How Accurate is the January Barometer?

Ben R. Marshall*
Department of Economics and Finance
Massey University
Private Bag 11-222
Palmerston North
New Zealand
Tel: +64 6 350 5799 Ext. 5402
Fax: +64 6 350 5651
B.Marshall@massey.ac.nz

Nuttawat Visaltanachoti
Department of Commerce
Massey University
Private Bag 102-904
Auckland
New Zealand
Tel: +64 9 414 0800 Ext. 9460
Fax: +64 9 441 8177
N.Visaltanachoti@massey.ac.nz

Abstract

The ability of positive (negative) equity market returns in January to predict positive (negative) returns in the following 11 months of the year is the subject of much discussion in the financial media. We find the January Barometer does not work in any of the 22 international equity markets we consider. At first glance, the January Barometer appears to work in US equity indices; however closer examination reveals this performance is spurious. We cannot rule out time-varying risk premia as an explanation and our US individual stock analysis is inconsistent with the US index results.

JEL Classification: G10, G11, G12, G14
Keywords: January Barometer, Seasonality, Return Predictability, Quantitative Investment

First Version: 29 September 2007
This Version: 14 March 2008

* Corresponding Author. We wish to thank Henk Berkman, Ben Jacobsen, Jeff Wongchoti, and Fei Wu for valuable comments.

How Accurate is the January Barometer?

Abstract

The ability of positive (negative) equity market returns in January to predict positive (negative) returns in the following 11 months of the year is the subject of much discussion in the financial media. We find the January Barometer does not work in any of the 22 international equity markets we consider. At first glance, the January Barometer appears to work in US equity indices; however closer examination reveals this performance is spurious. We cannot rule out time-varying risk premia as an explanation and our US individual stock analysis is inconsistent with the US index results.

JEL Classification: G10, G11, G12, G14
Keywords: January Barometer, Seasonality, Return Predictability, Quantitative Investment

First Version: 29 September 2007
This Version: 14 March 2008

1. Introduction

We conduct a comprehensive examination of the ability of returns in the month of January to predict returns in the remaining 11 months of the year in the US and 22 other equity markets. The worth of the "January Barometer" has been discussed by practitioners for decades[1], but it is only recently that academic researchers have subjected this technique to rigorous analysis. Cooper, McConnell, and Ovtchinnikov (2006) consider the ability of the January Barometer to predict returns for the remaining 11 months of the year in the US for the 1940 – 2003 period and find it has substantial power. Brown and Luo (2006) consider the performance of the January Barometer in the US over a very similar period (1941 – 2003) and show that it is particularly powerful at predicting yearly declines.

Whenever a new anomaly is documented it is important to determine its robustness by considering the profits on offer relative to transaction costs, risk, and its out-of-sample performance using the most appropriate econometric techniques. It is also essential to consider whether there are any institutional or theoretical explanations (either based on investor rationality or behavioral factors) for the anomaly. Investigating each of these factors in the context of the January Barometer, as we do in this paper, is particularly relevant as Cooper, McConnell, and Ovtchinnikov (2006) note that they are unaware of any theoretical justification for its success, so a null hypothesis of no predictive power might be a reasonable prior despite their statistical analysis in the US rejecting this.

Jensen (1978) points out that a market is still efficient if the gross profits on offer do not offset the costs incurred in exploiting an anomaly. For this reason, the January Barometer is an attractive anomaly to investigate. It only requires one round trip transaction per year so the profits generated only need to be compared to one estimate of transactions costs. In contrast, the net profits, and therefore implications for market efficiency, of anomalies which require frequent trading can be very sensitive to estimates of transactions costs (e.g. Lesmond, Schill, and Zhou, 2004).

Both Lakonishok and Smidt (1988) and Lo and MacKinlay (1990a) suggest that the best way to gain an insight into anomaly robustness is to use new data. One approach to the

[1] Cooper, McConnell, and Ovtchinnikov (2006), provide evidence of one of the first mentions of the January Barometer on p. 319. This is "We doubt that any technique or indicator ever devised has been so remarkably accurate as the January Barometer. The barometer, which indicates that as January goes, so will the market go for the total year, has proven correct in 20 of the last 24 years. The performance of this indicator becomes even more striking when you consider its simplicity, coupled with the fact that it is making its prediction eleven months in advance" (Hirsch, 1974, p. 11).

new data issue is to wait and observe whether the January Barometer persists over the next few decades in the US as Cooper, McConnell, and Ovtchinnikov (2006) suggest. We take another option and consider the January Barometer in international markets. This has the obvious advantage of enabling us to form conclusions about robustness now. It is also useful in that as Goeztmann and Jorion (1999) point out, the US is atypical in that it has the most successful capitalist system in the world so January Barometer results in the US may not necessarily hold elsewhere. A third reason to study international markets stems from their growing importance to international portfolio managers. A recent survey found the average allocation of money to international markets by global funds was 57 percent in 2006 compared with just 37 percent in 2002.[2]

Verifying the robustness of anomalies by checking their performance in international markets is not new. The fact that the momentum effect, as documented by Jegadeesh and Titman (1993) in the US equity market, performs well in international markets (e.g. Rouwenhorst, 1998) has gone a long way to establishing it as credible evidence against market efficiency. Fama (1998, p. 304) describes the momentum effect as an "open puzzle."

We use the stochastic dominance techniques of Anderson (1996), Davidson and Duclous (2000) and Barrett and Donald (2003), and the manipulation-proof performance measure of Ingersoll, Spiegel, Goetzmann, and Welch (2007) to check whether the performance of the January Barometer is driven by risk-based explanations. The stochastic dominance approach provides a framework which is not dependent on asset pricing benchmarks and requires only minimal assumptions about investors' utility functions and asset return distributions. The performance measure developed by Ingersoll, Spiegel, Goetzmann, and Welch (2007) does not rely on the assumptions of normal or lognormal distributed returns or independent and identically distributed variables so it is not subject to the same biases as popular performance measurement techniques such as the Sharpe ratio or Jensen's alpha.[3]

[2] http://www.iht.com/articles/2007/04/25/bloomberg/bxfund.php.

[3] Our use of these two techniques is driven by both practical and theoretical considerations. Firstly, the data required for commonly used equilibrium models of null returns, such as the Fama and French (1993) three factor model, are not available for all the markets we consider. Secondly, we suggest that our risk-adjustment techniques are theoretically superior. Fama (1998) cautions against rejecting market efficiency with tests based on a specific model of equilibrium returns as abnormal returns could simply indicate the equilibrium model is mis-specified. Furthermore, there is growing evidence that popular empirical asset pricing models are not robust. Ferson (2006) highlights that the commonly used Fama and French (1993) factors cannot distinguish legitimate risk factors from those that are unrelated to systematic risk.

Powell, Shi, Smith and Whaley (2007) show that standard OLS regression techniques can generate spurious results in time-series models involving indicator or dummy variables if the dummy variable is persistent. We find evidence of persistence in our dummy variable, which equals one if the return in January is positive and zero otherwise. For instance, in the US between 1983 and 2006 there were only two years with negative January returns compared to 22 years with positive January returns. We therefore adopt the Powell, Shi, Smith and Whaley (2007) methodology. The issue of data mining bias also needs to be carefully considered in empirical asset pricing tests. Out of sample tests are an important aspect of this, but this is not always possible so researchers have developed statistical techniques to formally account for data mining bias. We apply the technique proposed by Cooper, McConnell, and Ovtchinnikov (2006) and the method suggested by Powell, Shi, Smith and Whaley (2007).

Our paper contributes to the literature in two ways. First, we show that the January Barometer is not reliable in any international market. Of the 23 non-US series (22 countries and MSCI World Index) we consider 19 have, on average, larger 11-month returns following positive Januaries then negative Januaries. However, the differences in 11-month returns following positive and negative Januaries are only statistically significant in two countries – Spain and Switzerland, and the World Index. The January Barometer results in both these countries stand up to adjustment for risk using stochastic dominance techniques and the manipulation-proof performance measure. However, the World Index result does not. The results in all three markets are not statistically significant following joint adjustment for dummy variable persistence and data mining bias.[4]

Secondly, and more importantly, our study casts doubt on reliability of the US January Barometer findings. At first glance, the January Barometer appears to work in the US. It remains profitable after risk adjustment based on stochastic dominance and manipulation-proof performance techniques. In addition, it cannot be fully explained by dummy variable persistence or data mining bias. However, closer examination reveals the US index result is not robust. We are unable to rule out time-varying risk premia as an explanation. However, it

[4] Prior to our analysis, the evidence regarding the January Barometer in international markets was mixed. Hensel and Ziemba (1995b) find positive January returns predict positive 11-month returns in Australia, Canada, Japan, and the U.K. Easton and Pinder (2007) find that while the January Barometer does not generally work on international markets it does work in Italy, Norway, Thailand, and Zimbabwe. Bohl and Salm (2008) also show that the January Barometer does not hold in all markets they consider, but does work in Norway and the Netherlands. We discuss these papers in more detail in Section 2.

is our US individual stock analysis that casts the most doubt on accuracy of the January Barometer, in our opinion.

