Department of Finance
Working Paper Series 1998

FIN-98-013

The Dow Theory: William Peter Hamilton's Track Record Re-Considered

Stephen J. Brown, William N. Goetzmann, Alok Kumar

February 17, 1998

This working paper series has been generously supported by a grant from

CDC Investment Management Corporation

The Dow Theory: William Peter Hamilton's Track Record Re-Considered

Stephen J. Brown
NYU Stern School of Business

William N. Goetzmann
Yale School of Management

Alok Kumar
Yale School of Management

January 23, 1998

Abstract: Alfred Cowles' (1934) test of the Dow Theory apparently provided strong evidence against the ability of Wall Street's most famous chartist to forecast the stock market. In this paper, we review Cowles' evidence and find that it supports the contrary conclusion - that the Dow Theory, as applied by its major practitioner, William Peter Hamilton over the period 1902 to 1929, yielded positive risk-adjusted returns. A re-analysis of the Hamilton editorials suggests that his timing strategies yield high Sharpe ratios and positive alphas. Neural net modeling to replicate Hamilton's market calls provides interesting insight into the nature and content of the Dow Theory. This allows us to examine the properties of the Dow Theory itself out-of-sample.

For a current version of this paper, please contact:
William N. Goetzmann
Yale School of Management
Box 208200, New Haven, CT 06520-8200
william.goetzmann@yale.edu. http://viking.som.yale.edu/

Acknowledgements: We thank Ken French, Roger Ibbotson, Andrew Lo, N.R. Prabhala, Geert Rouwenhorst and participants in the faculty seminars at NYU and Yale School of Management for their comments.

I. Introduction

Alfred Cowles' (1934) test of the Dow Theory apparently provided strong evidence against the ability of Wall Street's most famous chartist to forecast the stock market. Cowles' analysis was a landmark in the development of empirical evidence about the informational efficiency of the market. He claimed that market timing based upon the Dow Theory resulted in returns that lagged 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, yielded positive risk-adjusted returns. The difference in the results is apparently due to the lack of adjustment for risk. Cowles compared 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 interpreted it, was frequently out of the market. Adjustment for systematic risk appears to vindicate Hamilton as a market timer.

In order 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 the *Wall Street Journal* editorials was 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 indicates 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 suggest 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 which has endured for decades.

Despite the fact that 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 Wall Street Journal editorials 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 multi-variate 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 indicate 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 III describes the empirical test of the Dow Theory published by Alfred Cowles in 1934, and discusses it's interpretation in light of current methods of risk adjustment. Section V describes our re-analysis of the Hamilton editorials, section VI reports the statistical modeling of Hamilton's editorials, and section VII concludes.

II. 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. 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 elucidates the basic outlines of the theory. The theory pre-supposes 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:

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

> **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

---

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 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

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 seconds 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.

III. 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

---

comparatively light transactions." We do not study price-volume relationships in this paper.
5 Ibid. p.13

efficient market theory. In the paper, Cowles tests the Dow Theory by coding each of Hamilton's editorials in the *Wall Street Journal* or *Baron's* as "bullish", "bearish" or "neutral". Cowles then assumes that on a bullish signal, an investor places 100% of his wealth in stocks (50% in the stocks comprising the Dow Industrial Index and 50% 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 transactions costs, in order 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%. He then compares this strategy to the alternative of investing 100% in the stock market over the same period. He concludes that the Dow Theory would have yielded 12% per annum, while an all-stock portfolio would have yielded 15.5% 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 was a far cry from 100%. 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 would seem to earn 290 basis points per year on a risk-adjusted basis.

Cowles also performs a non-parametric 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 of 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 neglected 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 that were 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 to fully appreciate the subtleties of conditioning in non-parametric tests. Nevertheless, a close look at the Cowles evidence suggests that the Dow Theory, as practiced by William Peter Hamilton, merits re-consideration.

IV. Analysis of the Hamilton Editorials

In order 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 non-parametric tests of trading strategies analogous to those evaluated by Cowles. Finally we examine the price dynamics of the Dow Industrials around editorial publication dates.

IV.1 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% of the time, in bills 38% of the time and short 16% 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%, neutral 24% and short 22%. 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.

IV.2 Non-Parametric 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 of 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 1 shows a contingency table indicating 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% level.

A natural test of the Dow Theory is the non-parametric 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 issue of potential importance 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.

6 See McCullagh, P. and J.A. Nelder, 1983, p.98, for details of the Fisher's exact test.

IV.3 Testing a Trading Strategy

Following Cowles, we simulate a trading strategy which moves from long stocks to short stocks to t-bills, depending upon the Hamilton editorial. While Cowles apparently used 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 refer to it as the S&P. This is widely considered to be 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 re-balanced 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 re-balancing at the end of the month. As a consequence, we do not pick up intra-month 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 to 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. In general, the figure indicates that the Hamilton portfolio was less volatile than the fully invested strategy.

The first column of Table 2 reports the results of the simulated investment strategy over the 27 year period. The annual arithmetic return to the Hamilton portfolio is 9.95% (9.83% geometric), slightly below the annual average return obtained by holding the S&P all-stock portfolio, which yields an annual arithmetic average of 10.90% (10.54% geometric). On a risk-adjusted basis, however, the Hamilton portfolio has a higher Sharpe ratio (1.2 compared to 5.25) and a positive Jensen measure of 3.12% - 300 basis points per year. This high Jensen measure is due to a beta of .31 with respect to the S&P index.

