Simple technical trading rules and stochastic properties of stock returns, *Journal of Finance* 47, 1731-1764.

Campbell, John, Andrew Lo and Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ.).

Casdagli, Martin and Stephen Eubank, 1992, *Nonlinear Modeling and Forecasting, A Proceedings Volume in the Santa Fe Studies in the Sciences of Complexity* (Addison-Wesley Publishing Company, Reading, MA.).

Cootner, Paul, 1964, *The Random Character of the Stock Market Prices*, (MIT Press, Cambridge, MA.).

Cowles, Alfred, 1934, Can stock market forecasters forecast? *Econometrica* 1, 309-324.

Elman, Jeffrey, 1990, Finding structure in time, *Cognitive Science* 14, 179-211.

Fogel, David, 1994, An introduction to simulated evolutionary optimization, *IEEE Transactions on Neural Networks* 5, 3-14.

Granger, Clive and Timo Terasvirta, 1992, Experiments in modeling nonlinear relationships between time series, in Casdagli, Martin and Stephen Eubank, ed: *Nonlinear Modeling and Forecasting, A Proceedings Volume in the Santa Fe Studies in the Sciences of Complexity* (Addison-Wesley Publishing Company, Reading, MA.).

Hamilton, William, 1922, *The Stock Market Barometer: A Study of its Forecast Value Based on Charles H. Dow's Theory of the Price Movement* (Barron's, New York, NY.).

Henricksson, Roy and Robert Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513-533.

Hertz, John, Anders Krogh and Richard Palmer, 1991, *An Introduction to the Theory of Neural Computation* (Addison-Wesley Publishing Company, Reading, MA.).

Hsieh, David, 1991, Chaos and nonlinear dynamics: Application to financial markets, *Journal of Finance* 46, 1839-1887.

Kohonen, Teuvo, 1995, *Self-Organizing Maps* (Springer Verlag, New York, NY.).

Kumar, Alok and Victor McGee, 1996, FEVA (Feature Vector Analysis): Explicitly looking for structure and forecastability in time series data, *Economic and Financial Computing*, Winter 1996, 165-189.

LeBaron, Blake, 1994, Nonlinear diagnostics and simple trading rules for high-frequency foreign exchange rates, in Andreas Weigend and Neil Gershenfeld, ed: *Time Series Prediction: Forecasting the Future and Understanding the Past* (Addison-Wesley Publishing Company, Reading, MA.).

McCullagh, Peter and John Nelder, 1983, *Generalized Linear Models* (Chapman and Hall, Ltd., Cambridge, England).

Michalewicz, Zbigniew, 1992, *Genetic Algorithms + Data Structure = Evolution Programs* (Springer Verlag, New York, NY.).

Mitchell, Melanie, 1996, *An Introduction to Genetic Algorithms* (MIT Press, Cambridge, MA.).

Neely, Christopher, Paul Weller, and Rob Dittmar, 1997, Is technical analysis in foreign exchange market profitable? A genetic programming approach, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Rhea, Robert, 1932, *The Dow Theory* (Barron’s, New York, NY.).

Scheinkman, Jose and Blake LeBaron, 1989, Non-linear dynamics and stock returns, *Journal of Business* 62, 311-337.

Smith, Murray, 1993, *Neural Networks for Statisticians* (Van Nostrand Reinhold, New York, NY.).

Wasserman, Philip, 1993, *Advanced Methods in Neural Computing* (Van Nostrand Reinhold, New York, NY.).

Weigend, Andreas, Bernado Huberman, and David Rumelhart, 1992, Predicting sunspots and exchange rates with connectionist networks, in Casdagli, Martin and Stephen Eubank, ed:

Nonlinear Modeling and Forecasting, A Proceedings Volume in the Santa Fe Studies in the Sciences of Complexity (Addison-Wesley Publishing Company, Reading, MA.).

Weigend, Andreas and Neil Gershenfeld, 1994, *Time Series Prediction: Forecasting the Future and Understanding the Past* (Addison-Wesley Publishing Company, Reading, MA.).

White, Halbert, 1989, Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings, *Neural Networks* 3, 535-549.

