

Window Dressing, Data Mining, Or Data Errors: A Re-Examination Of The Dogs Of The Dow Theory

Larry J. Prather, (Email: prather@etsu.edu), East Tennessee State University
Genell L. Webb, (Email: gennytn@juno.com), Tennessee State University

Abstract

We re-examine the Dow dividend yield anomaly to ascertain if data errors create the superior returns of the trading rule. Empirical testing, using both parametric and nonparametric methods, suggests that the trading rule outperforms the index. Additionally, data errors are not the drivers of superior trading rule returns. Moreover, the Chow breakpoint test of structural stability suggests that neither window dressing nor data mining explain this phenomenon. Finally, a turn of the year formation date fails to explain superior trading rule returns, further mitigating the data mining explanation.

1. Introduction

Recently, two contradictory articles were published on the efficacy of the Dow dividend yield theory, or the “Dogs of the Dow.” O’Higgins and Downes (1991) introduced the “Dogs of the Dow” trading rule in their best-selling publication *Beating the Dow*. The suggested trading rule is straightforward and only requires five steps. First, look up the “last” price and dividend yield for each of the thirty stocks that make up the DJIA. Second, rank the stocks in order of their dividend yields. Third, select the ten stocks with the highest dividend yields. Fourth, purchase an equal dollar amount of each of the selected stocks. Finally, wait one year and repeat the process.

Through introducing a method of stock selection and disposal that yields returns superior to the index, O’Higgins and Downes have implied that the market is not efficient. After studying this method, investors and researchers alike are left pondering its validity. Two such individuals are Prather (2000) and Hirschey (2000). However, the empirical evidence reported by these scholars has further complicated the matter since they draw different conclusions.

Prather empirically analyzes the theory as initially proposed by O’Higgins and Downes. As a result, he concludes that the theory is valid and that it leads investors to superior risk-adjusted returns. He also links this “anomaly” with other empirical anomalies such as the P/E ratio effect (Basu, 1977 & 1983), small firm effect (Banz, 1981), dividend yield effect (e.g., Fama and French, 1988; Hodrick, 1992), and evidence on market overreaction (e.g., De Bondt and Thaler, 1985; Lehmann, 1990; Jegadeesh, 1990; Lo and MacKinlay, 1990).

Hirschey focuses on why early tests of the theory work and is concerned that this method could be a result of data mining. He is also disturbed that several “data errors” exist in previous empirical tests. In response to his concerns, he extends the period of investigation that O’Higgins and Downes established to perform his own empirical testing. As a result, he reports no abnormal risk-adjusted returns. Hirschey interprets his findings as support that (1) the theory prevails only because it was applied to a concise time-period selected _____
Readers with comments or questions are encouraged to contact the authors via email.
through data mining, and (2) data errors in other tests drove the superior results. Therefore, he discounts the theory’s validity.

Because these two tests of this theory have led to such contradictory results, a third test is required to reconcile the conflicting results and extend the analysis. The focus of this paper will be to: (1) resolve Hirschey's argument that "index data errors" drive the results; (2) compare trading rule returns reported by Hirschey, returns reported by Prather, and returns computed in this study to ensure that trading rule return error does not drive the results; (3) ascertain if the anomaly is an artifact of a turn-of-the-year effect; and (4) examine the pre-1970 period to determine if data mining or window dressing may have caused this anomaly in the post-1970 period.

O'Higgins and Downes claim that a transition in the market took place around 1970 and that the market transitioned from being dominated by wealthy individual investors to institutional investors. Further, they claim that "window dressing" causes the institutional "dumping" of stocks at prices below their intrinsic value. This is plausible since it is consistent with the literature on window dressing by Falkenstein (1996), Lakonishok, Shleifer, Thaler, and Vishny (1991), and Musto (1999). An alternate explanation provided by Hirschey is that the theory is an outgrowth of data mining. Therefore, an examination of the pre-1970 and post-1970 periods and the existence of a structural break in the regression coefficient vector could possibly resolve the issue of data mining or window dressing.