IV.4 Bootstrapping Tests

The rest of Table 2 reports the results of significance tests generated by bootstrapping the Hamilton strategy. The bootstrap is performed 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, 95% (or 5% for standard deviations). The final column shows the rank represented by the actual value. The Hamilton portfolio yields an unusually high annual return compared to the null. The expected return from such a strategy appears to be around 5%. The actual return of 9.95% 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

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

the Hamilton portfolio exceeds all of the bootstrapped values, and the Jensen measure of the Hamilton portfolio exceeds the 99% 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% level. Note that the standard deviation of the Hamilton Jensen measure is 1.97%. This means we cannot reject the joint hypothesis null 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 simulated a null of random stock selection (using a deck of cards!) and then compared the distribution of actual analyst performance records to those generated under a null that forecasts were 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's could have done as well by flipping (an appropriately weighted) coin.

IV.5 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 and daily Dow Industrial Average data to examine the index dynamics around Hamilton's calls. 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 eighty-one trading days: forty days before publication date and forty days following publication of the editorial. Bull calls are followed by a 1.5% price increase over the next forty days on average, while bear calls are followed by 1.74% price decrease over the next forty days. The difference between these two, as measured by a two-tailed t-test allowing for unequal variance is significant at the 95% level (0.034 prob.value). The neutral calls have a 0.21% 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.

V. Recovering the Dow Theory

V.1 Step-wise Regression

Hamilton□s editorials provide us with 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 non-linear statistical methods, we can attempt to understand the basis for the theory. Was Hamilton simply lucky or did he adhere 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 non-linear 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 3 reports the results of a step-wise 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. Never-the-less, concentrating only on the editorials effectively throws away a lot of potentially valuable information about what constitutes a decision to be a "Bear".

Notice in Table 3 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 thirty day returns have been the same sign for the two indices is positive, but not significant. The interaction between this variable and sixty day returns is significant, as are interactions among sixty day and thirty day lagged returns for the two indices. In fact, the significance of these inter-temporal interactions is strongly suggestive of a non-linear response of the decision to call a bear market with respect to past price dynamics. Overall, the step-wise regression appears to confirm the conventional wisdom regarding the content of the Dow theory in general and the axioms set forth in the Rhea volume in particular.

V.2 Artificial Intelligence Methods for Detecting Patterns

Technical analysis (or charting) refers to a set of techniques that attempt to identify recurring patterns in time series data generated in the financial markets such as the stock market, futures market and the foreign exchange market. The belief is that these patterns reflect the dynamics of the marketplace and so on the basis of these identified patterns, predictions about the future market movements can be made.

Recurring patterns in the data can be detected using trading rules. For example, a simple moving average (MA) trading rule states that if the short-term (usually 1-5 days) moving average is greater than the long-term moving average (usually greater than 50 days), it is an indicator of a rising market and a buy signal should be issued. The MA trading rule would identify any pattern that is "topologically"[8] similar to the pattern shown in Figure 3. The important factors are the relative measures that define the shape of the pattern and the actual metric measures are not very important.

---

8 We do not use the word "topological" in a strict mathematical sense here. We mean "shape" without explicit metric.

Several studies provide support for the ability of technical analysis techniques to predict the market movements. In a study by Brock, Lakonishok and LeBaron (1992), the ability of two simple trading rules to predict the movements of 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, even though it is similar in spirit to Brock, Lakonishok and LeBaron (1992) and to Neely, Weller and Dittmar (1997), i.e., we also address the prediction problem. Our goal is neither to identify a set of optimal trading rules nor are we proposing an autonomous trading system[9]. In this study, we test if Hamilton's interpretation of Dow Theory (as manifested in 255 WSJ editorials that he wrote during his lifetime) has the ability to predict stock market movements. In more general terms we want to uncover the rules of Dow Theory (as interpreted by Hamilton) and we want to understand the implications of our results for the efficient market hypothesis.

Unlike 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 dataset and such features are precisely the kinds of dynamics typically cited by timers and technical traders as indicators of future market activity: rising trends, falling trends, "head and shoulders", resistance levels and so on. A recurrent

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.

neural net algorithm uses these shapes as inputs and, through "training" on the 1902 through 1929 period, identifies a set of characteristic features that are related to the state of Hamilton's recommendation at any given time, and then develops a non-linear function mapping these features into a recommendation. Clustering of the preprocessed data allows us to identify dominant patterns in the data. In general, it is difficult to interpret the model identified by the neural net but the preprocessing step, alongwith the clustering step, allows us to interpret the prediction function learned by the neural net. It is the data preprocessing step 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 will be no need to pre-specify the trading rules (or rule templates) and a search directly in the domain of patterns can be performed. Such issues are currently being investigated and the results will be reported in subsequent papers.

Many other authors have applied nonlinear and artificial intelligence (AI) based methods such as neural networks[10] (NN) and evolutionary computation[11], and in particular, there has been enormous interest in the past decade in the application of artificial neural network models to Finance. Neural network algorithms are statistical procedures to fit a reduced functional form to

---

10 Neural net models are generalized, semi-parametric, 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) provides 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.

data. They are similar in intent to step-wise regressions in that they try a wide range of model specifications to reduce in-sample residual variation. But unlike step-wise 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 non-linear specification. The interaction among input series is handled in NN models through the creation of intermediate series (a so-called "hidden layer" of synthesized data). These intermediate series are multiplicative and nonlinear transforms (through a combination of logistic functions) of the input data. The output is, in turn, a function of these intermediate series. This extra layer (or layers) of data processing allow even very sophisticated functional forms to develop as a result of estimation. Hutchison, Lo and Poggio (1994) find that a radial basis function neural net model recovered the Black-Scholes formula from option prices and inputs. Weigend (1992) uses a feedforward NN to predict currency exchange rates (DM vs USD). Campbell, Lo and MacKinlay (1997) provide a general overview of the applicability of neural net modeling to financial time series prediction[12] problems, and recent applications of neural net methods to financial markets suggest that non-linear dynamics are potentially important characteristics of markets.