Appendix : FEVA Modeling

Our neural net prediction model employs four distinct methodologies. In the first part, Feature Vector Analysis is used for pre-processing the raw dataset (Kumar and McGee (1996)). Second, a clustering[16] stage identifies feature vector “types”. At that point, a recurrent neural network model is employed to build a nonlinear prediction model (Elman (1990)). An evolutionary programming technique to select the feature vector size and its components (Fogel (1994)). The main idea in FEVA is to generate a vector of input information as of time ‘t’ which captures how a data point is “embedded” in its surroundings. It is a snapshot of metric and nonmetric features of the local neighborhood of a given datum. FEVA components can be fundamental metric measures (time, given value of the datum, etc.), derived metric measures (differences, ratios, rate of change, etc.), categorical measures, and non-metric measures (ordinal properties). By introducing these additional attributes for each period, we effectively increase the dimensionality of the data set, and with increased dimensionality, we expect a richer set of features to be identifiable.

The evolutionary feature selection algorithm selected an “enhanced” dataset consisting of 49 features for feature vector analysis, the components of which include various lagged values of the DJIA price level, the DJIA return and an up/down indicator. All calculations are performed with respect to a time period (focus point) preceding the variable to be predicted. We set this period to 40 days before the call to forecast. Backward lags from the focus point consist of 1, 2, 3, 4, 5, 10, 20 and 40 days and forward lags of 1, 2, 3, 4, 5, 10, 20 and 40 days are formed. Altogether we use 81 data points for constructing the feature vector. The coded Hamilton's editorials, DJIA, and the DJTA form the training set for the NN model (total of 7735 data points). We lose 81 data points to build feature vectors and are left with a 7654 x 49 matrix as the training data. The dataset, enhanced by the feature vectors, is used to train a NN: 69-20-3 Elman recurrent neural net model[17.] In an Elman recurrent NN, there are additional time series in the input layer that are simply copies of the values from the hidden layer output. Altogether there are 49 inputs from the dataset and 20 additional series taken by the Elman model from the hidden layer to the input layer giving a total of 69 inputs in the first layer. The output layer consists of 3 vectors predicting buy, sell and neutral calls, each one of them a mutually exclusive binary variable. For example, an output of (1,0,0) is a call to "buy" whereas (0,1,0) is a "sell" signal. The NN is trained for 200 iterations and a SSE of 4967.87 is obtained. The SSE continues to improve after 200 iterations, but at a very slow rate. So to prevent overfitting, we stop the training after 200 iterations. We find that increasing the number of series in the hidden layer improves the SSE but has an adverse effect on the predictive performance (as expected) of the NN because the network had started to "overlearn".

Table I: Non-Parametric Test of Hamilton's Market Calls

This table reports the frequency of successful versus unsuccessful "Bull" and "Bear" market calls by William Peter Hamilton in his column in *The Wall Street Journal* and in *Barron's* over the period December, 1903 through November, 1929. Identification of "Call Up" and "Call Down" is based upon a reading of the editorial to determine Hamilton's assessment of whether the "primary movement" of the market was up or down. "Neutral" calls, and calls for which the direction could not be assessed from the editorial are omitted. "Market Up" and "Market Down" refer to whether or not the rate of capital appreciation of the Dow Industrial index(m) exceeds the riskless rate (rf) of five percent per annum. We report the *t*-test for the nonparamateric Henriksson-Merton measure of the number of expected correct calls conditional upon a bear market. Fisher's Exact Test is a test about the log-odds ratio log[(upup*downdown)/(downup*downdown)]. Under the null, the variance of log odds ratio is 1/upup + 1/downdown + 1/downup + 1/updown (see McCullagh and Nelder (1983) p.98 for details).

Contingency Table Test

	Market Up	Market Down	Column Sum
Call Up	74	56	130
Call Down	18	36	54
Row Sum	92	92	

Fisher's Exact Test Statistic: 8.74

Henricksson-Merton (HM) Nonparametric Test

Number when m < rf	N1	92
Number when m> rf	N2	92
Number of Observations	N	184
Number right when m<rf	n1	36
Number wrong when m>rf	n2	18
Number of "Bear" calls	n	54
Expected number right when m<rf	E(n1)	27
Standard deviation of n1		2.53
t-test for HM Test		3.56

