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

Very Preliminary! Comments Welcome. Please do not Quote.

William N. Goetzmann
Yale School of Management

March 25, 1997

Abstract: Alfred Cowles' (1934) test of the Dow Theory apparently provided strong evidence against the ability of 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 timing strategies based upon the Dow Theory yield high Sharpe ratios and positive alphas.

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

Aknowledgements: I thank my colleagues Geert Rouwenhorst and N. Prabhala for helpful comments. All errors are the sole responsibility of the author.

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 in fact 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. The difference in the results is 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.

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 and section VI concludes.

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:

> **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.[1]

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. 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" is a useful guide. First, market movements may be decomposed into primary, secondary and tertiary trends, the most important of which are 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 seconds reflects selling due to decreased business and earnings, and the third is caused by distress selling of sound securities, regardless of their value.”[2]

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. 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, which 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.

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 *Baron's* as "bullish", "bearish" or "neutral." Cowles then assumes that on a bullish signal, an investo 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 was taken as a recommendation to invest in t-bills. 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 facia* 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 .55-.16 = .39, it seems that the Dow strategy earned a risk-adjusted return of .12 - [.05+.39(.155-.05)] = .029. In other words, Cowles' interpretation of Hamilton's strategy would seems 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. Never-the-less, a close look at the Cowles evidence suggests that the Dow Theory, as practiced by William Peter Hamilton merits re-consideration.

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 indeterminant. We then collect total return information on the U.S. stock market information 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.

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 indeterminant cases out of the 255 which we eliminate 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.

Non-Parametric Tests

To address the basic question of Hamilton's timing skill, we examine how often the Dow beats or lags 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. A natural test of the Dow Theory would be a Henriksson-Merton test, however that test is only appropriate for a bivariate investment choice.[3] As an alternative, we perform a related non-parametric test, Fisher's exact test, which indicates that the positive association between Hamilton's calls and subsequent results. Fisher's test is statistically significant at the 1% level. 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, an explicit trading test in necessary.

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. 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 short-term rate of 5% used by Cowles in his analysis. 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 first recommendation that appeared in a month, and then assume that this is used as a guide to rebalancing at the end of the month. In those months for which we have more than one recommendation, this means that we ignored the later call. As a consequence, we do not pick up intra-month returns to the Dow strategy.

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 1907, when he avoids the worst of the panic of that year. He also does well in 1917 and 1920, when the Dow portfolio is out of the market during both bear runs. In general, the figure indicates that the Dow 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 Dow portfolio is 9.95% (9.83% geometric), slightly below the annual average return obtained by holding the Cowles all-stock portfolio, which yields an annual arithmetic average of 10.90% (10.54% geometric). On a risk-adjusted basis, however, the Dow 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 Cowles index.

Bootstrapping Tests

The rest of Table 2 reports the results of significance tests generated by bootstrapping the Dow 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 generating stock return series' by drawing monthly returns with replacement from the Cowles 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 Dow 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.[4] This appears to provide evidence against the random walk assumption of the bootstrap. The Sharpe measure of the Dow portfolio exceeds all of the bootstrapped values, and the Jensen measure of the Dow 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 Dow 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

is in fact the procedure pioneered by Cowles himself in another part of his landmark 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 o f actual analyst performance records to those generated under a null that forecasts were simpl y 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 editoria l series. We thus generate 500 simulated track records under a null that the editor was, in effect, flipping coin, properly weighted so as to give the same expected proportions of "Bull" an d "Bear" forecasts as in the original series. The advanta ge of this is that we do not break the actual time-series characteristics pf the market history itself. Our bootstrap in the space of strategies now conditions upon the true market realization. We do, however, alter the time-serie s characteristics of Hamilton's 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.

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. 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 the forty trading days preceding and following the 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 (.034 prob.value). The neutral calls have a .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.

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 the Dow Theory is correct, or whether it simply means that Hamilton was one lucky forecaster among

many market analysts is another 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.

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 was a key element 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

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

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.

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

Rhea, Robert, 1932, *The Dow Theory*, Barron's, New York.