AI based methods are guided by different sets of principles and they differ in the representation techniques they use, but at a fundamental level these methods perform a search in a space of some kind. For example, in neural network based models used for market forecasting, the nonlinear prediction function is represented using a set of nonlinear information processing units and the objective is to perform a search in the space of functions to obtain the "optimal",

---

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

nonlinear prediction function. The prediction function encodes the various trading rules. These rules are not easy to interpret. In contrast, evolutionary methods explicitly represent each of the potential trading rules (or a rule template, as in Arthur et al (1997)) in the form of a vector and a stochastic, parallel search in the space of rules is performed to obtain the "optimal" trading rule (or a set of semi-optimal trading rules).

While current research shows that neural net models have the potential to uncover fairly sophisticated non-linear 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 step-wise regression above, we might not be able to check how the rules conform to statements about the Dow Theory. Not only are there are no "standard errors" about coefficients, there are not any clear 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. Even though the neural net can recover sophisticated rules, there may be no such consistent rules 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 Appendices A and B) attempts to eliminate these drawbacks associated with standard NN modeling procedure.

V.4 Feature Vectors

---

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

In order 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 100 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 25 most influential sell signals used by the algorithm. Figure 5 shows the 25 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 of the shapes represent basic forms like positive and negative trends, "U" shapes and "hump" shapes, some of the feature clusters appear noisy, or structure-less. These must result from over-fitting within sample. The "structure-less" clusters are likely to be limited in applications out of sample and their presence in the set is an indication that the model is over-fit 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.

V.5 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 predicted 4,464 sells, 469 buys and 2,721 neutrals for the in-sample data. In Figure 6, the predicted calls are superimposed on the DJIA series where the color represent the call: green is a "buy" call, red is "sell" and blue represents a "neutral" call. It is quite clear that the NN model is able to learn the patterns in the Dow series. 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.

V.6 Out-of-Sample Performance

We have an ideal hold-out 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 predicted 10,004 buys, 6,131 sells and 1,322 neutral calls. Figure 7 shows the result of a timing strategy based upon the Hamilton model. 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, i.e. 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 indicated as 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 termed "Second Day Hamilton Strategy". The reason for this is that on days with big drops, like the 1987 crash, opening prices differed considerably from their previous close. Thus, even if an investor bought the paper, or saw the signal before the opening of the market, it might be impossible to take full advantage of the implied signal.

Note that this change makes a big difference. The Next Day Hamilton Strategy appears to be a dramatic improvement over the buy and hold. It has a higher return with lower risk. The Second Day Hamilton Strategy is less dominant. It has a return over the entire history almost exactly equal to the buy and hold. The major benefit is that the strategy has less variance and less systematic risk. Evidently, the major advantage afforded by the Hamilton neural net "automaton" is in forecasting the next day trends in the DJIA. Whether these trends represent genuine daily persistence, or whether they are due to non-trading (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.

These results indicate that 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. 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 re-weighting. This may also affect the interpretation of the results to the extent that our summary statistics do not quite represent achievable returns.

Table 4 reports summary statistics for the three strategies over the entire out-of-sample period as well 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 1930's, the 1940's and the 1970's. Even the same-day strategy does not dominate the buy and hold in the 1980's.

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 decision-making 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 technical trading rules that our neural net algorithm developed appear to have some benefit out of sample. Lack of reliable daily return data over the period since 1930 prevents us from precisely calculating returns earned by following the Hamilton model vs. returns 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.

VI. 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 non-linear estimation methods.

The contribution of this paper is not simply to show that Hamilton was a successful market timer. Alfred Cowles' 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 the 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 R. Karjalainen, 1995, "Using Genetic Algorithms to Find Technical Trading Rules", Working Paper, The Wharton School, University of Pennsylvania, October 1995.

Angeline, Peter J. G.M. Sanders and J.B. Pollack,1994, "An Evolutionary Algorithm that Constructs Recurrent Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 5, January No. 1, pp. 54-65.

Arthur, W. Brian, J.H. Holland, B. LeBaron, R. Palmer, and P. Tayler, 1997, "Asset Pricing Under Exogeneous Expectations in an Artificial Stock Market", in *The Economy as an Evolving Complex System II,* edited by W. B. Arthur, S.N. Durlauf and D.A. Lane, A Proceedings Volume in the Santa Fe Studies in the Sciences of Complexity, Addison-Wesley Publishing Company, Reading, MA.

Bodie, Z., A. Kane and A. Marcus, *Essentials of Investments* Third Edition, Irwin McGraw Hill, 1997

Brock, William J., J. Lakonishok and B.LeBaron, 1992, "Simple Technical Trading Rules and Stochastic Properties of Stock Returns, *Journal of Finance,* 47, 1731-1764.

Campbell, John Y., A.W. Lo and A.C. MacKinlay, *The Econometrics of Financial Markets,* Princeton University Press, 1997.

Casdagli, Martin and S. Eubank (editors), 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 H., 1964, *The Random Character of the Stock Market Prices,* MIT Press, Cambridge, MA.

Cowles, Alfred, 1934, "Can Stock Market Forecasters Forecast?" *Econometrica*, pp.309-324.

Elman, Jeffrey L., 1990, "Finding Structure in Time," *Cognitive Science*, Vol. 14, pp. 179-211, 1990.

Fogel, David, 1994, "An Introduction to Simulated Evolutionary Optimization," *IEEE Transactions on Neural Networks*, Vol. 5, January No. 1, pp. 3-14.

Granger Clive W.J. and T. Terasvirta, "Experiments in Modeling Nonlinear Relationships Between Time Series", in *Nonlinear Modeling and Forecasting*, Casdagli, M. and S. Eubank (editors), Addison--Wesley Publishing Company, 1992

Hamilton, William Peter, 1922, *The Stock Market Barometer: A Study of its Forecast Value Based on Charles H. Dow□s Theory of the Price Movement*, Barrons, New York.