Table II: Summary of Simulated Trading Strategy Based on Hamilton's Editorials

Statistics for the trading strategy are reported in Column 1. The strategy follows Cowles (1934) and assumes a short position in the stock market is taken at the end of the month in which a down call is made, while a long position in the market is taken at the end of the month in which an up call is made. Neutral calls are taken as a signal to invest in riskless securities. Randomizing returns bootstrap results are based upon 500 outcomes under a null in which market returns are i.i.d. Pseudo-histories of total monthly returns for the 27 year period are generated by random draws with replacement from the actual distribution of monthly returns. Randomizing strategies bootstrap results are based upon 500 outcomes of a null in which market forecasts are random. Pseudo-strategies are generated by drawing with replacement from the actual distribution of Hamilton forecasts with replacement. The median, mean, and standard deviation of bootstrap values are provided, together with a *t*-test for the significance of the difference between realized and bootstrap values, and percentile values (fifth percentile for standard deviation and 95th percentile for other statistics)

		Randomizing Returns: Bootstrap Results					
	Actual Values	Mean	Median	Standard deviation	*t*-test	Percentile values	Rank
Hamilton Beta	0.326	30.58%	30.88%	8.95%	0.220	0.448	0.561
Hamilton Annual Return	10.73%	5.32%	5.36%	2.10%	2.580	8.8%	0.995
Hamilton Std	10.44%	10.16%	10.15%	0.76%	0.378	8.9%	0.657
Hamilton Sharpe Ratio	0.559	0.045	0.043	0.208	2.468	0.401	0.990
Hamilton Jensen alpha	4.04%	-1.28%	-1.29%	2.11%	2.518	2.0%	0.996
S&P Annual Return	10.75%	11.09%	11.14%	2.39%	-0.140	15%	0.455
S&P Std.	12.83%	12.74%	12.68%	0.82%	0.111	11.5%	0.561
S&P Sharpe Ratio	.456	.491	.495	.199	-0.175	0.827	0.441

		Randomizing Strategies: Bootstrap Results					
	Actual Values	Mean	Median	Standard deviation	*t*-test	Percentile values	Rank
Hamilton Beta	0.326	30.93%	30.68%	9.98%	0.162	0.468	0.562
Hamilton Annual Return	10.73%	5.12%	5.07%	1.82%	3.090	8.2%	1.00
Hamilton Std. Deviation	10.44%	10.25%	10.29%	0.52%	0.369	9.3%	0.619
Hamilton Sharpe Ratio	0.559	0.023	0.016	0.178	3.012	0.328	1.00
Hamilton Jensen alpha	4.04%	-1.48%	-1.62%	1.94%	2.844	2.0%	0.997

Table III: Stepwise Regression of Hamilton's Bear Market Calls

This table reports the results of a step-wise regression of Hamilton's bear market calls on a number of variable constructed from the preceding values of Dow Industrial and Dow Transportation indices over the period 1902 through 1929. Industrial Index returns are specified unless the "tr" suffix indicates the transportation index. "Same sign" indicates that the past thirty day returns of the indiceshave the same sign. Colons indicate interactions among variables.

	Value	*t*-stat
(Intercept)	-1.72	-19.94
sixty.day.returns	-14.24	-6.67
sixty.day.returns.tr	-9.72	-6.38
thirty.day.returns	-3.70	-2.70
thirty.day.returns.tr	6.14	3.18
same.sign	0.12	1.30
sixty.day.returns:same.sign	12.19	5.19
sixty.day.returns:thirty.day.returns	-93.76	-6.45
thirty.day.returns:thirty.day.returns.tr	64.39	3.82
sixty.day.returns.tr:thirty.day.returns.tr	-235.84	-6.73
sixty.day.returns:thirty.day.returns.tr	100.69	4.21
sixty.day.returns.tr:thirty.day.returns	110.31	3.64
sixty.day.returns.tr:thirty.day.returns:thirty.day.returns.tr	-951.33	-3.96
sixty.day.returns:thirty.day.returns:thirty.day.returns.tr	345.69	2.73