We are unaware of any institutional reasons why the January Barometer should work in the US but not in other markets so we investigate the US market further by focusing on individual stocks. Proponents of the January Barometer highlight its success on the market in general rather than individual stocks (see Section 2 for more detail), but we suggest that if it is in fact a strong predicator in the US then we would expect to see those stocks that have positive January returns out-performing those stocks that have negative January returns for the remaining 11 months of the year. Interestingly, we find no evidence of this. As a final step, we investigate whether the January Barometer holds in US traded commodities. If there is a genuine behavioural reason for its existence it should be expected to work in these series. To the contrary we find no evidence of predictability for the January Barometer in this market. Overall, we suggest that our results provide support for the conclusion that the apparent profitability of the January Barometer in the US equity indices is simply due to chance.

The remainder of this paper is organized as follows. Section 2 provides evidence of the current focus on the January Barometer in the popular press, a discussion of the academic papers that have investigated the January Barometer, and possible explanations for its success. Our methodology and results are presented in Section 3. Section 4 contains robustness checks, which Section 5 concludes the paper.

2. Background on the January Barometer

In this section we provide evidence of the current focus on the January Barometer in the popular press, discuss the academic literature, and consider possible explanations for the existence of the January Barometer.

2.1. Practitioner Focus

According to Cooper, McConnell, and Ovtchinnikov (2006), the January Barometer appears to have been first mentioned by Yale Hirsch in his Stock Traders' Almanac publication in the 1970s. An example of Hirsch's endorsement of this indicator from his 1974 publication, which they provide on p. 319, is included below:

“We doubt that any technique or indicator ever devised has been so remarkably accurate as the January Barometer. The barometer, which indicates that as January goes, so will the market go for the total year, has proven correct in 20 of the last 24 years. The performance of this indicator becomes even more striking when you consider its simplicity, coupled with the fact that it is making its prediction eleven months in advance (Hirsch, 1974, p. 11).

The January Barometer continues to attract widespread media attention. Selections of the headlines of stories devoted to it are included below:

“The January Barometers Score Again; How the S&P 500 goes for the first month can give a good reading of the year.” (Sam Stovall, *Business Week Online*, 2 July 2003)

“January Barometer predicts a pretty lousy year ; Month is turning out to be a loser, and chances are 2005 will be, too” (Adam Shell, *USA Today*, 31 January 2005).

“Molasses in January; A weak start -- like this year's -- can signal a bear market. Time to tread lightly?” (Marcia Vickers, *Business Week*, 31 January 2005).

“The January Barometer really works: Foretells market's future better than any other month” (John Dorfman, *Bloomberg News*, 1 February 2006).

“January barometer' points to prosperity” (Tom Walker, *The Atlanta Journal – Constitution*, 1 February 2006).

“January Barometer'bodes well for 2007; Rosy start seen as reliably good sign for investors” (Adam Shell, *USA Today*, 1 February 2007).

2.2. Academic Papers

The conclusions reached by the first academic papers regarding the validity of the January Barometer are mixed. Hensel and Ziemba (1995a) find support using the S&P 500 index over the 1926 – 1993 period, while Fuller (1978) finds, using S&P 500 and DJIA data, that the January Barometer produces returns no better than a buy-and-hold strategy over the 1898 – 1977 period. However, neither of these papers consider the risk-adjusted performance of the January Barometer. This is important as any one trading based on the signals emitted by the January Barometer would spend long periods of the time out of the market (i.e. all 11 month periods following a negative January return) and would therefore likely incur less risk than if they adopted a buy-and-hold strategy.

Two recent papers consider the profitability of the January Barometer after risk is accounted for. Cooper, McConnell, and Ovtchinnikov (2006) consider the ability of the January Barometer to predict returns for the remaining 11 months of the year in the US and find it has substantial power. They focus on the 1940 – 2003 period but they consider the robustness of their result using data dating back to 1825. Cooper, McConnell, and Ovtchinnikov (2006) show the January Barometer works in small and large stocks, value and growth stocks and persists after adjustment for macroeconomic / business cycle variables, investor sentiment, and the presidential cycle in stock returns. Brown and Luo (2006) consider the performance of the January Barometer in the US over a very similar period (1941 – 2003) using Sharpe Ratios to account for risk and show that it is particularly powerful at predicting yearly declines.

To the best of our knowledge, there are three papers that consider the performance of the January Barometer in international markets. Hensel and Ziemba (1995b) find positive January returns predict positive 11-month returns in Australia, Canada, Japan, and the U.K. Easton and Pinder (2007) find that while the January Barometer is usually unreliable but it does work in Italy, Norway, Thailand, and Zimbabwe. Bohl and Salm (2008) also show that the January Barometer does not hold in all markets they consider, but does work in Norway and the Netherlands.

As mentioned in the introduction, we suggest that our paper contributes to the literature by comprehensively showing that the January Barometer does not work in the US or any of the 22 other countries we consider. Our conclusions are at variance with the previously mentioned literature due to them not considering one or all of the following: biases introduced by dummy variable persistence and data snooping, time varying risk-premia explanations, and possible institutional and behavioral explanations.

2.3. Explanations

The January Barometer is inconsistent with the tenets of modern finance theory. It runs counter to the concept of weak form market efficiency (e.g. Fama, 1970) which suggests that past price information is not useful when it comes to predicting future price movements. Any anomaly which lacks a theoretical explanation is particularly exposed to the criticism that it is simply a statistical illusion so many researchers have turned their attention to potential explanations for anomalies. For instance, Hong and Stein (1999) develop the gradual information diffusion hypothesis where investors react slowly to information as an explanation for the momentum effect.

To the best of our knowledge there are no conclusive explanations for the January Barometer. However, we briefly discuss practitioner claims regarding the factors behind the January Barometer. Little and Albrecht (2006, p.3) state:

"The major marginal players in world equity markets are the major institutions. Powerful investment committees run these institutions. Calendar years for most start in January and the first investment results appear before these powerful committees in early February. These committees cannot afford to ignore what seems to be working. They launch funds in what seem to be "hot" areas and allocate assets to likely winners, pushing up prices in those sectors. So – the argument goes – if you see what's been working at the end of January, you get an "inside feel" for what might work for the rest of the year."

Such behavior by investment committees appears consistent with Representativeness Bias, which was introduced by Kahneman and Tversky (1974). In other words, investors try and determine if it will be a positive year for the equity market via the use of the representativeness heuristic of January performance.

While the Little and Albrecht (2006) theory may go some way to explaining the January Barometer there are many questions left unanswered. For instance, why don't rational market participants arbitrage away its profits? Barberis and Thaler (2003) highlight that many examples that are inconsistent with market efficiency can be explained by limits to arbitrage, such as short sale constraints, which prevent investors from exploiting the situation. A large portion of the profits to the January Barometer occur from entering the market with a long position following a positive January return so there does not appear to be any restriction on investors exploiting this. As the month of January progresses and the likelihood of a positive return in January increases one would expect rational market participants to start entering the market and in the process impound any impact of January Barometer into price before the end of January which would leave no reaction the positive January return for the rest of the year. Given the lack of theory supporting the January Barometer we devote much of Section 4 to the examination of whether institutional factors or behavioral factors are behind the apparent success of the January Barometer in the US.

3. Test Procedures and Empirical Results

In this section we discuss our data, present our methodology and results and discuss their implications. We source CRSP equally weighted (EW) and value weighted (VW) indices for the 1940 – 2006 period. Our start point of 1940 is chosen to coincide with Cooper, McConnell, and Ovtchinnikov (2006) to ensure our results are comparable to theirs. Cooper, McConnell, and Ovtchinnikov (2006) investigate earlier periods than this but begin their core analysis at this point because it represents the approximate start point that practitioners started to mention the January Barometer. Our international equity market data are the total return indices calculated by Morgan Stanley for 22 developed markets (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, United Kingdom). We also include the Morgan Stanley World Index and the Morgan Stanley US index. All these series cover the 1970 – 2006 period. These data are obtained from DataStream. We do not include data for emerging countries due to the lack of total return data for a reasonable time period.

We simply source all available MSCI data, however the start date of 1970 is very appropriate as it was the early 1970s that the January Barometer started to be mentioned in the media. The starting date of 1970 allows us to achieve a balance between focusing on a recent period, which is more relevant than a longer term historical period, and having enough data points to make meaningful conclusions. Since there is only one signal per year, focusing on short time frame, such as ten years, makes it impossible to carry out meaningful statistical analysis. In accordance with Cooper, McConnell, and Ovtchinnikov (2006), we focus our results presentation on excess returns, although the raw return results we generate in each instance are qualitatively similar. Our risk free rate proxies are 1-month Treasury Bill rates (or equivalent) from Global Financial Data. All international results are presented for indices in USD. We re-run all results in local currency terms and find they are qualitatively similar.