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. 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.[5]

	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

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
Dow Beta	0.311	0.305	0.311	0.091	0.060	0.446	0.501
Dow Annual Return	9.95%	5.14%	4.98%	1.98%	2.435	8.38%	0.992
Dow Std.	8.24%	10.18%	10.14%	0.93%	-2.088	8.89%	0.007
Dow Sharpe Ratio	1.208	0.510	0.497	0.207	3.371	0.856	1.000
Dow Jensen Measure	3.12%	-1.55%	-1.68%	1.97%	2.364	1.79%	0.990
Cowles Annual Return	10.90%	10.80%	10.86%	2.64%	0.036	15.06%	0.519
Cowles Std.	11.24%	12.77%	12.76%	1.01%	-1.511	11.45%	0.027
Cowles Sharpe Ratio	0.525	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
Dow Beta	0.311	.306	.306	.099	.051	.467	.509
Dow Annual Return	9.95%	4.97%	4.93%	1.80%	276.60%	8.00%	1.00
Dow Std.	8.24%	9.04%	9.03%	0.36%	-222.00%	8.40%	0.03
Dow Sharpe Ratio	1.208	.551	.547	.204	3.225	.86	1.00
Dow Jensen Measure	3.12%	-1.76%	-1.75%	1.97%	247.70%	1.48%	.990