Table IV: Summary Statistics for Out-of-Sample Performance

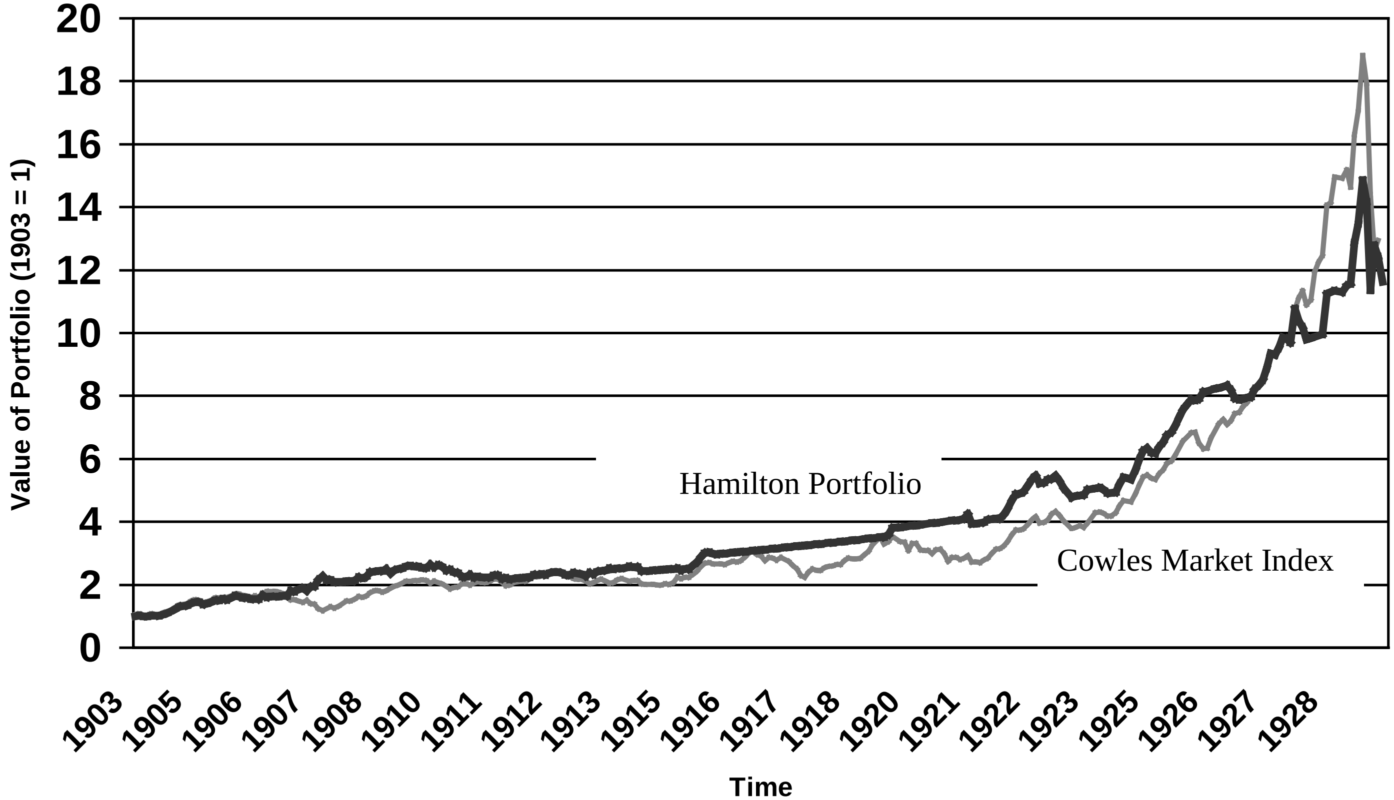
This table reports the returns two three investment strategies over the period September 1, 1930 through December 1, 1997. DJIA is the capital appreciation returns of the Dow Industrials, without dividends. "HamNext" are the returns to investing at the opening of the day of the out-of-sample Hamilton call generated by the neural net model. A neutral or buy call is taken as a signal to be fully invested in the Dow Jones Industrial Average. A sell call is take as a signal to be invested in cash, which is assumed to earn no interest. "Hamilton 2^{nd}" are the returns to investing at the close of the day prices on the day on which the forecast appears. Arithmetic returns only are reported for the sub-periods.