3.1. Basic Tests

We follow the approach of Cooper, McConnell, and Ovtchinnikov (2006) and test the statistical significance of the January Barometer using indicator or dummy variable regressions. As the regression specification in equation 1 indicates, we regress the 11-month (February – December) return on a dummy variable that equals one if the return in January is positive and zero if the January return is negative. Effectively, we are conducting a difference in means test between 11-month returns following positive and negative January returns.

$$R_t = \alpha_t + \beta_t D_t + \epsilon_t \tag{1}$$

Where $D_t = 1$ if January return is positive and 0 otherwise

We run this regression over six US series and 22 developed markets series and a World Market Index. Following Cooper, McConnell, and Ovtchinnikov (2006) we use the CRSP equally weighted (EW) and value weighted (VW) indices starting in 1940. We consider results for raw indices and excess indices, which we calculate as the raw index less the risk free rate. To ensure comparability with our international market results we also obtain the US total return index from Morgan Stanley (sourced from DataStream), which begins in 1970. All our series end in 2006, three years later than those of Cooper, McConnell, and Ovtchinnikov (2006). The regression specified in equation 1 is run over each series using three techniques to determine the standard errors: Ordinary Least Squares (OLS), Newey-West (1987) (NW), and the randomized bootstrap procedure of Cooper, McConnell, and Ovtchinnikov (2006) (BS). The bootstrap procedure is particularly important as promoters of the January Barometer may have considered using performance in non-January months as indicators for performance in the following 11 months in the first instance and only arrived at the January Barometer as this search. If this is the case then the OLS and Newey West (1987) techniques will overstate statistical significance as they do not account for data mining bias. The bootstrapping procedure adopted by Cooper, McConnell, and Ovtchinnikov (2006) explicitly controls for potential data mining bias by accounting for the other 11 months of the years as potential future 11 month performance indicators.

The results presented in Table 1 Panel A indicate there is an apparent January Barometer in the US. Appling this timing tool to raw and excess returns for both time periods generates superior returns. The spread between 11 month periods following positive January returns and negative January returns is as high as 19.73% (EW – Rf) and is above 11% in all series and time periods. The statistical significance is strong. The null hypothesis of a spread of zero is strongly rejected for each series by all three regression techniques. The bootstrapped p-values are less significant than their ordinary least squares and Newey West (1987) counterparts in the 1970 – 2006 period which indicated that some data mining bias may be present but they are still highly significant, so the January Barometer prevails after adjustment for this bias.

Our results are very similar to those of Cooper, McConnell, and Ovtchinnikov (2006). For instance, the spread we document for EW – Rf is 19.73% and theirs is 20.04%. For our VW – Rf series we find a spread of 14.58% whereas they document 14.71%. These minor differences can no doubt be attributed to our different data ending points. We use three more years of data (2004 – 2006) than they do.

We now consider how well the January Barometer works internationally. All data are sourced from DataStream. We report results for excess returns in USD but re-run all results for raw USD returns and local currency returns and find they are qualitatively similar. We generate results using the method described earlier and equation 1. The results displayed in Panel B of Table 1 indicate that the January Barometer is far less effective in markets outside the US. The spread is positive in all but four of the markets but the null hypothesis that the spread is no different to zero can only be rejected at the 10% level based on the Ordinary Least Squares, Newey West (1987), and Bootstrap approach in Spain and Switzerland. In addition, the null hypothesis can be rejected at the 10% level based on the Newey West (1987) and Bootstrap approach in the World Index. The difference between 11-month returns following positive and negative January returns is 18.63%, 13.47%, and 9.14% in Spain, Switzerland, and the World Index respectively.

[Please Insert Table 1 About Here]

As a further check to the robustness of the January Barometer we test the ability of other month "barometers" to predict 11-month returns. For instance, we see if returns in February can predict March – January returns and so on. These results, which are presented in the Appendix, suggest that non-January months in the US are poor predictors. No other month is statistically significant, with the exception of April which is only significant in one of our US series. The results are however quite the opposite for international markets. Of the three markets that have a statistically significant January Barometer, Spain and Switzerland each have two other months that appear to predict 11-month returns just as well as January does. Looking across all 22 international markets, it is evident that January is not the best predicating month. While there are only two countries and the World Index with a January Barometer, there are six countries with a February Barometer and five countries with a June Barometer. Taken together, these results are further evidence that the January Barometer works well in the US but performs poorly internationally.

3.2. Is the January Barometer Explained by Risk Adjustment?

As pointed out by Fama (1998), testing the robustness of an anomaly versus a benchmark model of expected returns effectively involves a joint test of the anomaly itself and the model that is being used to generate the expected returns. Therefore a finding of superior risk-adjusted returns to the anomaly may simply mean that the model being used is mis-specified. We consider the risk-adjusted returns of the January Barometer using two techniques which make very few assumptions about the underlying returns generation process.

The first method we apply is the Manipulation Proof Performance Measure (hereafter, MPPM) of Ingersoll, Spiegel, Goetzmann, and Welch (2007). These authors point out that traditional performance measures, such as Sharpe Ratio, Jensen's Alpha, and the Henriksson and Merton (1981) market timing techniques, each suffer from two weaknesses. Firstly, they are based around the assumption that return distributions are normal or lognormal. Secondly, they must be estimated using statistical techniques which assume independent and identically distributed variables. The MPPM, which is not dependent on these limiting assumptions, generates a score which is "(1) increasing in returns (to recognize arbitrage opportunities), (2) concave (to avoid increasing the score via leverage or adding unpriced risk), (3) time separable to prevent dynamic manipulation of the estimated statistics, and (4) has a power form to be consistent with an economic equilibrium." (Ingersoll, Spiegel, Goetzmann, and Welch, 2007, p. 1506).

The MPPM is given below:

$$\hat{\Theta} \equiv \frac{1}{(1-\rho)\Delta t} \ell n\left(\frac{1}{T} \sum_{t=1}^{T} \left[(1+r_t)/(1+r_{ft}) \right]^{1-\rho} \right) \tag{2}$$

In effect, the $\hat{\Theta}$ statistic is an estimate of the excess returns of a portfolio (over an above the risk-free asset) generated after adjusting for risk. The portfolio's un-annualized return at time t is r_t, and the risk-free rate is r_{ft}. T is the total number of observations, and Δt is the length of time between observations. Together these two variables annualize the measure. ρ is risk aversion coefficient. Higher values of ρ penalise risk more strongly.[5] We utilize the MPPM measure to test the risk-adjusted returns of the January Barometer by applying it a portfolio that is invested based on the January Barometer (i.e. returns are equal to the market return from February – December if the January return is positive and equal to the risk-free rate if January is negative) and a portfolio invested based on a buy-and-hold strategy (i.e. returns are equal to the market portfolio. In accordance with Ingersoll, Spiegel, Goetzmann, and Welch (2007) we test three different risk-aversion coefficients ($\rho = 2$, $\rho = 3$, $\rho = 4$).

[5] The interested reader should refer to Ingersoll, Spiegel, Goetzmann, and Welch (2007) for a more detailed description of the MPPM.

The MPPM results presented in Table 2 Panel A provide strong evidence that the superior performance of the January Barometer in the US is not compensation for risk. The January Barometer generates a larger excess return over and above the risk-free rate than a buy-and-hold strategy in each of the three US indices we consider. For instance, based on ρ = 2 the January Barometer generates a MPPM of 5.01% in the CRSP EW index compared to 2.57% for a buy-and-hold strategy. In other words, the extra excess return accruing to someone adopting the January Barometer rather than a buy-and-hold strategy is 3.28%. As expected the out-performance of the January Barometer improves further as the risk-aversion coefficient is increased to 3 and 4. The out-performance increases to 3.09% and 3.77% respectively. The January Barometer signals investors to spend 11 month periods out of the market following a negative January return so someone adopting it as a timing technique incurs less risk than they would with a buy-and-hold strategy.

Consistent with our US analysis we use the MPPM measure to investigate the risk-adjusted performance of the January Barometer in international markets. The results presented in Table 2 Panel B indicate that, on the whole, the January Barometer out-performs a buy-and-hold strategy after risk is taken into account. As with the US results, it is evident that a trading strategy based around the January Barometer is superior to a buy-and-hold strategy. Based on the lowest risk-aversion coefficient (ρ = 2) just 5 of the 23 markets have higher risk-adjusted buy-and-hold returns than those generated by the January Barometer. When the highest risk-aversion coefficient (ρ = 4) is applied only 1 of the 23 markets has higher risk-adjusted buy-and-hold returns. Overall, we conclude that there is strong evidence that risk does not explain the performance of the January Barometer in either the US or international markets.