2. Data Sources

To implement our test, we attempted to replicate the strategy proposed by O'Higgins and Downes as closely as possible. Clearly, data errors are a potential problem in this type of research due to the complexity of computing returns and forming portfolios. These complexities are discussed below. Investors are encouraged to use the *Wall Street Journal* (WSJ) to find information. Therefore, we consulted the first issue of the WSJ for each calendar year to ascertain the closing price of each of the 30 Dow stocks for the prior year and their dividend yields. Before 1977, the WSJ reported annual dividends and price. This required investors to compute yields. During 1977, the WSJ began publishing yields (to one decimal place), which simplified the process. This is one potential (minor) source of conflict in the construction of trading rule portfolios and may explain some differences in reported trading rule portfolio returns. The reason for this is that the procedure of O'Higgins and Downes requires calculation of dividend yields and selection of the highest yield stocks. In case of a tie, they recommend selecting the lower priced stock. In several instances, our calculated yields were very close and the portfolio would differ slightly depending on the number of decimal places used in calculating dividend yields. Therefore, during years when researchers had to compute dividend yields instead of using the published yields, the composition of their portfolios could differ slightly (by one stock) due to differences in the precision used in computations. Once the portfolio of the ten highest yielding stocks is formed, it will be held until the last trading day of the year and then liquidated.

Once portfolios are formed, returns for each stock must be computed. Many issues are involved in computing portfolio returns. These various issues arise due to stock splits, special dividends, and rights distributions. We identified firm specific events by consulting *Capital History & Capital Changes* provided by the Hilliard Lyons research department. When dealing with stock splits, one must consider the effects of both the acquisition of additional shares and additional dividends. For example, General Motors experienced a 2 for 1 common stock split in 1989. In a case such as this, the closing price used in computing the return must be multiplied by the number of stocks a holder of one share GM would have after the split (2 shares). Additionally, dividends for the acquired shares must be accounted for. In 1989, GM paid quarterly dividends. Because the record date of the split was 2-17-89, a holder of one share GM stock before the split would now receive the dividends of 2 shares GM stock for each quarter.

Numerous types of special dividends must also be accounted for when computing portfolio returns. Many companies comprising our portfolios distributed stock of another company to its stockholders. When this occurred, we assumed that the shares acquired were held until the end of the calendar year and then liquidated at their market price. The price received from the sale of the stock, and any dividends paid to shareholders of that stock are included in the portfolio return. For example, for every share of AT&T stock held in 1984, a shareholder received 0.10 shares of each of seven "baby bells." In addition, all four quarterly dividends paid by each "baby bell" were received. Several of our companies also issued special cash dividends. In such a case, the cash is held until the end of the calendar year. When a company issued a common stock dividend, we adjusted the ending price (as was done

with a split) to reflect the additional stock. A failure to account for these distributions would negatively bias trading rule portfolio returns and bias against finding superior performance.

A final complication involves the distribution of rights. Accurate computation of returns requires valuing these rights. The dilemma is establishing a legitimate value. Holding a right is similar to holding an American call option. Since the value of an American call option consists of both intrinsic value and time value, an option holder is better off selling the option than exercising the option early. Additionally, in a world with transaction costs, an investor would choose to let the option expire if the transaction costs of selling the option exceeded its value. A right holder would do so as well. We have chosen to ignore rights distributions in our calculations except in the cases where firms redeemed rights by distributing cash to the rights holders in exchange for the rights. Our procedure causes the computed trading rule returns to have a very small negative bias. However, our procedure provides a floor on actual returns.