Whole Period

	Arithmetic Mean	Geometric Mean	Annual standard deviation
DJIA	7.07	5.48	18.30
HamNext	9.97	9.87	11.90
Hamilton2nd	5.91	5.52	12.10

Sub-Periods

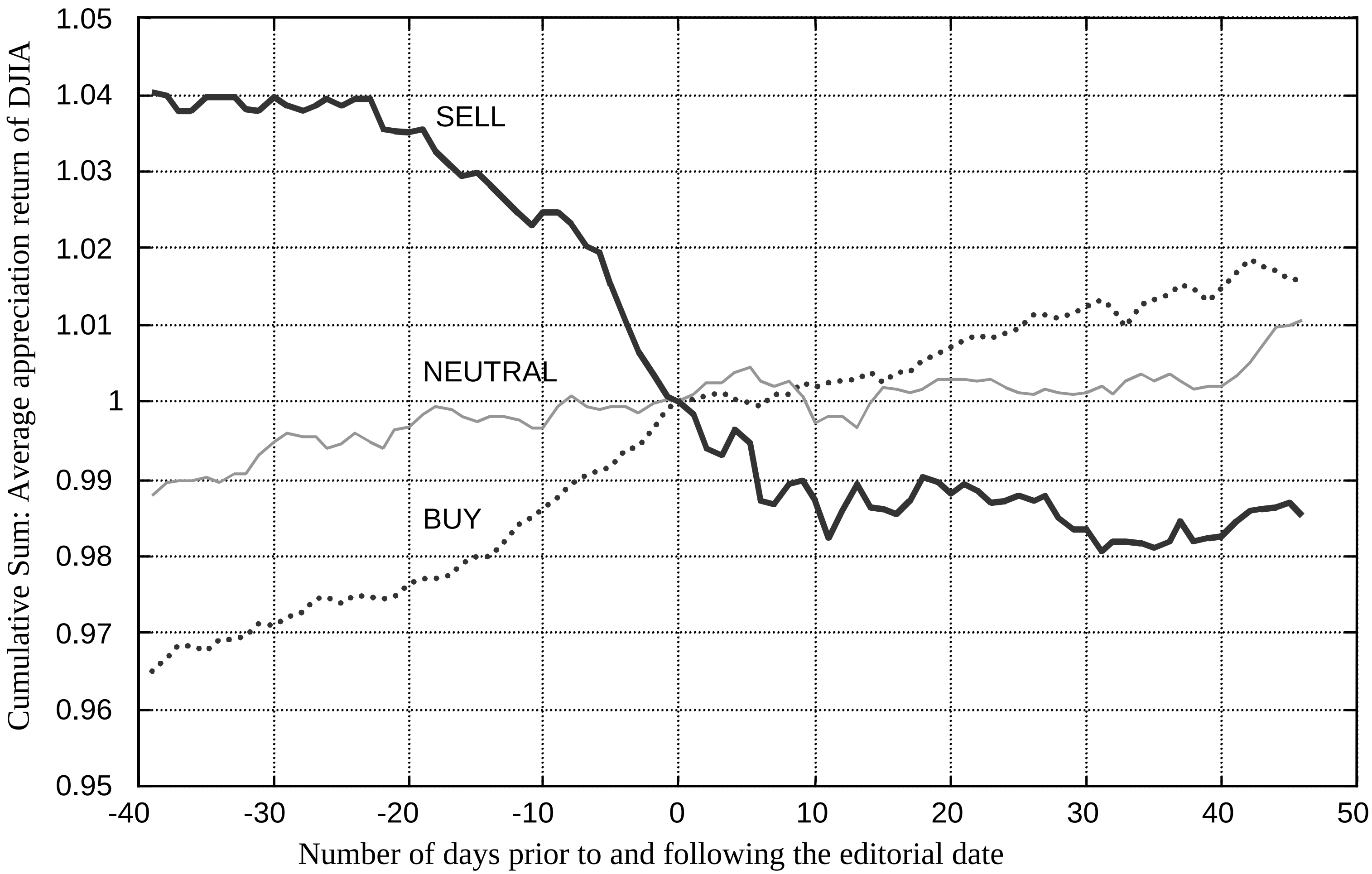
DJIA	Cap App	HamNext	Ham2nd
1930-39	1.477	11.10	2.43
1940-49	3.213	6.04	5.66
1950-59	9.641	9.91	5.27
1960-69	7.712	9.68	6.53
1970-79	0.409	6.74	4.30
1980-89	12.626	11.29	7.01
1990-97	15.442	16.24	10.72



20
18
16
14
12
10
8
6
4
2
0
Value of Portfolio (1903 = 1)
Hamilton Portfolio
Cowles Market Index
1903
1905
1906
1907
1908
1910
1911
1912
1913
1915
1916
1917
1918
1920
1921
1922
1923
1925
1926
1927
1928
Time

Figure 1: Hamilton Portfolio vs. 100% Stocks

Relative performance of the Hamilton portfolio compared to a portfolio invested entirely in the market over the 27 years. The figure indicates that the Hamilton portfolio was less volatile than the fully invested strategy.