[Please Insert Table 2 About Here]

Our second risk adjustment technique is based the principles of stochastic dominance. This technique, which is discussed extensively by Anderson (1996), Davidson and Duclous (2000) and Barrett and Donald (2003), uses a very general framework to assess portfolio choice which is not reliant on asset pricing benchmarks.[6] In our case there are two hypotheses as follows: The first tests whether a positive January return indicator is superior to a negative

[6] The stochastic dominance specification we adopt comes from Barrett and Donald (2003) so we refer the interested reader to this paper for more detail on the description we provide.

January indicator for predicting a remainder of the year return (H_0: JB positive SD JB negative). The second is used to test whether a negative January return indicator is superior to a positive January return indicator (H_0: JB negative SD JB positive). Each null hypothesis is applied for stochastic dominance of order one, two, and three. The rejection of the negative January dominance hypothesis combined with the non-rejection of the positive January dominance hypothesis would confirm the dominance of positive January indicator.

In Table 3 Panel A we present the p-value of the bootstrapped Komogorov-Smirnov test statistics for our US indices. The first part of our test has a null hypothesis that a positive January return indicator is superior to a negative January indicator for predicting a remainder of the year return. As we can see in Table 3, the p-values of the positive January dominance hypothesis are all greater than 0.10. For instance, the p-values of equally weighted and value-weighed market returns from 1940-2006 for the second-order stochastic dominance test are 0.226 and 0.146. This suggests that there is no evidence supporting the rejection of positive January dominance hypotheses.

Nevertheless, such findings are insufficient to suggest the dominance of a positive January indicator compared to a negative January indicator. As a result we have to test whether a negative January return indicator is superior to a positive January return indicator. The rejection of the negative January dominance hypothesis combined with the non-rejection finding of the positive January dominance hypothesis would confirm the dominance of positive January indicator. We find consistent evidence across all market returns rejecting the superiority of negative January return indicator hypotheses. For the second-order stochastic dominance test, we find that the p-values of an equally weighted and value weighted returns are 0.000 and 0.001 which are well below 0.05. Hence, the findings based on the stochastic dominance test clearly indicate that the 11-month returns following positive January returns statistically dominate the 11-month returns following the negative January return.

The international stochastic dominance test results are shown in Table 3 Panel B. These indicate that we can not reject the null hypothesis that a positive January return indicator is superior to a negative January indicator for predicting the return for the remaining 11 months of the year in any countries. On the other hand, we also unable to reject the null hypothesis that a negative January return indicator is superior to a positive January return indicator for all orders of stochastic dominance at the 5% level in each market except Norway, Spain, and Switzerland. The rejection of both hypotheses is required for us to conclude that a positive January return is superior to a negative January return for predicting the return for the remaining 11 month of the years so we are only able to conclude this for these three countries.

[Please Insert Table 3 About Here]

3.3. Is the January Barometer Explained by Dummy Variable Persistence or Data Mining?

Recent work by Powell, Shi, Smith, and Whaley (2007) shows that standard OLS tests involving dummy variables that are persistent can lead to spurious inferences. This is a particular problem when long-standing regimes are compared using data of a higher frequency than the regime itself. For instance, using monthly data and dummy variable analysis to investigate the presidential puzzle ensures dummy variable "runs" of either zeros or ones of a minimum of 48 observations (i.e. one observation per month for a four year presidential term). The dummy variable used in our analysis relates to the return in the month of January, and we find evidence that this is persistent. For example, in the US between 1983 and 2006 there were only two years with negative January returns compared to 22 years with positive January returns.

We therefore apply the Powell, Shi, Smith and Whaley (2007) approach, which involves determining whether the coefficient significance levels estimated in our equation 1 are less than would be expected by chance. This involves simulating cut-off values of the coefficients, the t-statistics and the $\bar{R}^2$. The dependent return variable is simulated based on a starting point of an unconditional mean and variance of zero and one respectively, and an error term drawn from a normal distribution with a mean of zero and a variance of $1 - \rho^2$, where ρ is the first-order correlation coefficient of the dependant return variable. The dummy variable is generated based on a transition matrix which accounts for the conditional probability of remaining in a dummy variable state. Ninety five percent confidence bands are created for the coefficients, the t-statistics and the $\bar{R}^2$ by running 10,000 simulations of the process outlined above. A vector of estimates is formed for each of coefficients, t-statistics and the $\bar{R}^2$. The contents of each vector are ranked from lowest to highest and the 2.5th and

97.5th percentile coefficient estimates and t-statistics are recorded as the 5% critical cut-off values. The 95th percentile $\overline{R}^2$ is used for this statistic.[7]

Powell, Shi, Smith and Whaley (2007) also advocate calculating a second set of cut-off statistics based on the Bonferonni correction to account for data mining. In the presidential puzzle setting it is difficult to determine the appropriate factor to adjust the critical cut-off values by, but in our analysis it is more obvious. There are twelve months in a year it is possible that the eleven other months we examined before arriving at the January Barometer. We thus use a factor of 12. This is equivalent to requiring a significance level of 0.42% (5% / 12).

The results presented in Table 4 Panel A contain the coefficient estimates, associated t-statistics, $\overline{R}^2$, and the corresponding simulated lower and upper cut-off values for our US indices. The results suggest that spurious regression due to dummy variable persistence does not explain all the success of the January Barometer in the US. The dummy variable (β) coefficients are outside the simulated bands for each of our six series and the actual t-statistics for the beta coefficient are also outside the bands. The results are however less consistent when spurious regression and data mining adjusted cut-off points are used. The January Barometer dummy variable coefficient falls outside the spurious regression and data mining range for all four of the series relating to the 1940 – 2006 period, and the associated t-statistics fall outside the spurious regression and data mining range for three of the four series. However, both coefficients and t-statistics for the more recent series (1970 – 2006) are contained within the spurious regression and data mining range. In summary, we conclude that data mining bias is only a partial explanation for the success of the January Barometer in the US. The results in Table 1 relating to the Cooper, McConnell, and Ovtchinnikov (2006) data mining adjustment procedure, which indicated that all US series survive adjustment for data mining bias reinforce this view.

We now turn our attention to the international results. Given that the dummy variable coefficient is only statistically significant in our core tests (Table 1 Panel B) in the case of Spain, Switzerland, and the World Index, we only consider these three markets at this point. (Table 4 Panel B) includes the coefficients, t-statistics, and $\overline{R}^2$, and associated lower and upper cut-off points. The results indicate that the statistical significance of the January

[7] The interested reader should refer to Powell, Shi, Smith and Whaley (2007) for a more detailed description of the spurious regression adjustment techniques they develop.

Barometer in Spain is not due to persistence in the dummy variable. The coefficient of the January Barometer dummy variable and associated t-statistic both fall outside the spurious regression bands. However, in the case of both Switzerland and the World Index both the coefficient and associated t-statistic are contained within the spurious regression bands so we are unable to reject the null hypothesis that the January Barometer in both these markets is due to dummy variable persistence. Returning, to the case of Spain, we see that both the dummy variable coefficient and t-statistics are within the spurious regression and data mining bands so we cannot reject the null hypothesis that the performance of the January Barometer in Spain is due to dummy variable persistence and data mining bias. Overall, we conclude the evidence against the ability of the January Barometer in international markets is overwhelming. There is no statistically significant relationship in 20 of the 23 markets and in the three markets there is a relationship there is evidence it could simply be due to spurious regression techniques.

[Please Insert Table 4 About Here]

4. Robustness Checks

In this section we subject the performance of the January Barometer in the US to a range of robustness tests. We consider its performance in different size deciles, book-to-market equity deciles, and industry groups. We determine whether time-varying risk premia explains its performance. Finally, we investigate how individual stocks within the market index behave from the perspective of the January Barometer. We understand that supporters of the January Barometer suggest that it should be applied to the market in general rather than individual stocks, but we suggest that if it is in fact a strong predicator in the US then we would expect to see those stocks that have positive January returns out-perform those stocks that have negative January returns for the remaining 11 months of the year.

4.1. Does the January Barometer work in different Size, Book-to-Market and Industry Portfolios?

Our size, book-to-market equity, and industry results are presented in Panels A, B, and C of Table 5 respectively. All data are derived from CRSP for the period 1940-2006 and are sourced from the Ken French website. We report excess returns in each instance. The Panel A results indicate that the January Barometer is profitable in each of the size decile indices we consider. It is most profitable on the stocks within the smallest decile (spread of 24.89%), but all spreads are over 5%. There is no clear pattern of different profitability across the size deciles. Our book-to-market equity results indicate the January Barometer is most profitable in low book-to-market equity or value stocks (spread of 13.90%), but it remains profitable in all other book-to-market deciles apart from the decile with the highest book-to-market value. The January Barometer does not appear to work in the Shops, Technology, and Consumer NonDurables industry, but it is profitable in each of the other industries. It is most profitable in the Consumer Durables industry (spread of 12.62%).