2.1. Computation of trading rule portfolio returns

Portfolios used in this study are formed and disposed of on the first trading day of the year and the annual returns on the trading rule portfolios are computed using equation (1)

$$R_{pt} = \sum_{i=1}^N \frac{P_{it} - P_{it-1} + D_{it} + I_t + S_{it}}{P_{it-1}} \tag{1}$$

where R_{pt} is the return on the trading rule portfolio during period t , P_{it} is the closing price of security i at time t , P_{it-1} is the closing price of security i at time $t-1$, D_{it} is the dividend paid on security i during the twelve-month holding period, I_t is the interest earned from reinvesting dividend distributions (to provide consistent results for investors who do not reinvest, we do not assume dividend reinvestment and set $I_t = 0$), and S_{it} is the value at time t of any stock received as a distribution.

2.2. Data errors

Hirschey argues that some studies contain data errors that bias results in favor of the trading rule portfolios. Therefore, it is essential to examine the returns of both the DJIA and trading rule portfolios to reconcile the work of Prather and Hirschey. Table 1 presents the DJIA and trading rule returns used by Hirschey, Prather, and those used in this study.

An examination of the DJIA returns from 1961-1998 suggests that the returns for two years are quite different. In 1962, Hirschey reports returns of 7.64% compared with -7.62 for Prather. Other sources also report negative DJIA returns for 1962 (The Motley Fool reports -7.37%, Chase Investment Performance Digest reports -7.43%). This acts to bias Hirschey’s results against finding superior trading rule returns since DJIA returns are overstated. In 1996, Hirschey reports DJIA returns of 28.04 whereas Prather reports DJIA returns of 22.6%. Other studies also report returns above 24% (Barron’s reports 28.6%, Merrill Lynch reports 28.57%, and The Motley Fool reports 24.33%) suggesting that Prather’s DJIA returns are understated. This creates a bias in favor of the trading rule portfolios in Prather’s study.

Table 1 DJIA and Trading Rule Returns

The returns for the DJIA and trading rule portfolios are presented below. Column one is the year and columns two through four are the DJIA returns used by Hirschey (Hirdow), Prather (Pradow), and those of this study (Newdow). Columns five through seven are the trading rule returns used by Hirschey (Hirdogs), Prather (Pradogs), and those of this study (Newdogs).

<u>Year</u>	<u>Hirdow</u>	<u>Pradow</u>	<u>Newdow</u>	<u>Hirdogs</u>	<u>Pradogs</u>	<u>Newdogs</u>
1961	21.82	22.40	22.60	26.06	25.03	25.03

1962	7.24	-7.62	-7.43	-2.48	8.21	-0.59
1963	20.07	20.59	20.83	19.03	18.92	20.90
1964	18.14	18.67	18.85	19.23	23.08	21.00
1965	13.83	14.16	14.39	16.64	20.17	20.30
1966	-14.88	-15.65	-15.78	-14.22	-18.02	-16.66
1967	18.43	19.04	19.16	24.22	26.58	26.32
1968	7.59	7.73	7.93	13.78	14.18	11.49
1969	-10.95	-11.60	-11.78	-15.92	-14.14	-14.10
1970	8.58	8.76	9.21	0.57	7.60	7.60
1971	9.58	9.79	9.83	4.88	4.21	4.01
1972	17.74	18.22	18.48	22.70	23.28	23.30
1973	-12.43	-13.14	-13.28	0.32	3.96	4.00
1974	-21.45	-23.14	-23.58	-2.95	-0.70	-0.61
1975	42.71	44.40	44.75	47.28	56.99	56.45
1976	21.98	22.72	22.82	32.97	33.61	33.74
1977	-11.76	-12.70	-12.84	0.97	-2.42	-2.43
1978	2.88	2.69	2.79	1.15	-1.13	-0.54
1979	10.27	10.52	10.55	6.40	8.24	13.25
1980	20.57	21.41	22.17	28.41	27.23	27.98
1981	-2.81	-3.40	-3.57	2.21	3.73	3.74
1982	24.77	25.79	27.11	17.66	29.25	27.49
1983	24.74	25.68	25.94	37.97	38.53	38.81
1984	1.26	1.05	1.31	4.85	6.26	6.35
1985	31.67	32.78	33.55	27.72	29.21	30.22
1986	26.12	26.92	27.10	24.73	27.32	32.45
1987	5.93	6.02	5.48	7.45	1.46	6.13
1988	15.52	15.95	16.14	17.71	21.18	24.03
1989	30.70	31.71	32.19	27.62	26.53	25.98
1990	-0.40	-0.58	-0.56	-12.95	-7.56	-7.48
1991	23.32	23.93	24.19	34.34	36.25	33.85
1992	7.22	7.35	7.41	2.94	11.52	11.24
1993	16.37	15.10	16.94	22.80	27.31	24.29
1994	4.89	5.02	5.02	0.73	10.21	4.11
1995	35.75	35.60	36.94	35.17	37.44	36.83
1996	28.04	22.60	28.91	27.25	24.05	32.90
1997	24.36	22.76	24.91	19.80	21.56	20.72
1998	17.75	17.28	18.04	10.97	12.08	11.70