Cumulative Sum: Average appreciation return of DJIA
1.05
1.04
1.03
1.02
1.01
1
0.99
0.98
0.97
0.96
0.95
-40
-30
-20
-10
0
10
20
30
40
50
SELL
NEUTRAL
BUY
Number of days prior to and following the editorial date

Figure 2: Cumulated value of Dow around editorial dates

This figure represents the cumulated sum of the equal-weighted average appreciation return of the Dow Industrial Index over a window of eighty-one trading days: forty days before publication date and forty days following publication of the editorial.

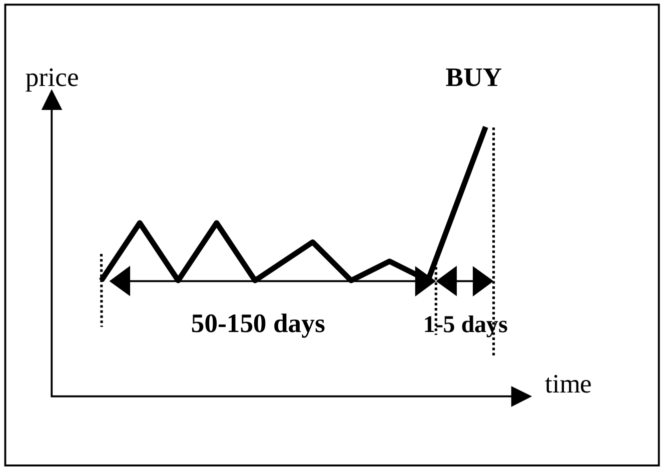

price
BUY
50-150 days
1-5 days
time

Figure 3

An example (schematic) of the type of pattern a simple moving average (MA) trading rule would detect.