[Please Insert Table 5 About Here]

4.2. Is Time-Varying Risk Premia an Explanation?

Driesprong, Jacobsen, and Maat (2008) highlight that return predictability may in fact be driven by time-varying equilibrium expected returns. As a result, it is important to determine whether return predictability is indicative of market inefficiency or is simply a result of time-varying risk premia during the business cycle. One approach to this is to include well-known business cycle related variables, such the default spread and term spread (see Fama and French (1989)) in a regression environment to determine whether the anomaly in question remains robust after controlling for these. However, there is also a second, stricter approach. Schwert (2003) suggests that the strongest test involves determining whether excess returns are predictively negative, as such a finding cannot be attributed to risk. Schwert (2003) documents many instances of supposed anomalies not meeting this criterion (i.e. they are successful at predicting positive excess returns by relatively unsuccessful at predicting negative expected returns.) However, there is also evidence of anomalies passing this test. For instance, Driesprong, Jacobsen, and Maat (2008) show that changes in the price of Oil are just as good at predicting negative excess returns as they are at predicting positive excess returns.

We apply the Schwert (2003) test to the January Barometer by investigating the ability of negative January months to predict negative excess returns for the remaining 11 months of the year. We use both equally weighted and value weighted CRSP indices over the 1940-2006 period and apply both parametric and non-parametric tests. The results in Table 6 Panel A indicate that we are unable to reject the null hypothesis that the January Barometer is simply a result of time varying risk premia at the 5% level in the equally weighted index and at the

10% level in the value weighted index. Overall, we conclude that there is no conclusive evidence that the January Barometer is not due to time-varying risk premia.

[Please Insert Table 6 About Here]

4.3. Are US Institutional Factors an Explanation?

We now focus on the results generated from applying the January Barometer to individual stocks. We begin this section with a stylized example of how entirely different levels of performance by the January Barometer in individual stocks can result in an identical view of the accuracy of the January Barometer if one solely considers an index. We realize that supporters of the January Barometer suggest that it should be applied to the market in general rather than individual stocks. However, we believe that there should be a pattern of the individual stocks that have positive (negative) January returns being those stocks that have positive (negative) 11-month returns if there is an institutional reason behind the January Barometer. In order to illustrate that an accurate reading for the January Barometer at the market level in any given year does not necessarily imply accuracy among individual stocks we consider a simple setting where there are just four stocks, although our logic applies equally well to settings where there are numerous stocks.

[Please Table 7 About Here]

Under Scenario A the January Barometer is 100% accurate on individual stocks. Each time a stock records a positive (negative) return in January it goes on to record a positive (negative) return for the remaining 11 months of the year. A market index would also show that the January Barometer was accurate in this particular year (i.e. the January market return is positive at 1 % and the February – December market return is also positive at 10%) We assume an equally weighted market index to keep things simple but our logic holds for a market weighted index as well. In Scenario B the January Barometer is 100% inaccurate. Each time a stock records a positive January return it goes on to experience a negative February – December return and vice versa. However, despite the inaccurate performance at stock level, the January Barometer still has precisely the same performance at index level. As in Scenario A the index gains 1% in January and 10% for the year. Finally, in Scenario C the January Barometer is accurate on individual stocks 50% of the time, yet at the market level its performance is identical to Scenarios A and B. We contend that this indicates that the performance of the January Barometer in an index does not imply anything about its performance in individual stocks.

In order to test the performance of the January Barometer in individual stocks we use all stocks in the CRSP database for the 1940-2006 period and report excess returns. In Figure 1 we graph, for each year, the mean 11-month returns for stocks that had a positive return in January (light colored bars) versus the mean 11-month returns for stocks that had a negative return in January (dark colored bars). If there is an institutional reason behind the apparent success of the January Barometer in the US it seems reasonable to expect this to show up in individual stocks. In other words, one would expect the 11-month returns for stocks with a positive return in January to be higher than the 11-month returns for stocks with a negative return in January, on average. The results presented in Figure 1 illustrate that this is not the case. There is no clear trend of higher 11-month returns for positive January return stocks. The first six years of data have a pattern which is indicative of the entire data set. Positive January return stocks have lower 11-month returns (than their negative January return counterparts) in 1940, 1941, 1946 (three out of the six year). Over our entire 67 year period, stocks with positive January returns only have 11 months returns higher than their negative January return counterparts in 31 years, or 46% of the time.

[Please Insert Figure 1 About Here]

We present the results of the formal comparison of the mean and median 11-month returns for individual stocks that have positive and negative January returns in Table 8. This analysis includes all stocks in the CRSP database over the 1940-2006 period. We present results for excess returns. Over the 67 year period we study the mean 11-month return for stocks that had a positive return in January is 7.42%, compared to 6.52% for stocks that had a negative January return. This difference is not statistically significant. In the sub-samples of years where the market was up in January and down in January the 11-month returns for stocks with a positive January are slightly higher than the 11-month returns for stocks with a negative January, but these differences are also not significant.

The median results in Panel B also strongly indicate that there is no statistically significant difference between 11-month returns for stocks with positive and negative January returns. Over all 67 years stocks with positive January returns actually under perform stocks with negative January returns. This situation is reversed in the sub-sample of years with positive market performance in January and negative market performance in January. However, none of these differences are statistically significant based on the non-parametric Wilcoxon test.

[Please Insert Table 8 About Here]

In unreported results, we conduct an additional test into differences in the predictive power of stocks with positive January performance versus those with negative January performance. More specifically, for all years where the market is positive in January we record the proportion of stocks with a positive January return that have a positive 11-month return. For all years where the market is negative in January we record the proportion of stocks with a negative January return that have a negative 11-month return. We then conduct a simply binomial test in both instances, based on the null hypothesis that if the January Barometer has no predictive power at the individual stock level then the proportion of "successful" predictions in each instance should not be different to 0.5 (i.e. the level expected by chance). We are not able to reject this null hypothesis in either instance.

Overall, we conclude that the results presented in Table 8, together with the unreported binomial test results, provide strong evidence that there is no intuitional explanation for the performance of January Barometer in US stock market indices. There is no pattern of stocks that have positive January returns outperforming those with negative January returns for the remaining 11 months of the year.

4.4. Are US Investor Behavioral Biases an Explanation?

We conclude our analysis by considering whether it is possible that there is a behavioral explanation for the apparent succuss of the January Barometer in US equity indices. We deem this unlikely, but apply a simple test to check. More specifically, we investigate whether the January Barometer works on a range of US traded commodities series. Previous researchers have shown that anomalies that exist in equity markets also exist in commodity markets. Wang and Yu (2004) show that short term contrarian strategies produce similar profits in commodity markets to those documented in equity markets by Lehmann (1990) and Lo and MacKinlay (1990b). Miffre and Rallis (2007) find that the momentum strategies of Jegadeesh and Titman (1993) are highly profitable in commodity markets. This is unsurprising because if there is a behavioral reason for their existence, like the underreaction hypothesis of Hong and Stein (1999) in the case of momentum, one would expect this explanation of investor behaviour to hold across different markets.

[Please Insert Table 9 About Here]

We apply the January Barometer to 21 commodity series, which we source from Global Financial Data. Each series starts at different points - the earliest is Corn in 1947 and the latest is Natural Gas in 1998. We report the results for excess returns in each instance. The results presented in Table 9 indicate that the spread between 11-month returns following positive and negative Januaries is not statistically significant form zero at the 10% level in each of the three significance tests we apply in any of the commodity series. The spread is positive, and at times large, in all but five of the commodity series, but it is not statistically significant. We interpret this as further evidence against the predictive power of the January Barometer.

5. Conclusions

We investigate the performance of the January Barometer in the US and 22 other equity markets. The worth of this indicator, which suggests that positive (negative) returns in the month of January predict positive (negative) returns in the remaining 11 months of the year, is constantly mentioned in the popular press but has received surprising little attention in the academic literature.

We present comprehensive evidence that the January Barometer is not a reliable return predictor in any of the markets we consider. Of the 23 non-US series (22 countries and MSCI World Index) we consider 19 have, on average, larger 11-month returns following positive Januaries then negative Januaries. However, the differences in 11-month returns following positive and negative Januaries are only statistically significant in two countries – Spain and Switzerland, and the World Index. The results in both Spain and Switzerland are robust to adjustment for risk using stochastic dominance techniques and the manipulation-proof performance measure, but the World Index result does not. However, the results in all three markets cease to exist following joint adjustment for dummy variable persistence and data mining bias.