Despite these differences, the average returns presented by these two studies do not differ dramatically. Table 2 shows that the average DJIA returns for Hirschey are 12.77% while Prather’s average DJIA returns are 12.34%. A paired sample t-test fails to reject the null hypothesis of equal returns. One plausible explanation for the minor differences in returns, for years other than 1962 and 1996, is that the authors did not use identical sample formation periods for computing returns (i.e., end of December vs. beginning of January). While it is unlikely that these minor differences drive the results in either study, it provides an impetus for examining another set of DJIA returns. For this study, DJIA returns and the three-month T-bill rates were obtained from Thomson Financial Company.

Table 2 Comparison of Average DJIA Returns 1961-1998 of Prather and Hirschey

Columns one through three list the study, average DJIA return, and standard deviation of return respectively. Columns four and five present the t-statistic and p-value for a paired sample t-test for a difference in means.

<u>Study</u>	<u>Mean Return</u>	<u>Standard Deviation</u>	<u>T-statistic</u>	<u>P-value</u>
Hirschey (2000)	12.77%	14.77	0.99	0.33
Prather (2000)	12.34%	15.51		

Trading rule portfolio returns are another source of data errors. These are subject to greater differences and greater error since return computation is much more complex. An examination of the trading rule returns reported by Hirschey and Prather suggest that numerous discrepancies exist. One possible explanation is that the formation dates differed slightly. Prather uses *Wall Street Journal* information for the last trading day of the year while Hirschey uses the first trading day of the year. However, since the first issue of the *Wall Street Journal* each year reports year-end data, it is not possible to determine if the two studies use the same formation day or whether they are a day apart. While this may have little impact on index returns, it could have a greater impact on the returns of any given portfolio. The reason for the greater impact on a portfolio is that with a constant dividend, daily changes in price will cause daily changes in yields. Since the trading rule is based on yields, even small daily changes could alter portfolio composition (by one or two stocks) and therefore the portfolio return.

Two other sources of differences between the portfolios in the two studies exist. The first is the precision with which yield is computed (as discussed in section 2 above). The second difference is that Prather uses the historical yield whereas Hirschey uses the expected yield for the next year. Given the number of potential factors that could drive trading rule return differences, completely reconciling the previous research is not possible. However, it is possible to compare the trading rule returns from the previous studies to those from this study to ascertain whether the trading rule returns are significantly different from one another.

Table 3 reports paired sample t-tests of the hypothesis that the average returns of the trading rule portfolios are equal. Results suggest that significant differences exist between Hirschey’s returns and the returns reported by both Prather and those in this study. However, there is no statistical difference between Prather’s returns and those reported here.