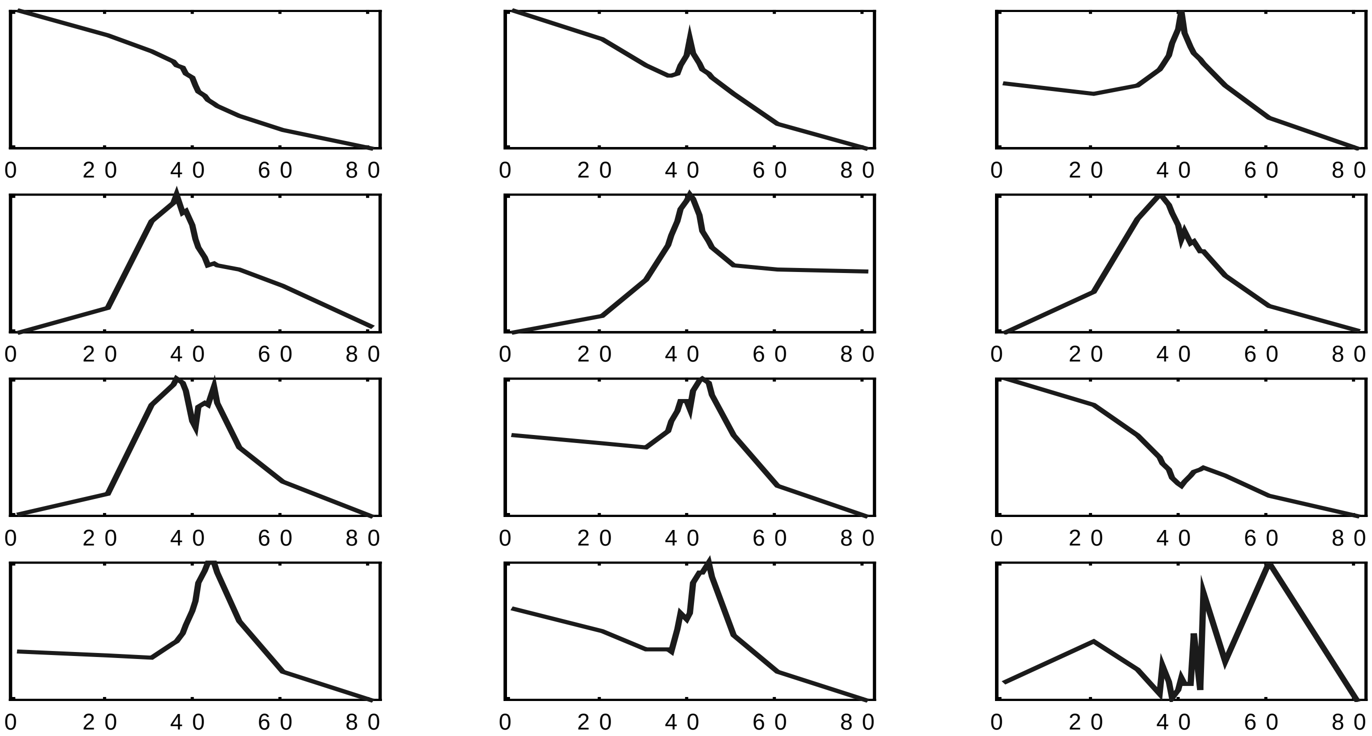

0 20 40 60 80

Figure 4: 12 Prominent SELL Indicators.

FEVA Plots representing the “structure” of the DJIA time series. We obtain a 100 cluster solution for the enhanced dataset (27774 x 49 matrix) and selected the 12 prominent FEVA SELL indicators from the solution. Patterns represent the price path of the DJIA over the eighty days preceding the forecast. The vertical axis is the price and the horizontal axis represents the number of days used to construct the feature vector.

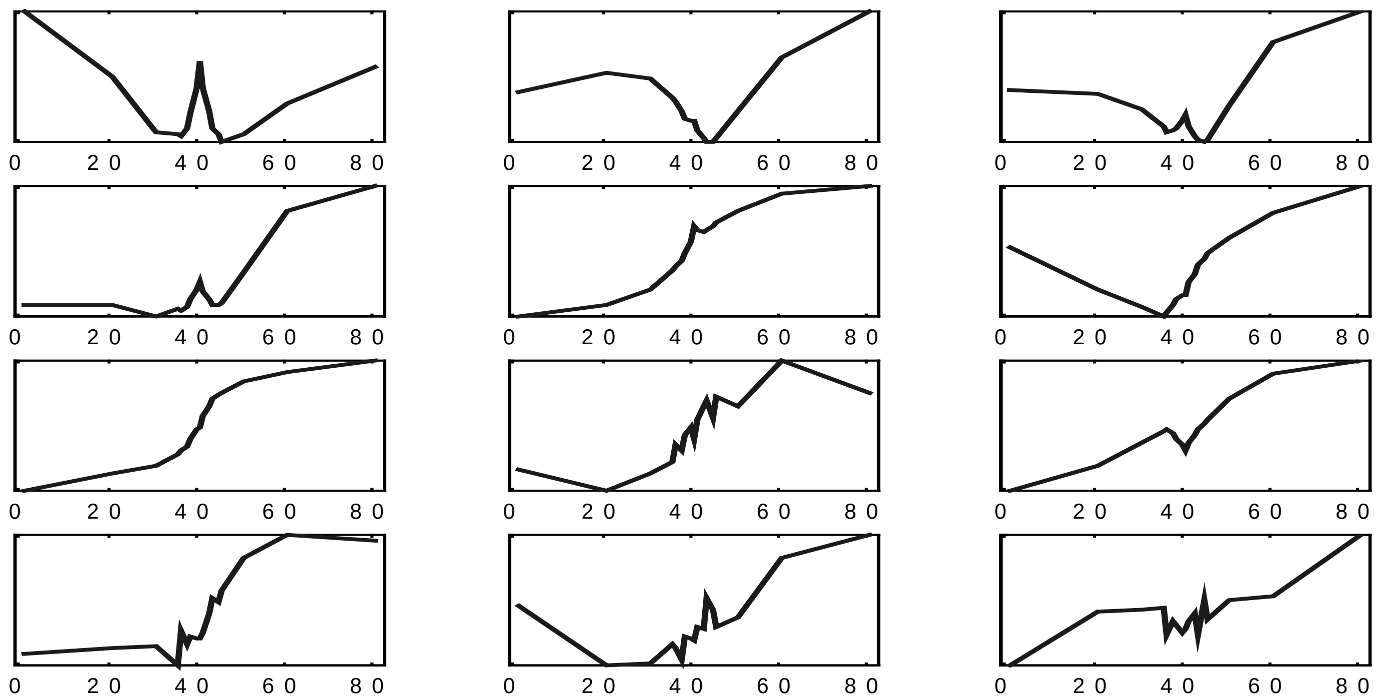

0 20 40 60 80

Figure 5: 12 Prominent BUY Indicators.