In contrast, the January Barometer appears to reliable in the US as it remains profitable after risk adjustment based on stochastic dominance and manipulation-proof performance techniques and cannot be fully explained by dummy variable persistence or data mining bias. However, we are unable to rule out time-varying risk premia as an explanation for the January Barometer in US indices. Further, our finding of a lack of institutional or behavioral reason for the apparent success of the January Barometer in the US cast major doubt on regarding its accuracy.

There is no obvious institutional reason why the January Barometer should work in the US but not in other markets that we are aware of, but we investigate this further by focusing on different size, and book-to-market equity deciles, and different US sectors. We find that the January Barometer works reasonably consistently in these subsets of the entire market. We then focus on individual US stocks. Proponents of the January Barometer claim it works on the market in general rather than individual stocks, but we suggest that if it is in fact a strong predicator in the US then we would expect to see those stocks that have positive January returns out-perform those stocks that have negative January returns for the remaining 11 months of the year. We find no evidence of this.

To conclude our analysis we consider whether the January Barometer holds in US traded commodities. This would be expected if there is a genuine behavioural reason for its existence. We find no evidence of predictability for the January Barometer in this market. Taken together, we suggest that our results provide support for the conclusion that the profitability of the January Barometer in US equity indices is simply due to chance.

References

Anderson, G. (1996). Nonparametric Tests for Stochastic Dominance, *Econometrica*, 64, 1183-1193.

Barberis, N. and Thaler, R. (2002). A survey of behavioural finance, Handbook of the economics of finance, edited by George Constantinides, Milt Harris and Rene Stulz.

Barrett, G. and Donald, S. (2003). Consistent Tests for Stochastic Dominance, *Econometrica*, 71(1), 71-104.

Bohl, M. and Salm, C. (2008). *The Other January Effect: Nothing More than a Statistical Artifcat*. Working Paper, Westfalische Wilhelms-University Munster, Germany

Brown, L. and Luo, L. (2006). The January Barometer. Further Evidence. *Journal of Investing*, Spring, 25-31.

Cooper, M., McConnell, J. and Ovtchinnikov, A. (2006). The Other January Effect. *Journal of Financial Economics*, 82, 315-341.

Davidson, R. and Duclos, J.Y. (2000). Statistical Inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, 68 (6), 1435–1464.

Driesprong, G., Jacobsen, B. and Maat, B. (2008). Striking Oil: Another Puzzle? *Journal of Financial Economics* – forthcoming.

Easton, S. and Pinder, S. (2007). A Refutation of the Existence of the Other January Effect. *International Review of Finance*, 3-4 (7), 89-104.

Fama, E. (1970). Efficient Capital Markets, A Review of Theory and Empirical Work. *Journal of Finance*, 25, 383-417.

Fama, E. (1998). Market Efficiency, Long-term Returns, and Behavioural Finance. J*ournal of Financial Economics*, 49, 283-306.

Fama, E., and MacBeth, J., (1973). Risk, Return, and Equilibrium: Empirical Tests. The *Journal of Political Economy*, 81(3), 607-636

Fama, E., and French, K. (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-49.

Fama, E., and French, K. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3–56.

Ferson, W. (2006). Note from the Editor, Wayne E. Ferson on Shanken, Jay and Mark I. Weinstein, Economic Forces and the Stock Market Revisited, Journal of Empirical Finance 13, Issue 2, 2006, 129–144, *Journal of Empirical Finance*, Volume 13, Issue 3, June 2006, Pages 389-391

Fuller, R., (1978). The January Barometer: What's Its Batting Average. *Journal of Portfolio Management*, Winter, 5–7.

Goetzmann, W.N. and Jorion, P. (1999). Global Stock Markets in the Twentieth Century, *Journal of Finance*, 54(3), 953-980.

Henriksson, R.D. and Merton, R.C. (1981). On the Market Timing Ability and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business*, 54, 513-533.

Hensel, C. and Ziemba, W. (1995a). The January Barometer. *Journal of Investing*, Summer, 67-70.

Hensel, C. R. and W. T. Ziemba (1995b). The January barometer: European, North American, Pacific and worldwide results. *Finanzmarkt und Portfolio Management* 9 (2), 187–196.

Hirsch, Y., 1974. Stock Trader's Almanac. The Hirsch Organization, Nyack, NY.

Hong H. and Stein J.C. (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance*, 54, 2143-2148.

Ingersoll, J., Spiegel, M., Goetzmann, W., and Welch, I. (2007). Portfolio Performance Manipulation and Manipulation-proof Performance Measures, *Review of Financial Studies*, 20(5), 1503-1546.

Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48, 65-91.

Jensen, M. (1978). Some Anomalous Evidence Regarding Market Efficiency. *Journal of Financial Economics*, 6, 95-101.

Kahneman, D. and Tversky, A. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.

Lakonishok, J. and Smidt, S. (1988). Are Seasonal Anomalies Real? A Ninety-year Perspective. *Review of Financial Studies*, 1, 403-425.

Lehmann, B. (1990). Fads, Martingales, and Market Efficiency. *Quarterly Journal of Economics*, 105, 1-28.

Lesmond, D., Schill, M. and Zhou, C. (2004). The Illusory Nature of Momentum Profits. *Journal of Financial Economics*, 71, 349-380.

Little, I. and Albrecht, B.(2006). *The Fallacy of the Price Earnings Ratio ("PER")*, *Global Thematic Investors Newsletter*, April 2006, 1-14.

Lo, A.W., MacKinlay, A.C. (1990a). Data-Snooping Biases in Tests of Financial Asset Pricing Models. *Review of Financial Studies*, 3, 431–467.

Lo, A. W., MacKinlay, A.C. (1990b). When are Contrarian Profits due to Stock Market Overreaction? R*eview of Financial Studies* 3, 175-208.

Miffre, J., Rallis, G. (2007). Momentum Strategies in Commodity Futures Markets. *Journal of Banking and Finance* 31, 1863-1886.

Newey, W. and West, K. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance, *Econometrica*, 55(3), 703-708.

Powell, J, Shi, J, Smith, T., and Whaley, R. (2007). The Persistent Presidential Dummy. *Journal of Portfolio Management*, Winter, 133-143.

Rouwenhorst, K.G. (1998). International Momentum Strategies. *Journal of Finance*, 53, 263-284.

Schwert, G.W. (2003). Anomalies and Market Efficiency, In: The Handbook of the Economics of Finance, Edited by Constantinides, G.M., M. Harris and R. Stulz, Chapter 15, 2003, 939-972.

Wang, C., Yu, M. (2004). Trading Activity and Price Reversals in Futures Markets. *Journal of Banking and Finance*, 28, 1337-1361.

Table 1. Core Results

		Positive Jan Return	N	Negative Jan Return	N	Spread	P-Value OLS	P-Value NW	P-Value BS
		Panel A: US (11-Month Return from Feb-Dec)							
EW	1940-2006	14.09	54	-2.98	13	17.06	0.01	0.00	0.01
VW	1940-2006	14.62	44	3.25	23	11.37	0.00	0.00	0.00
EW - Rf	1940-2006	11.18	52	-8.54	15	19.73	0.00	0.00	0.00
VW - Rf	1940-2006	11.93	43	-2.65	24	14.58	0.00	0.00	0.00
MSCI	1970-2006	14.43	23	2.63	14	11.80	0.01	0.01	0.02
MSCI - Rf	1970-2006	8.60	23	-3.15	14	11.75	0.01	0.01	0.03
		Panel B: International Markets (11-Month Return from Feb-Dec)							
Australia	1970-2006	2.35	22	3.02	15	-0.67	0.91	0.92	0.65
Austria	1970-2006	8.32	18	2.32	19	6.00	0.49	0.46	0.65
Belgium	1970-2006	7.98	24	0.13	13	7.84	0.26	0.28	0.28
Canada	1970-2006	2.45	21	3.85	16	-1.40	0.79	0.77	0.76
Denmark	1970-2006	6.92	25	-3.06	12	9.98	0.31	0.21	0.47
Finland	1988-2006	22.84	13	-10.46	6	33.30	0.17	0.06	0.24
France	1970-2006	2.22	26	8.61	11	-6.39	0.47	0.50	0.84
Germany	1970-2006	6.92	23	5.76	14	1.15	0.89	0.88	0.66
Greece	2002-2006	19.86	3	2.51	2	17.35	0.62	0.59	0.44
Hong Kong	1970-2006	15.50	24	14.55	13	0.95	0.95	0.95	0.63
Ireland	1988-2006	7.00	15	-5.34	4	12.34	0.32	0.26	0.27
Italy	1970-2006	2.24	26	-6.93	11	9.16	0.32	0.26	0.45
Japan	1970-2006	6.27	24	4.42	13	1.84	0.84	0.83	0.56
Netherlands	1970-2006	8.13	26	-0.56	11	8.69	0.20	0.19	0.28
New Zealand	1988-2006	1.03	9	-0.68	10	1.70	0.85	0.85	0.45
Norway	1970-2006	12.09	25	-8.12	12	20.21	0.10	0.02	0.23
Portugal	1988-2006	4.43	14	-12.48	5	16.91	0.20	0.11	0.30
Singapore	1970-2006	8.84	27	7.94	10	0.90	0.96	0.94	0.74
Spain	1970-2006	9.10	25	-9.53	12	18.63	0.02	0.01	0.08
Sweden	1970-2006	6.67	27	10.78	10	-4.11	0.69	0.72	0.70
Switzerland	1970-2006	10.82	23	-2.65	14	13.47	0.05	0.04	0.07
United Kingdom	1970-2006	5.82	23	-0.83	14	6.65	0.29	0.38	0.15
World	1970-2006	4.65	26	-4.49	11	9.14	0.10	0.06	0.07