Figure 1: Dow Theory vs. 100% Stocks

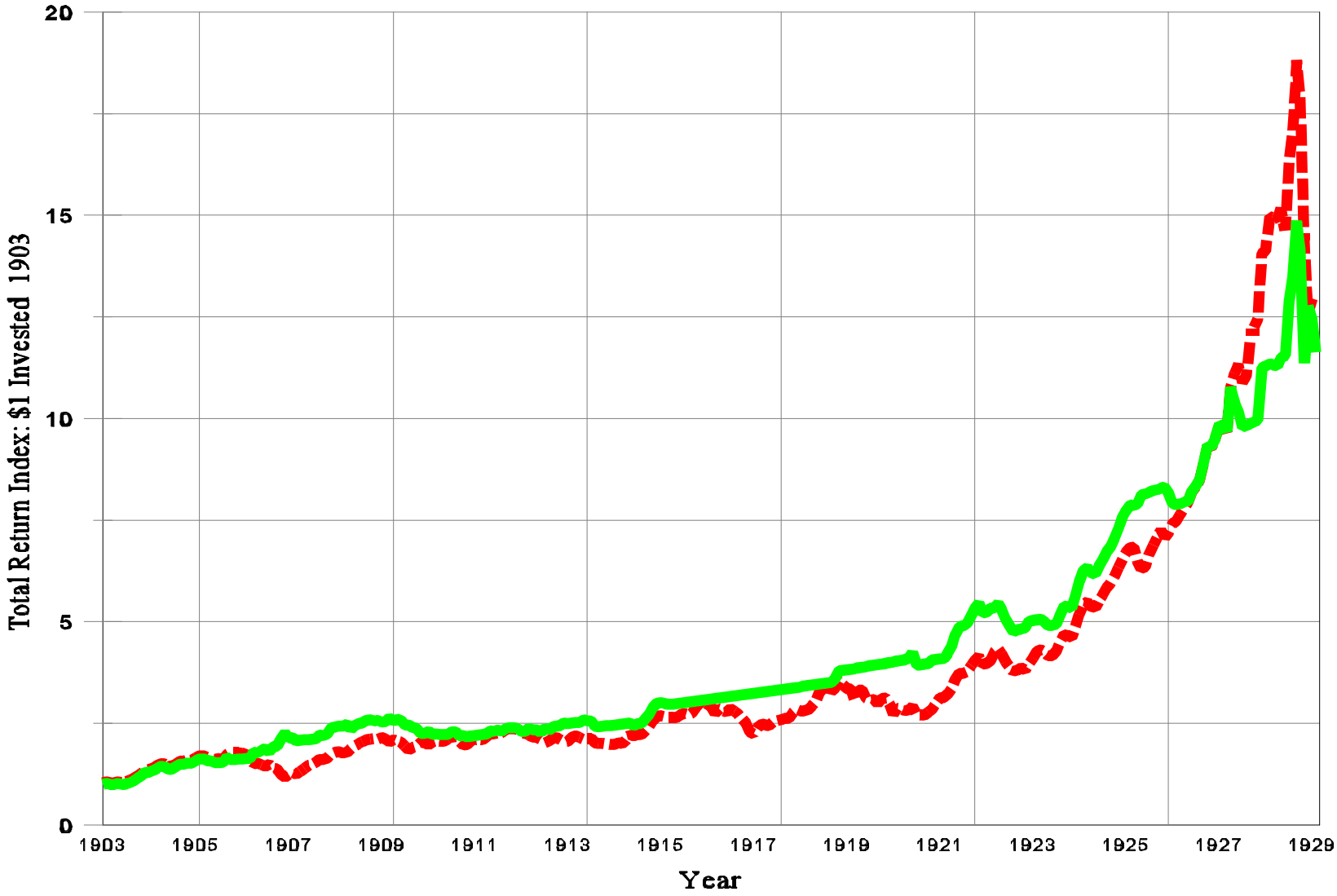
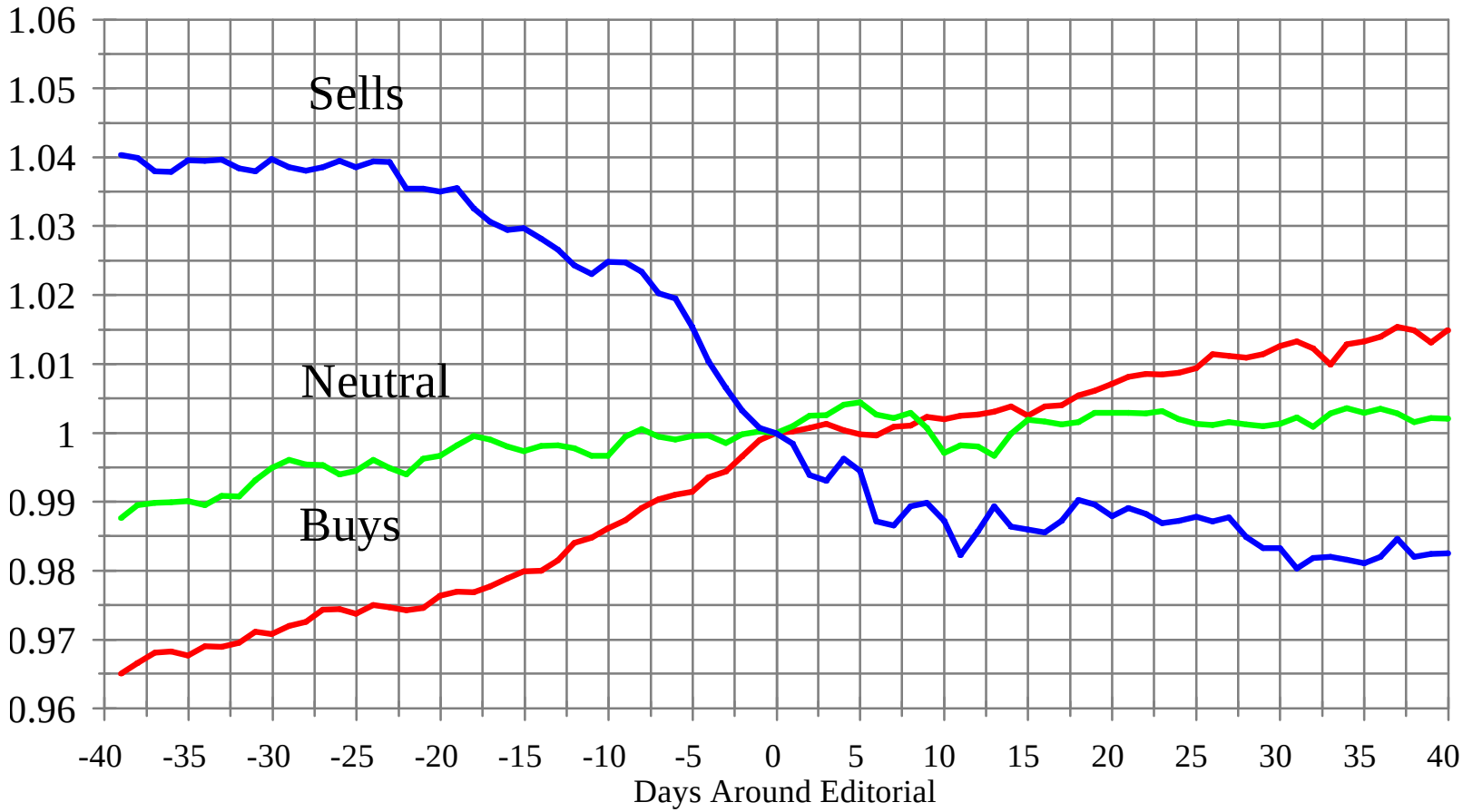


Figure 2. DJIA Around Editorials



Notes

1. Rhea (1932) p. 12

2.Ibid. P.13.

3. The Henriksson-Merton test assumes that the investor has the choice only to invest or not invest in the market. The timer is thus credited only with success at calling when the market underperforms a benchmark. Thus, remaining in the market is effectively a passive decision, while getting out of it is an active one. As a consequence, the H-M is a function only of the ratio of successful to unsuccessful "down" calls. The Cowles strategy attributes an active portfolio choice to both positive and negative market forecasts.

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

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