FEVA Plots representing the "structure" of the DJIA time series. We obtain a 100 cluster solution for the enhanced dataset (27774 x 49 matrix) and selected the 12 prominent FEVA BUY indicators from the solution. Patterns represent the price path of the DJIA over the eighty days preceding the forecast. The vertical axis is the price and the horizontal axis represents the number of days used to construct the feature vector.

Notes

* Acknowledgments: We thank Ned Elton, Ken French, Roger Ibbotson, Andrew Lo, N.R. Prabhala, Geert Rouwenhorst and participants in the faculty seminars at New York University and Yale School of Management for their comments

1. For a typical (albeit quite excellent) textbook treatment, see Bodie, Kane and Marcus (1997) pp. 414-417

2. Rhea (1932) p. 12

3. The other two axioms emphasize the existence of a primary trend in market movements and assert the fact that even though the theory is not infallible, it still is an invaluable aid for making speculations about the market movements.

4. Rhea(1932) proposed 12 theorems but only the relevant ones are discussed in our paper. There is a theorem relating price movements and trading volume - “Bull markets terminate in a period of excessive activity and begin with comparatively light transactions.” We do not study price-volume relationships in this paper.

5. Ibid. p.13

6. See McCullagh and Nelder (1983) p.98, for details of the Fisher's exact test.

7. This is consistent with the hypothesis that the market over this period displayed mean-reversion.

8. We use the Dow Industrial Index here in place of the Cowles market index because this value-weighted index of U.S. stocks including income return is not available on a daily basis for the period in question.

9. The results shown in our paper must be interpreted with caution. We have not considered transaction costs in our analysis and a successful implementation of our methodology must take this cost into account.

10. Neural net models are generalized, semiparametric, nonlinear function estimators. Smith (1993) is a good introductory book which concentrates on one type of neural networks, namely, feedforward neural networks (one of the most commonly used NN). The BASIC code for implementing a two-layered network is provided. Wasserman (1993) is an intermediate level book and provides a good introduction to different flavors of neural networks. Other learning paradigms are also discussed. Hertz, Krogh, and Palmer (1991) provide a clear and concise description of the theoretical foundations of neural networks using a statistical mechanical framework.

11. Evolutionary Computation refers to a set of algorithms inspired by the process of natural evolution. These algorithms try to simulate the “survival of the fittest” strategy - a key feature of natural evolution. A “parallel search” is performed in a multi-dimensional space (problem space) for optimal “good” solutions where the search process is probabilistic but not random. Four methodologies developed independently fall into this class: (1) Genetic Algorithms (GA), (2) Genetic Programming (GP), (3) Evolutionary Programming (EP), and (4) Evolutionary Strategies (ES). See Mitchell (1996) for an introduction.

12. A collection of papers that investigate the applicability of nonlinear and artificial intelligence-based techniques for time series prediction tasks appear in Casdagli and Eubank (1992) and Weigend and Gershenfeld (1994).

[13] In fact White (1989) has shown that theoretically a neural net model is capable of representing any nonlinear functional form.

[14] However, absent a realistic estimate of transaction costs over this extended interval, the numbers reported in Table IV can best be considered an upper bound on the performance generated by the Dow strategy.

[15] In addition, at the present time, we do not have a daily income return series for the Dow that would allow us to exactly calculate returns. To mitigate this, we make the assumption that income is not earned by the riskless asset. Thus, when dividend yields are close to the riskless rate, these effects should be offsetting. Furthermore, the DJIA is a price-weighted index, not value-weighted index, and thus it is not strictly investable without active reweighting. This may also affect the interpretation of the results to the extent that our summary statistics do not represent achievable returns.

[16] We use the clustering method developed in Kohonen (1995) for grouping features into a reduced set to examine the structural similarity of the DJIA series.

[17] Elman (1990) develops a neural net model that allows the intermediate series -- the so-called “hidden layer” to be used as original inputs. This simply allows for a richer potential nonlinear interaction among series and is well suited for time series analysis where the objective is to identify spatial and temporal patterns.