EW and VW are the CRSP equally weighted and value weighted indices respectively. MSCI is the US total return index calculated by Morgan Stanley. All our international country indices are also based on MSCI USD total return data. These data are sourced from DataStream. Our risk free rate proxies are 1-month Treasury Bill rates (or equivalent) from Global Financial Data. All international results are for excess returns. We calculate the 11-month return from February to December following positive (negative) January returns and then calculate the spread, which is the difference in the 11-month returns. The statistical significance of the spread (null hypothesis is spread equals 0) is determined using OLS t-statistics (OLS), Newey West (1987) t-statistics (NW) and bootstrapped t-statistics (BS), which are calculated using the method advocated by Cooper, McConnell, and Ovtchinnikov (2006). All returns are presented as percentages.

Table 2. Portfolio Performance Manipulation-Proof Performance Measures

	$\rho = 2$			$\rho = 3$			$\rho = 4$		
	MPPM (JB)	MPPM (Mkt)	Diff	MPPM (JB)	MPPM (Mkt)	Diff	MPPM (JB)	MPPM (Mkt)	Diff
Panel A: US									
EW	5.01	2.57	2.44	3.28	0.19	3.09	1.48	-2.29	3.77
VW	5.81	4.77	1.04	4.98	3.59	1.40	4.09	2.33	1.76
MSCI	4.85	2.28	2.57	4.43	1.06	3.37	4.00	-0.26	4.25
Panel B: International Markets									
Australia	0.15	-0.92	1.07	-1.30	-3.01	1.71	-2.97	-5.24	2.27
Austria	0.04	0.70	-0.67	-1.09	-1.17	0.09	-2.09	-2.79	0.70
Belgium	2.59	1.75	0.84	1.83	-0.17	2.00	1.12	-2.08	3.20
Canada	0.97	0.68	0.28	0.02	-0.72	0.73	-1.02	-2.18	1.17
Denmark	-0.52	-3.15	2.63	-2.19	-6.41	4.22	-3.74	-9.43	5.70
Finland	-0.74	-2.85	2.11	-5.01	-9.37	4.36	-9.40	-15.24	5.84
France	-2.46	-1.59	-0.86	-4.44	-4.72	0.28	-6.58	-7.86	1.28
Germany	1.55	1.17	0.38	0.54	-1.94	2.48	-0.44	-5.22	4.78
Greece	6.39	3.73	2.66	6.01	-2.45	8.45	5.65	-8.90	14.56
Hong Kong	-0.09	-0.68	0.59	-8.37	-10.91	2.54	-19.14	-22.34	3.20
Ireland	0.82	-0.54	1.36	-1.59	-3.65	2.05	-4.29	-6.98	2.69
Italy	-2.28	-7.21	4.93	-4.34	-10.38	6.03	-6.35	-13.35	7.00
Japan	4.41	0.50	3.90	3.21	-2.11	5.32	2.09	-4.71	6.79
Netherlands	3.14	2.21	0.93	2.37	0.20	2.17	1.63	-1.89	3.53
New Zealand	1.09	-3.81	4.90	0.71	-5.96	6.67	0.33	-8.26	8.60
Norway	-1.67	-4.30	2.63	-5.55	-9.16	3.61	-9.89	-14.03	4.13
Portugal	-1.38	-6.80	5.42	-3.51	-10.31	6.80	-5.71	-13.75	8.04
Singapore	-4.83	-3.77	-1.06	-10.48	-9.99	-0.49	-16.48	-16.39	-0.08
Spain	3.19	-1.76	4.95	1.95	-4.21	6.16	0.76	-6.54	7.30
Sweden	1.35	1.53	-0.17	-0.04	-1.83	1.79	-1.40	-5.14	3.75
Switzerland	-0.18	1.65	-1.82	-1.53	-0.75	-0.77	-3.00	-3.24	0.24
United Kingdom	2.24	-0.87	3.11	1.73	-4.13	5.86	1.25	-8.47	9.73
World	0.61	-0.82	1.43	-0.44	-2.54	2.10	-1.62	-4.42	2.80

EW and VW are the CRSP equally weighted and value weighted indices respectively. MSCI is the US total return index calculated by Morgan Stanley. All other equity series are also total return indices in US dollars. Our risk free rate proxies are 1-month Treasury Bill rates (or equivalent) from Global Financial Data. We utilize the MPPM measure to test the risk-adjusted returns of the January Barometer. We apply the MPPM to a portfolio that is invested based on the January Barometer (JB) and a portfolio invested based on a buy-and-hold strategy (Mkt). We then calculate the difference in risk-adjusted returns. In accordance with Ingersoll, Spiegel, Goetzmann, and Welch (2007) we test three different risk-aversion coefficients ($\rho = 2$, $\rho = 3$, $\rho = 4$). All numbers are presented as percentages.

Table 3. Stochastic Dominance Results

	H_0: JB positive SD JB negative			H_0: JB negative SD JB positive		
	SD1	SD2	SD3	SD1	SD2	SD3
			Panel A: US			
EW	0.57	0.23	0.37	0.00	0.00	0.00
VW	0.20	0.15	0.38	0.00	0.00	0.01
EW - Rf	0.99	0.74	0.70	0.01	0.00	0.00
VW - Rf	1.00	0.74	0.69	0.00	0.00	0.00
MSCI	0.99	0.71	0.66	0.01	0.01	0.01
MSCI - Rf	0.92	0.71	0.65	0.01	0.01	0.01
			Panel B: International Markets			
Australia	0.72	0.46	0.35	0.71	0.74	0.69
Austria	0.53	0.41	0.43	0.40	0.25	0.43
Belgium	0.92	0.75	0.70	0.30	0.14	0.09
Canada	0.44	0.44	0.36	0.62	0.72	0.66
Denmark	0.96	0.75	0.71	0.52	0.14	0.17
Finland	0.95	0.68	0.64	0.17	0.03	0.03
France	0.51	0.28	0.38	0.84	0.64	0.63
Germany	0.36	0.72	0.68	0.46	0.32	0.34
Greece	0.55	0.44	0.44	0.55	0.25	0.25
Hong Kong	0.81	0.61	0.57	0.17	0.35	0.33
Ireland	0.97	0.61	0.58	0.25	0.13	0.19
Italy	0.95	0.73	0.69	0.33	0.14	0.18
Japan	0.69	0.73	0.68	0.49	0.41	0.40
Netherlands	0.97	0.71	0.65	0.28	0.10	0.12
New Zealand	0.69	0.69	0.64	0.45	0.28	0.23
Norway	0.97	0.63	0.60	0.03	0.02	0.02
Portugal	1.00	0.67	0.62	0.24	0.07	0.10
Singapore	0.28	0.32	0.34	0.70	0.68	0.75
Spain	1.00	0.78	0.74	0.03	0.00	0.00
Sweden	0.04	0.37	0.67	0.40	0.31	0.37
Switzerland	0.97	0.64	0.60	0.05	0.02	0.03
United Kingdom	0.50	0.69	0.64	0.24	0.13	0.11
World	0.98	0.67	0.63	0.15	0.06	0.08

EW and VW are the CRSP equally weighted and value weighted indices respectively. MSCI is the US total return index calculated by Morgan Stanley. All other equity series are also excess total return indices in US dollars. Our risk free rate proxies are 1-month Treasury Bill rates (or equivalent) from Global Financial Data. This Table reports the p-value of the bootstrapped Komogorov-Smirnov test statistics for two null hypotheses. The first tests whether a positive January return indicator is superior to a negative January indicator for predicting a remainder of the year return (H_0: JB positive SD JB negative). The second is used to test whether a negative January return indicator is superior to a positive January return indicator (H_0: JB negative SD JB positive). The rejection of the negative January dominance hypothesis combined with the non-rejection finding of the positive January dominance hypothesis would confirm the dominance of positive January indicator. SD1, SD2, SD3, refer to first, second, and third orders of stochastic dominance respectively.