Henricksson, Roy D. and Robert C., 1981, "On Market Timing and Investment Performance. II.

Statistical Procedures for Evaluating Forecasting Skills," *Journal of Business*, 54:4, pp.513-533.

Hertz, J., A. Krogh and R.G. Palmer, 1991, *An Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company.

Hsieh, David, 1991, "Chaos and Nonlinear Dynamics: Application to Financial Markets," *Journal of Finance*, 46(5) pp. 1839-1887.

Kohonen, Teuvo, 1995, *Self-Organizing Maps*, Springer Verlag Berlin, 1995.

Kumar, Alok and V.E. McGee, 1992, "FEVA (Feature Vector Analysis): Explicitly Looking for Structure and Forecastability in Time Series Data," *Economic and Financial Computing*, pp.165-189, Winter 1996.

LeBaron, Blake, 1994, "Nonlinear Diagnostics and Simple Trading Rules for High-Frequency Foreign Exchange Rates", in *Time Series Prediction: Forecasting the Future and Understanding the Past,* edited by A. Weigend and G.A. Gershenfeld, Addison-Wesley Publishing Company, Reading, MA, pp. 457-474.

Scheinkman, Jose' and B. LeBaron, 1989, "Non-linear Dynamics and Stock Returns," *Journal of Business* 62, 311-337.

McCullagh, P. and J.A. Nelder, 1983, *Generalized Linear Models*, Chapman and Hall, Ltd., Cambridge, England.

Michalewicz, Z., 1992, *Genetic Algorithms + Data Structure = Evolution Programs*, Springer Verlag Berlin, 1992.

Mitchell, Melanie, 1996, "An Introduction to Genetic Algorithms", MIT Press, Cambridge, MA.

Neely, C., P. Weller, and R.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.

Smith, Murray, 1993, *Neural Networks for Statistician*, Van Nostrand Reinhold, NY.

Wasserman, Philip D., 1993, *Advanced Methods in Neural Computing*, Van Nostrand Reinhold, NY.

Weigend, Andreas S., B.A. Huberman, and D.E. Rumelhart, 1992, "Predicting sunspots and exchange rates with connectionist networks", in *Nonlinear Modeling and Forecasting,* edited by M. Casdagli and S. Eubank, Addison-Wesley Publishing Company, Reading, MA, pp. 395-432.

Weigend, Andreas S. and N.A. Gershenfeld (editors), 1994, *Time Series Prediction: Forecasting the Future and Understanding thePast*, Proceedings of the NATO Advanced Research Workshop on Comparative Time Series Analysis, Santa Fe, May 14--17, 1992, Addison-Wesley Publishing Company, Reading, MA.

White, H., 1989, "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings", *Neural Networks,* 3 (5), 535-549.

Appendix A: 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[14] 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 non-metric 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

14 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.

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 used 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[15]. 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 were 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 consisted 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 over-fitting, 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 "over-learn".

Appendix B: Creating Feature Vectors

---

15 Elman (1990) developed 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 non-linear interaction among series and is well-suited for time-series analysis where the objective is to identify spatial and temporal patterns.

In our FEVA based prediction model, data points from x(k-80) to x(k) were used to predict x(k+1). Figure 8 shows the details of the feature vector calculations. x(k-40) is the *focus point* and we look forward and backward *with respect to the focus point.* Different types of looks can be formed by changing the focus point.

Let z = x(k-40) be the *focus point.* Define

$$
\begin{aligned}
\text{Ord}(x) &= 1 \quad \text{if } x > 0 \\
&= 0 \quad \text{if } x = 0 \\
&= -1 \quad \text{if } x < 0
\end{aligned}
$$

as an up/down indicator function. The details of the feature vector components used in our model are shown below:

Feature Vector Components (fv)

fv(1) = z;

Backward Looks: Long

fv(2) = x(k-80);

fv(3) = (z - x(k-80)) / x(k-80);

fv(4) = Ord(z - x(k-80));

fv(5) = x(k-60);

fv(6) = (z - x(k-60)) / x(k-60);

fv(7) = Ord(z - x(k-60));

fv(8) = x(k-50);

fv(9) = (z - x(k-50)) / x(k-50);

fv(10) = Ord(z - x(k-50));

Backward Looks: Short

fv(11) = x(k-45);

fv(12) = (z - x(k-45)) / x(k-45);

fv(13) = Ord(z - x(k-45));

fv(14) = x(k-44);

fv(15) = (z - x(k-44)) / x(k-44);

fv(16) = Ord(z - x(k-44));

fv(17) = x(k-43);

fv(18) = (z - x(k-43)) / x(k-43);

fv(19) = Ord(z - x(k-43));

fv(20) = x(k-42);

fv(21) = (z - x(k-42)) / x(k-42);

fv(22) = Ord(z - x(k-42));

fv(23) = x(k-41);

fv(24) = (z - x(k-41)) / x(k-41);

fv(25) = Ord(z - x(k-41));

Forward Looks: Long

fv(26) = x(k);

fv(27) = (z - x(k)) / x(k);

fv(28) = Ord(z - x(k));

fv(29) = x(k-20);

fv(30) = (z - x(k-20)) / x(k-20);

fv(31) = Ord(z - x(k-20));

fv(32) = x(k-30);

fv(33) = (z - x(k-30)) / x(k-30);

fv(34) = Ord(z - x(k-30));

Forward Looks: Short

fv(35) = x(k-35);

fv(36) = (z - x(k-35)) / x(k-35);

fv(37) = Ord(z - x(k-35));

fv(38) = x(k-36);

fv(39) = (z - x(k-36)) / x(k-36);

fv(40) = Ord(z - x(k-36));

fv(41) = x(k-37);

fv(42) = (z - x(k-37)) / x(k-37);

fv(43) = Ord(z - x(k-37));

fv(44) = x(k-38);

fv(45) = (z - x(k-38)) / x(k-38);

fv(46) = Ord(z - x(k-38));

fv(47) = x(k-39);

fv(48) = (z - x(k-39)) / x(k-39);

fv(49) = Ord(z - x(k-39));

Table 1: 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 exceeded the riskless rate of 5% per annum. We report the t-test for the non-paramateric 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[16].