Table 3 Paired Sample T-tests of Trading Rule Portfolio Returns

Columns one through nine present the trading rule return test pairs, mean difference between returns, standard deviation, standard error of the mean, lower and upper 95% confidence intervals, the t-statistic for testing the null hypothesis that returns are equal (two-sided test), the number of degrees of freedom, and the p-value. The descriptors in column one are as follows: HIRDOGS represents trading rule returns reported by Hirschey, PRADOGS are Prather’s trading rule returns, and NEWDOGS are the trading rule returns in this study.

Test Pairs	Mean Difference	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t-statistic	df	Sig. (2-tailed)
				Lower	Upper			
HIRDOGS – PRADOGS	-2.1895	4.0327	.6542	-3.5150	-.8640	-3.347	37	.002
HIRDOGS – NEWDOGS	-2.2576	3.3954	.5508	-3.3737	-1.1416	-4.099	37	.000
PRADOGS – NEWDOGS	-.0682	2.9110	.4722	-1.0205	.8887	-.144	37	.886

3. Methodology

Since it is well documented that risk and return bear a nearly linear relationship, one explanation for superior trading rule returns is that the selected portfolios have higher risk than the DJIA. Therefore, risk-adjusted returns must be examined. The Jensen (1968) measure of risk-adjusted abnormal returns were computed using the market model in equation (2)

$$R_{p_t} - R_f = \alpha + \beta_p (R_{mt} - R_f) + \varepsilon_{p_t} \tag{2}$$

where R_p is the annual return on the portfolio under investigation, R_f is the risk-free rate (three-month T-bill), α is the abnormal risk-adjusted return, β_p is the estimated systematic risk of the portfolio, R_{mt} is the annual return on the

DJIA (used as a proxy for the market portfolio), and ϵ_p is the error term.

4. Empirical Results

Table 4 presents results of market model regressions. Results in Panel A suggest that with this sample, the “Dogs of the Dow” outperformed the DJIA by more than 4% annually on a risk-adjusted basis. The point estimate of the systematic risk is also lower than the index. However, a t-test suggests that the difference in risk is not statistically significant.

As a test of robustness, a portfolio formation date of July 1st was also examined. This permits a test of the impact of both the formation date and the January effect on trading rule profits. Results in Panel B suggest that the “Dogs of the Dow” outperformed the DJIA by more than 3% annually on a risk-adjusted basis. This finding is consistent with the results of De Bondt and Thaler (1985) who report that loser portfolios exhibit higher returns and lower risk than winner portfolios in subsequent periods and that this result is not related to a January effect.

Panels C and D present results of using the trading rule returns reported by Hirschey and Prather, respectively. In both cases, our index returns and risk-free rate are used to standardize results. In all four cases, the trading rule exhibited statistically superior risk-adjusted performance.

Another approach to examining the significance of return differences involves using nonparametric procedures. Using nonparametric procedures is helpful since it can mitigate criticism that results are driven by a violation of the assumptions of the CAPM. Therefore, to examine the robustness of the results above, we use a Sign test and Wilcoxon Signed Ranks Test.

The Sign test is a nonparametric test that can be used with two related samples. This test makes no assumptions about the shape of the distributions but tests the hypothesis that the distributions of the variables are the same using only the direction of the differences between pairs.

The Wilcoxon Signed Ranks Test can also be used to test whether two related samples are the same. It differs from the Sign test in that it uses information about the magnitude of differences in addition to their direction.

The results of the nonparametric tests are presented in Table 5. Table 5 shows that the trading rule dominates the index, despite the choice of index returns. The trading rule returns were greater than the index returns in at least 26 of the 38 years. Based on the Wilcoxon Matched-Pair Sign Rank Test Z-statistic and p-value, presented in Panel B, the null hypothesis of no difference in returns between the trading rule and the index can be rejected at the 1% level. Panel C shows that the Sign test confirms the above results at the 5% level. This evidence suggests that the trading rule consistently outperforms the index and that the superior performance is not simply due to a few large outliers.