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 Non-Parametric Test

# when m < rf	N1	92
# when m> rf	N2	92
Number of Obs.	N	184
# right when m<rf	n1	36
# wrong when m>rf	n2	18
# of bear calls	n	54
Exp. of dist	E(n1)	27
Std of distribution		2.53
T-test for HM Test		3.56

16 See McCullagh, P. and J.A. Nelder, 1983, p.98 for details.

Table 2: 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.

		Randomizing Returns: Bootstrap Results					
	Actual Values	mean	median	std	t-test	.95 percentile	rank
Hamilton Beta	0.326	0.305	0.311	0.091	0.060	0.446	0.501
Hamilton Annual Return	10.73%	5.14%	4.98%	1.98%	2.435	8.38%	0.992
Hamilton Std.	10.44%	10.18%	10.14%	0.93%	-2.088	8.89%	0.007
Hamilton Sharpe Ratio	.559	0.510	0.497	0.207	3.371	0.856	1.000
Hamilton Jensen Measure	4.04%	-1.55%	-1.68%	1.97%	2.364	1.79%	0.990
S&P Annual Return	10.75%	10.80%	10.86%	2.64%	0.036	15.06%	0.519
S&P Std.	12.83%	12.77%	12.76%	1.01%	-1.511	11.45%	0.027
S&P Sharpe Ratio	0.456	0.460	0.453	0.214	0.303	0.812	0.634

		Randomizing Strategies: Bootstrap Results					
	Actual Values	mean	median	std	t-test	.95 percentile	rank
Hamilton Beta	0.311	.306	.306	.099	.051	.467	.509
Hamilton Annual Return	9.95%	4.97%	4.93%	1.80%	276.60%	8.00%	1.00
Hamilton Std.	8.24%	9.04%	9.03%	0.36%	-222.00%	8.40%	0.03
Hamilton Sharpe Ratio	1.208	.551	.547	.204	3.225	.86	1.00
HamiltonJensen Measure	3.12%	-1.76%	-1.75%	1.97%	247.70%	1.48%	.990

Table 3: 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 4: 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 2nd" 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 STD
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's	1.477	11.10	2.43
1940's	3.213	6.04	5.66
1950's	9.641	9.91	5.27
1960's	7.712	9.68	6.53
1970's	0.409	6.74	4.30
1980's	12.626	11.29	7.01
1990's	15.442	16.24	10.72

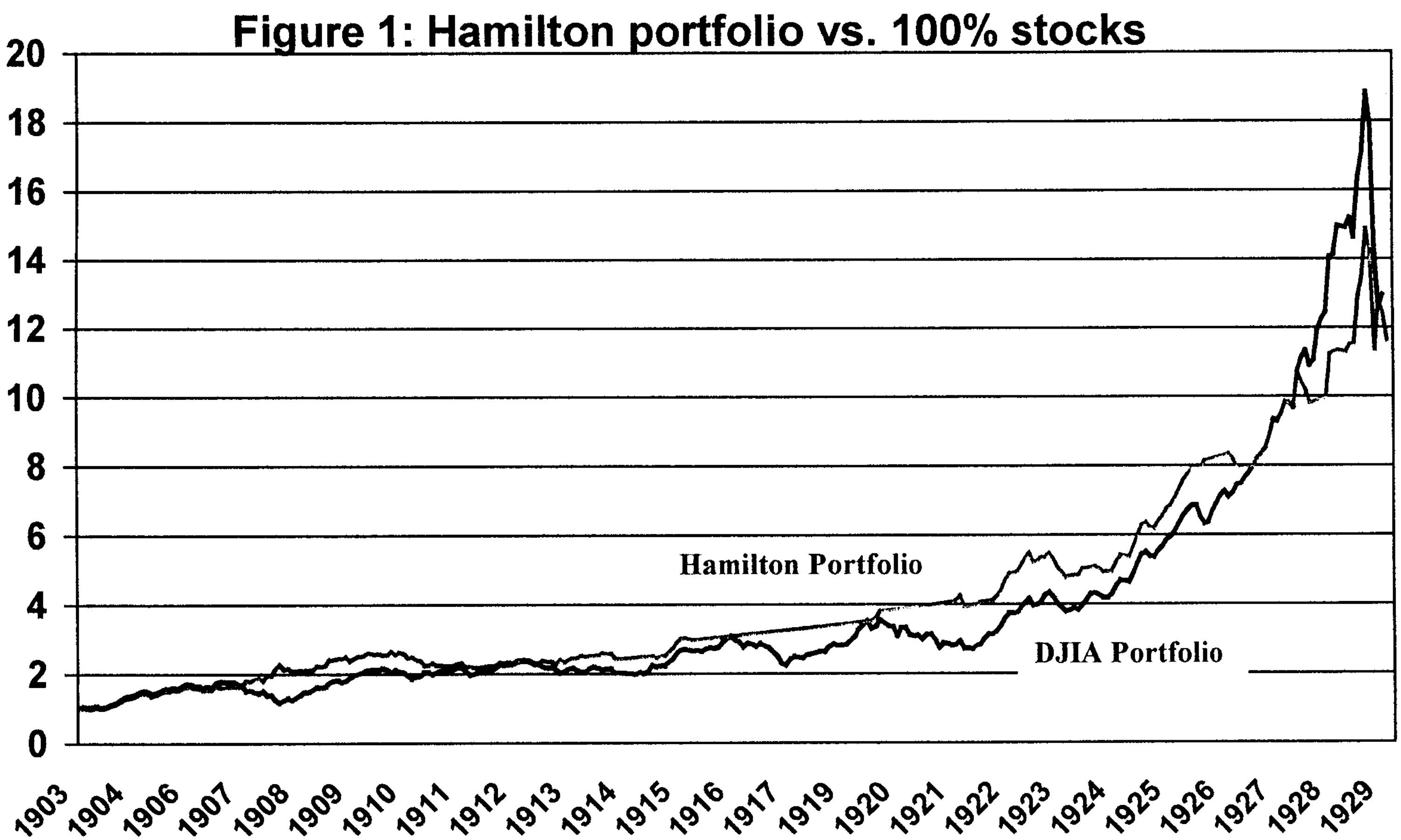

Figure 1: *Relative performance of the Hamilton portfolio compared to a portfolio invested entirely in the market over the 27 years. The figure indicates that the Hamilton portfolio was less volatile than the fully invested strategy.*

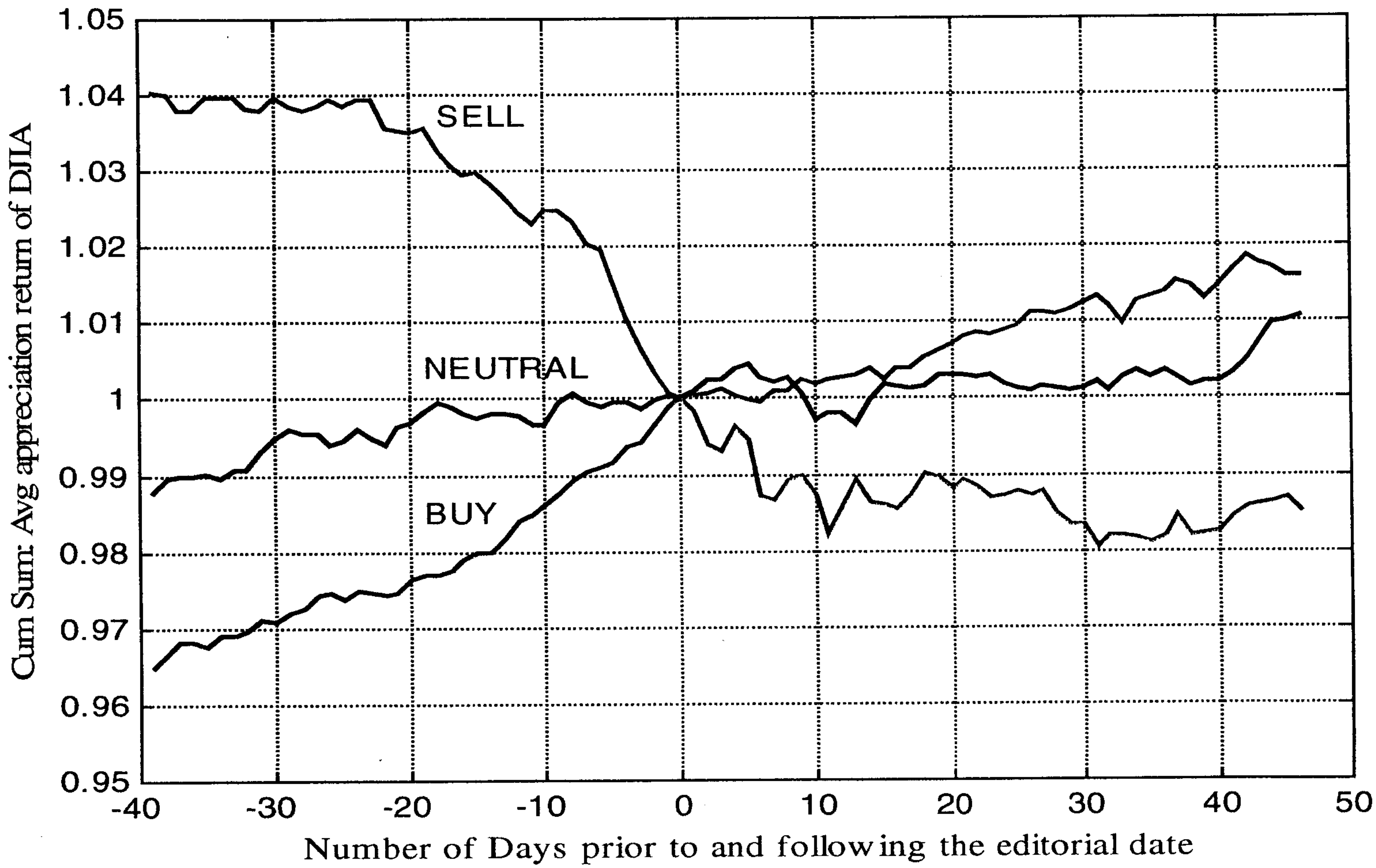

Figure 2: *Paths representing the cumulated sum of the equal-weighted average appreciation return of the Dow Industrial Index over a window of eighty-one trading days: forty days before publication date and forty days following publication of the editorial.*