Table 4 Dow Dog Risk-Adjusted Returns Using Thomson's Index Returns

Columns one through six list the coefficients, their estimates, standard errors, t-statistic, significance level (p-value), and the r-squared of the model.

Panel A January 1st Formation Date

	<i>Coefficients</i>	<i>Std. Error</i>	<i>t</i>	<i>Sig.</i>	<i>R²</i>
∇	4.188	1.153	3.361	.001	.840
∩	.911	.066	13.754	.000	

Panel B July 1st Formation Date

	<i>Coefficients</i>	<i>Std. Error</i>	<i>t</i>	<i>Sig.</i>	<i>R²</i>
∇	3.055	1.371	2.228	.032	.866
∩	.767	.074	10.388	.000	

Panel C Hirschey’s Trading Rule Returns

	<i>Coefficients</i>	<i>Std. Error</i>	<i>t</i>	<i>Sig.</i>	<i>R²</i>
∇	2.223	1.211	1.836	.075	.812
∃	.866	.070	12.452	.000	

Panel D Prather’s Trading Rule Returns

	<i>Coefficients</i>	<i>Std. Error</i>	<i>t</i>	<i>Sig.</i>	<i>R²</i>
∇	4.268	1.258	3.392	.002	.808
∃	.889	.072	12.291	.000	

4.1. Robustness of results during pre-1970 and post 1970 periods

To ascertain whether results are confined to the post-1970 period, we segmented the sample into the pre-1970 period (1961-69) and post-1970 period (1970-98) and repeated empirical tests on each sub sample. During the 1961-69 period, the market model provides an abnormal risk-adjusted return of 2.50% (P=0.067) suggesting that the trading rule worked before 1970. The Wilcoxon Matched Pair Sign Rank Test shows that the trading rule produced a larger return than the DJIA in seven of the nine years, which was sufficient to reject the hypothesis of no difference in returns (P=0.051).

During the 1970-98 period, the market model produced an abnormal risk-adjusted return of 4.91% (p=0.002). The Wilcoxon Matched Pair Sign Rank Test shows that the trading rule exhibited higher returns than the DJIA in 19 of 29 years and that the hypothesis of equal returns could be rejected (p=0.016).

The combined results of the pre-1970 and post-1970 periods suggest that the trading rule worked before 1970 and continued to work. This weakens Hirschey’s claim that data mining is the cause of the superior returns that O’Higgins and Downes reported during the 1973-1989 period. However, it also casts doubt on O’Higgins and Downes’ “window dressing” explanation for why the rule works.

4.2. Window dressing or data mining

Two explanations for the original results reported by O’Higgins and Downes have been proposed. Hirschey (2000) believes that the result is due to data mining. However, O’Higgins and Downes (1991) claim that the anomaly is a result of window dressing by portfolio managers. Therefore, O’Higgins and Downes

Table 5 Wilcoxon Signed Ranks Tests

Column one in Panel A lists the test-pairs. DJIARTN is the return of the DJIA as reported by Thomson Financial Company and NEWDOGS are the trading rule portfolio returns computed in this study. HIRDOW and PRADOW are the index returns reported by Hirschey (2000) and Prather (2000). Column three presents the number of negative differences, positive differences, and ties between the index and trading rule returns. The mean rank and sum of ranks are presented in columns four and five respectively. Panel B presents the Z-value and p-value of the Wilcoxon Matched-Pair Sign Rank test for testing the null hypothesis that the trading rule returns and index returns are equal and Panel C presents the Z-value and p-value of the Sign test for testing the null hypothesis that the trading rule returns and index returns are equal.