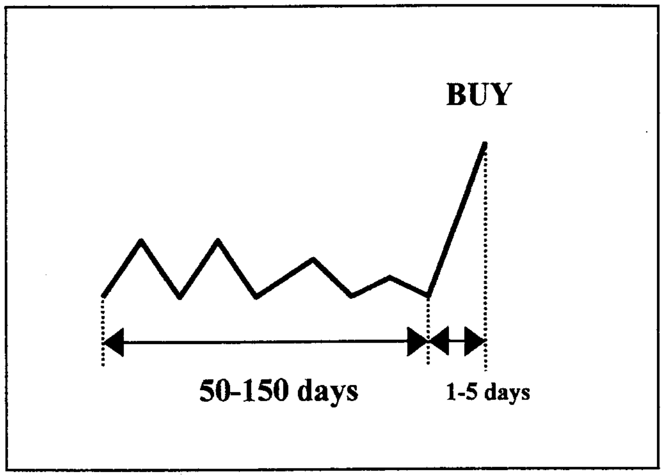

Figure 3: *An example (schematic) of the type of pattern a simple moving average (MA) trading rule would detect.*

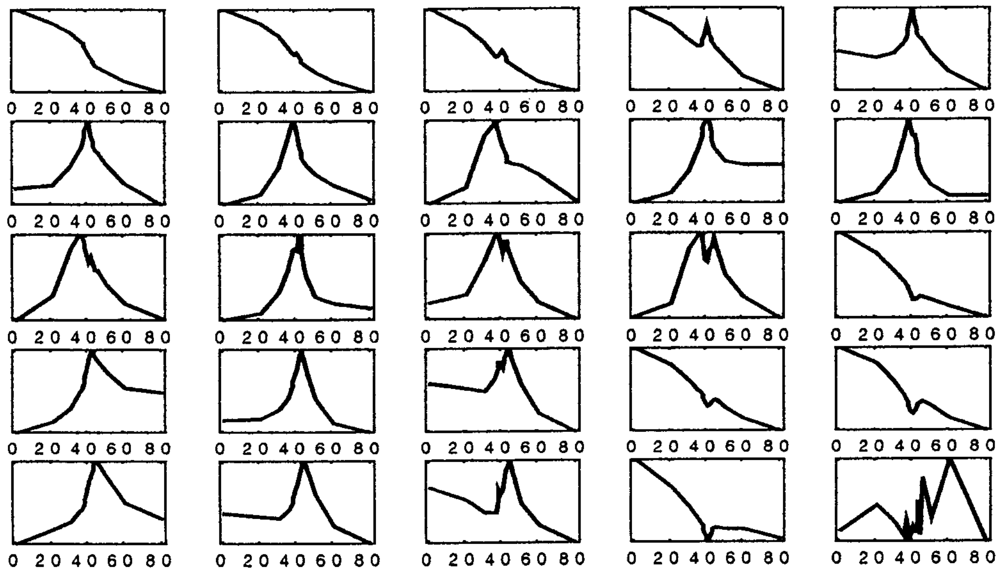

Figure 4: *25 Prominent SELL Indicators. FEVA Plots representing the "structure" of the DJIA time series. We obtained a 100 cluster solution for the enhanced dataset (27774 x 49 matrix) and selected the prominent FEVA (25 BUY indicators and 25 SELL indicators) from the solution. Patterns represent the price path of the DJIA over the eighty days preceding the forecast. The horizontal axis represents the number of days used to construct the feature vector.*

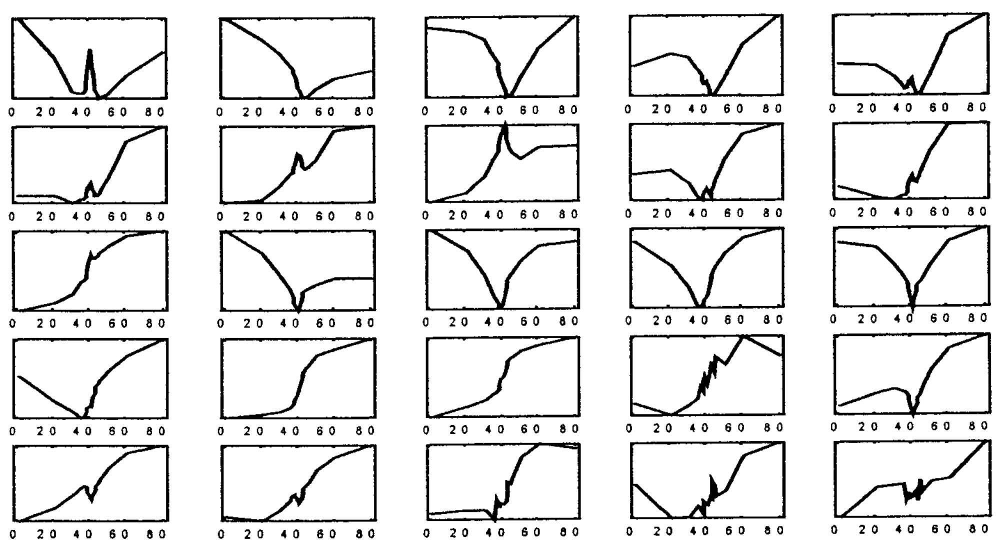

Figure 5: *FEVA Plots: 25 Prominent BUY indicators.*

Figure 6: *Neural Net In-Sample Predictions: the predicted calls are superimposed on the DJIA series where the color represent the call: green is a "buy" call, red is "sell" and blue represents a "neutral" call. It is quite clear that the NN model is able to learn the patterns in the Dow series.*