Panel A Ranks

		<i>N</i>	<i>Mean Rank</i>	<i>Sum of Ranks</i>
DJIARTN - NEWDOGS	Negative Ranks	26	22.00	572.00
	Positive Ranks	12	14.08	169.00
	Ties	0		
	Total	38		
HIRDOW - NEWDOGS	Negative Ranks	26	21.92	570.00
	Positive Ranks	12	14.25	171.00

	Ties	0		
	Total	38		
PRADOW - NEWDOGS	Negative Ranks	27	22.15	598.00
	Positive Ranks	11	13.00	143.00
	Ties	0		
	Total	38		

Panel B Wilcoxon Matched Pair Test Statistics

	DJIARTN - NEWDOGS	HIRDOW - NEWDOGS	PRADOW - NEWDOGS
Z	-2.922	-2.893	-3.299
Asymp. Sig. (2-tailed)	.003	.004	.001

Panel C Sign Test Statistics

	DJIARTN - NEWDOGS	HIRDOW - NEWDOGS	PRADOW - NEWDOGS
Z	-2.109	-2.109	-2.433
Asymp. Sig. (2-tailed)	.035	.035	.015

hypothesize that the increased institutional ownership of stock created the anomaly around 1970. This may be due to the stronger focus toward tax-efficient investing. Lakonishok, Shleifer, Thaler, and Vishny (1991) believe that if window dressing occurs, it is in response to the end of year evaluation of fund managers by the plan sponsor. If this evaluation goes beyond a strict quantitative performance evaluation to include the current holdings of the funds, managers may sell losers to improve the appearance of holdings. Musto (1997) also believes that this is the case. Musto’s findings suggest that managers oversell extreme losers, particularly in the fourth quarter. This could be the result of the funds trying to be tax-efficient by offsetting gains with losses. Falkenstein (1996) also finds that funds hold proportionately more stocks that have recently appreciated. These findings are consistent with the proposition made by O’Higgins and Downes.

Table 6 examines the perpetual growth of institutional ownership. Clearly, as O’Higgins and Downes (1991) claim, institutional investors are becoming more important. Further, the average percentage of institutional ownership of Dow stocks was 50 percent in 1986 compared with an average of 43 percent for the 1,000 largest firms ranked by market capitalization (*Business Week*, April 18, 1986). However, the impact of institutional ownership on the proposed trading rule is unknown and requires investigation. This issue is addressed in the following section.

Table 6 Percentage of Institutional Ownership of NYSE-Listed Stocks

Row one lists the year. Row two lists the percentage of institutional ownership for all NYSE-listed stocks.

1950	1960	1970	1980	1990	1999
6.1	18.7	26.7	35.4	44.1	50.2

4.3. Structural breaks and stability of coefficient vectors

A common problem in econometrics is that structural shifts may occur during the sample period. Therefore, fitting a regression equation to the full sample can result in a poor fit of the regression equation and biased estimates of the regression coefficients. An example of a structural shift would be the change in the impact that institutional investors have on the stock market if O’Higgins and Downes’ assertion is correct.

To examine data for structural shifts, the Chow breakpoint-test can be used. The Chow test requires data to be partitioned into two or more periods (before and after the structural shift). The regression equation is fitted separately to each period and the coefficients are then tested for stability. If coefficient vectors are unstable (a structural shift), the null hypothesis of stability can be rejected. This is important in examining either the “data mining” theory or the “window dressing” theory. If the original results of O’Higgins and Downes were due to data mining, as Hirschey suggests, we would expect a structural shift. Furthermore, this shift should occur before the period of their investigation. Additionally, if O’Higgins and Downes are correct in their theory that institutional investors caused this anomaly, a structural break should occur. If, however, there is no structural break, it may be that the trading

rule is related to other well-documented anomalies.


To ascertain if a structural break occurred in the data, a Chow Breakpoint test was conducted on the coefficient vectors of the regression equation. Since O'Higgins and Downes (1991) claim that structural changes took place around 1970 that created this phenomenon, we tested the stability of the regression coefficients for the pre-1970 and post-1970 periods. The Chow test was unable to reject the hypothesis of stable coefficients ($P=0.354$) suggesting that neither data mining nor window dressing is the sole cause of this anomaly.

5. Conclusion

We examine the possibility of outperforming the market by employing a simple trading strategy based on dividend yields. Through this examination, it becomes evident that market anomalies exist which permit the achievement of abnormal returns. Empirical results suggest that the trading rule outperforms the index by more than 4%. In addition, while any rigorous empirical test is prone to human error, the results reported in this paper suggest that data errors do not drive the superior trading rule results.

Furthermore, we cannot conclude that the success of this trading strategy is attributable to window dressing by portfolio managers or data mining. Given this evidence, it is possible that the trading rule is related to the many well-documented CAPM anomalies that remain unexplained. These include the P/E ratio (Basu, 1977 & 1983), the small firm effect (Banz, 1981), the evidence on dividend yields (e.g., Fama and French, 1988; Hodrick, 1992), and evidence on market overreaction (e.g., De Bondt and Thaler, 1985; Lehmann, 1990; Jegadeesh, 1990; Lo and MacKinlay, 1990).

6. Suggestions for Future Research

The practical importance of implementing simple trading rules that can outperform an index is significant. However, there is no magic in the DJIA. If this anomaly is related to other documented anomalies, it is likely that similar trading rules could be used to outperform other indices. Therefore, examining other U.S. indices, sector indices, and foreign indices would provide an interesting extension. However, this rule may only work on indices where all, or nearly all, of the stocks pay dividends. 

References

1. Banz, Rolf W., "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, Vol. 9, No.1, pp. 3-18, 1981.
2. Basu, Sanjoy, "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, Vol. 32, No. 3, pp. 663-682, 1977.
3. Basu, Sanjoy, "The Relationship Between Earnings Yield, Market Value, and Return for NYSE Common Stocks: Further Evidence," *Journal of Financial Economics*, Vol. 12, No. 1, pp. 129-156, 1983.
4. De Bondt, Werner F. M., and Richard Thaler, "Does the Stock Market Overreact?" *Journal of Finance*, Vol. 40, No. 3, pp. 793-808, 1985.
5. Falkenstein, Eric G., "Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings," *Journal of Finance*, Vol. 51, No. 1, pp. 111-136, 1996.
6. Fama, Eugene F., and Kenneth R. French. "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, Vol. 22, No. 1, pp. 3-25, 1988.
7. Hirschey, Mark, "The Dogs of the Dow Myth," *Financial Review*, Vol. 35, No. 2, pp. 1-15, 2000.
8. Hodrick, Robert J., "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference Measurement," *Review of Financial Studies*, Vol. 5, No. 3, pp. 357-386, 1992.
9. Jegadeesh, Narasimhan, "Evidence of Predictable Behavior of Security Returns," *Journal of Finance*, Vol. 45, No. 3, pp. 881-898, 1990.
10. Jensen, Michael C., "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance*, Vol. 23, No. 2, pp. 389-416, 1968.
11. Lakonishock, Josef, Andrei Shleifer, Richard Thaler, and Robert Vishny, "Window Dressing by Pension Fund Managers," *American Economic Review*, Vol. 81, No. 2, pp. 227-231, 1991.
12. Lehman, Bruce N., "Fads, martingales, and market efficiency," *Quarterly Journal of Economics*, Vol. 105,

- No. 1, pp. 1-28, 1990.
13. Lo, Andrew W, and Craig A. MacKinlay, "When Are Contrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies*, Vol. 3, No. 2, pp. 175-205, 1990.
 14. Musto, David K., "Investment Decisions Depend on Portfolio Disclosures," *Journal of Finance*, Vol. 54, No. 3, pp. 935-952, 1999.
 15. O'Higgins, Michael, and James Downes, *Beating the Dow*, Harper Collins, New York, New York, 1991.
 16. Prather, Larry J., "An Empirical Test of the Dow Dividend Theory," *The Southern Business and Economic Journal*, Vol. 23, No. 3, pp. 170-184, 2000.