Figure 7: Out of Sample Forecast Portfolios vs. DJIA Index

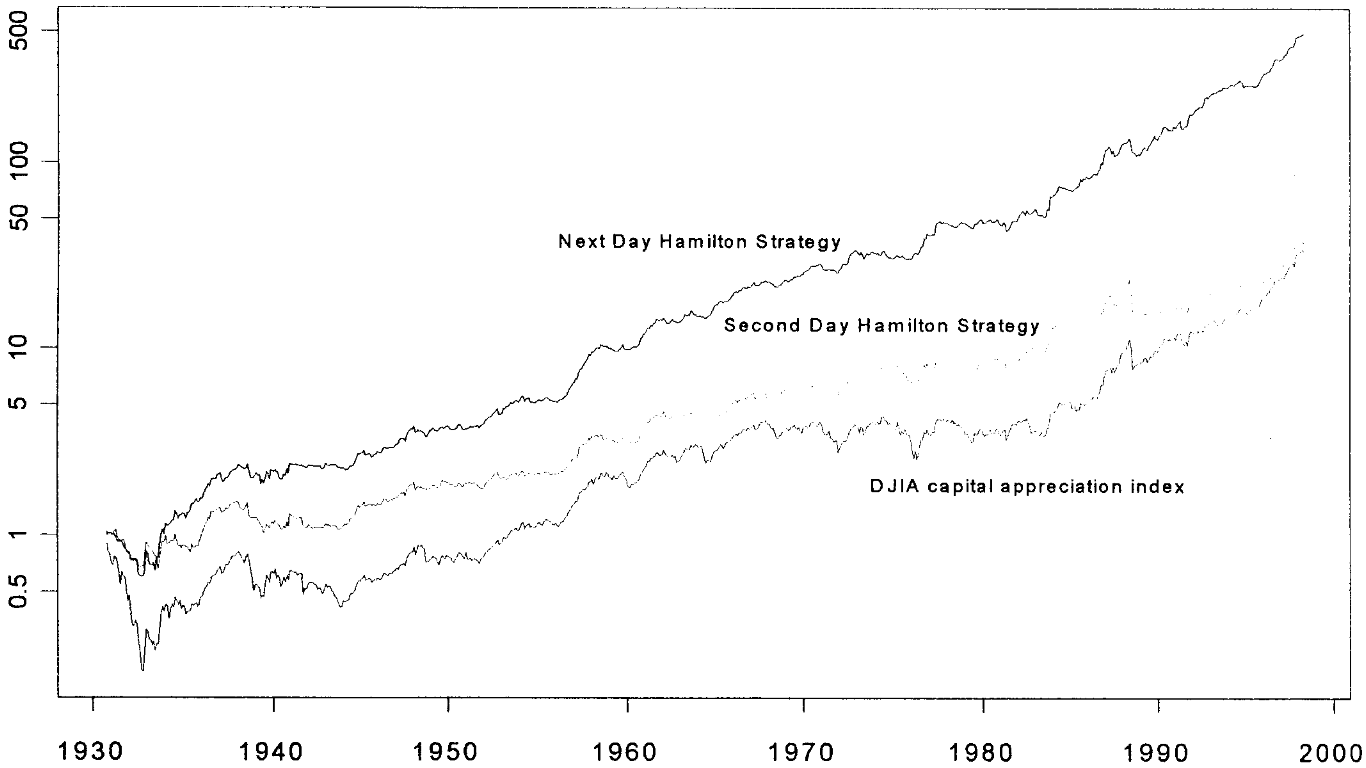

Figure 7: *Results from timing strategies based upon the Hamilton model. Two strategies are considered: (1) the investor can act immediately on the call (indicated as "Next Day Hamilton Strategy") and (2) the investor can only trade at the close of the day that the forecast comes out (indicated as "Second Day Hamilton Strategy"). The "Next Day Hamilton Strategy" appears to be a dramatic improvement over the buy and hold. It has a higher return with lower risk. The "Second Day Hamilton Strategy" is less dominant. It has a return over the entire history almost exactly equal to the buy and hold. The major benefit is that the strategy has less variance and less systematic risk.*

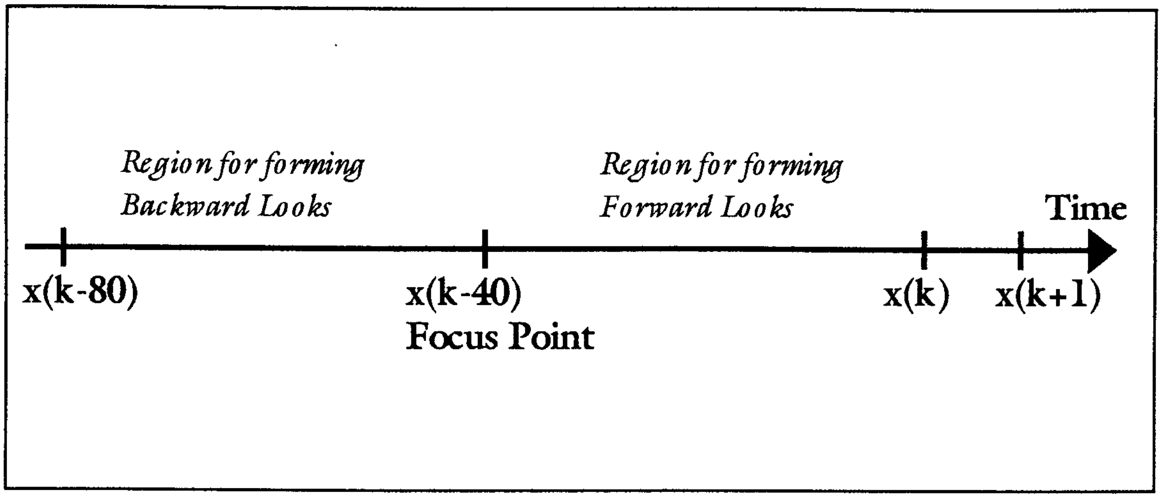

Figure 8: *A schematic diagram describing the feature vector (FEVA) calculations: Focus point, backward looks and forward looks (with respect to the focus point). Features obtained using data points between x(k-80) and x(k) are used to predict x(k+1).*