Table 4. Spurious Regression and Data Mining Adjustment Results

		N	α	$t(\alpha)$	β	$t(\beta)$	$\bar{R}^2$
				Panel A: US			
EW		67	-0.030	-0.52	0.170	2.69	8.62%
	Spurious regression bias		0.036 / 0.185	0.93 / 5.07	-0.104 / 0.103	-1.99 / 1.97	
	Spurious regression bias and data mining		0.000 / 0.219	0.01 / 6.19	-0.151 / 0.151	-2.95 / 2.94	
VW		67	0.0325	1.15	0.114	3.26	12.71%
	Spurious regression bias		0.0588 / 0.1603	2.24 / 6.57	-0.070 / 0.070	-1.98 / 1.99	
	Spurious regression bias and data mining		0.0356 / 0.1839	1.35 / 7.73	-0.102 / 0.102	-2.96 / 2.98	
EW - Rf		67	-0.085	-1.6	0.197	3.25	12.69%
	Spurious regression bias		-0.007 / 0.147	-0.18 / 3.89	-0.105 / 0.105	-1.94 / 1.96	
	Spurious regression bias and data mining		-0.044 / 0.181	-1.13 / 4.93	-0.151 / 0.153	-2.89 / 2.90	
VW - Rf		67	-0.027	-0.97	0.146	4.28	20.82%
	Spurious regression bias		0.016 / 0.122	0.61 / 4.77	-0.072 / 0.072	-1.96 / 1.98	
	Spurious regression bias and data mining		-0.009 / 0.146	-0.33 / 5.87	-0.106 / 0.105	-2.94 / 2.96	
MSCI		37	0.026	0.73	0.118	2.57	13.46%
	Spurious regression bias		0.024 / 0.174	0.64 / 5.12	-0.095 / 0.094	-1.98 / 1.96	
	Spurious regression bias and data mining		-0.013 / 0.210	-0.34 / 6.44	-0.139 / 0.137	-2.98 / 2.95	
MSCI - Rf		37	-0.032	-0.91	0.118	2.68	14.65%
	Spurious regression bias		-0.031 / 0.113	-0.87 / 3.34	-0.091 / 0.091	-1.99 / 1.97	
	Spurious regression bias and data mining		-0.064 / 0.145	-1.85 / 4.49	-0.133 / 0.133	-3.00 / 2.96	

Panel B: International Markets

Spain	37	-0.004	-0.06	0.227	2.83	16.27%
Spurious regression bias		-0.012 / 0.272	-0.21 / 4.91	-0.173 / 0.174	-2.00 / 2.00	
Spurious regression bias and data mining		-0.080 / 0.339	-1.34 / 6.43	-0.253 / 0.257	-3.01 / 3.03	
Switzerland	37	-0.0265	-0.50	0.1347	2.00	7.68%
Spurious regression bias		-0.0284 / 0.1539	-0.61 / 3.46	-0.1387 / 0.1375	-2.05 / 2.03	
Spurious regression bias and data mining		-0.0710 / 0.1975	-1.55 / 4.58	-0.2037 / 0.1979	-3.10 / 3.04	
World	37	-0.0449	-0.99	0.0914	1.68	4.83%
Spurious regression bias		-0.0512 / 0.0925	-1.41 / 2.58	-0.1079 / 0.1070	-2.12 / 2.12	
Spurious regression bias and data mining		-0.0852 / 0.1259	-2.39 / 3.64	-0.1574 / 0.1593	-3.23 / 3.22	

This table contains parameter estimates and t-statistics and ranges for these variables within which the null hypothesis of spurious regression bias due to dummy variable persistence or spurious regression and data mining bias cannot be discounted. These are determined based on the Powell, Shi, Smith, and Whaley (2007) approach. All equity series are excess total return indices in US dollars.

Table 5. US Size, Book-to-Market and Industry Results

	Positive Jan Return	N	Negative Jan Return	N	Spread	P-Value OLS	P-Value NW	P-Value BS
Panel A: Size								
Size 1 (Small)	11.56	54	-13.34	13	24.89	0.00	0.00	0.01
Size 2	11.21	49	-3.70	18	14.91	0.02	0.01	0.03
Size 3	10.66	49	0.11	18	10.54	0.06	0.04	0.04
Size 4	11.81	44	0.64	23	11.17	0.03	0.01	0.03
Size 5	12.29	44	-0.37	23	12.66	0.01	0.00	0.01
Size 6	11.47	41	2.14	26	9.33	0.03	0.01	0.03
Size 7	11.70	45	1.74	22	9.96	0.03	0.02	0.01
Size 8	10.74	44	1.90	23	8.84	0.03	0.01	0.02
Size 9	10.97	43	0.93	24	10.04	0.01	0.00	0.01
Size 10 (Large)	12.04	38	-0.90	29	12.94	0.00	0.00	0.00
Panel B: Book-to-Market Equity								
BM 1 (Low BE/ME)	12.64	36	-1.26	31	13.90	0.00	0.00	0.00
BM 2	10.64	42	-0.42	25	11.06	0.00	0.01	0.00
BM 3	10.20	39	0.88	28	9.32	0.01	0.00	0.01
BM 4	12.68	37	-0.16	30	12.83	0.00	0.00	0.00
BM 5	12.09	39	2.97	28	9.13	0.01	0.00	0.02
BM 6	10.66	42	3.47	25	7.18	0.05	0.01	0.02
BM 7	12.29	41	1.15	26	11.14	0.00	0.01	0.00
BM 8	12.31	47	3.05	20	9.25	0.04	0.03	0.02
BM 9	12.14	48	1.54	19	10.60	0.03	0.00	0.03
BM 10 (Hi BE/ME)	10.35	52	6.85	15	3.50	0.60	0.60	0.35
Panel C: Industry								
Consumer NonDur	9.47	42	3.85	25	5.61	0.14	0.09	0.13
Consumer Durables	11.36	43	-1.26	24	12.62	0.02	0.01	0.03
Manufacturing	10.98	38	1.58	29	9.40	0.01	0.01	0.01
Energy	13.46	34	6.17	33	7.29	0.06	0.08	0.04
Technology	9.62	41	2.56	26	7.06	0.19	0.15	0.16
Telecommunications	7.64	45	-0.66	22	8.30	0.04	0.03	0.08
Shops	10.68	37	3.98	30	6.70	0.14	0.13	0.09
Healthcare	14.11	35	4.48	32	9.62	0.03	0.06	0.04
Utilities	9.30	42	0.30	25	9.00	0.02	0.01	0.06
Other	9.82	45	0.27	22	9.55	0.03	0.01	0.03

All equity data are derived from CSRP and sourced from Ken French's website for the 1940-2006 period. Our risk free rate proxy is the 1-month Treasury Bill rates from Global Financial Data. All results are for excess returns. We calculate the 11-month return from February to December following positive (negative) January returns and then calculate the spread, which is the difference in the 11-month returns. The statistical significance of the spread (null hypothesis is spread equals 0) is determined using OLS t-statistics (OLS), Newey West (1987) t-statistics (NW) and bootstrapped t-statistics (BS), which are calculated using the method advocated by Cooper, McConnell, and Ovtchinnikov (2006). All returns are presented as percentages.

Table 6. US Time-Varying Risk Premia Results

	Equally Weighted Index	Value Weighted Index
N	15	24
Mean	-8.54	-2.65
p-value	0.08	0.39
Median	-7.79	-4.46
p-value	0.07	0.40

All equity data are sourced from CSRP for the 1940-2006 period. We include all stocks in the CRSP database. Our risk free rate proxy is the 1-month Treasury Bill rate from Global Financial Data. All results are for excess returns, presented as percentages. We apply the Schwert (2003) test to the January Barometer by investigating ability of negative January months to predict negative excess returns for the remaining 11 months of the year. Results relate to a basic t-test and the non-parametric Wilcoxon test.

Table 7. Example of Disparity Between the Performance of the January Barometer in Individual Stocks versus Indices

	Scenario A JB works for ALL stocks		Scenario B JB works for NO stocks		Scenario C JB works for 50% of stocks	
	Jan Return	Feb - Dec Return	Jan Return	Feb - Dec Return	Jan Return	Feb - Dec Return
Stock 1	4%	25%	4%	-3%	4%	25%
Stock 2	3%	20%	3%	-2%	3%	-2%
Stock 3	-1%	-3%	-1%	25%	-1%	-3%
Stock 4	-2%	-2%	-2%	20%	-2%	20%
Equally Weighted Market Index	1%	10%	1%	10%	1%	10%

This table contains a stylized example of how the January Barometer can perform quite differently at the individual stock level (i.e. be 100% accurate, 50% accurate or 100% inaccurate) and yet still produce the same result at the market index level (i.e. being accurate for any given year).

Table 8. Individual Stock Analysis

	All Years	Years with a Positive January Market Return	Years with a Negative January Market Return
N	67	52	15
Panel A: Mean Analysis			
Mean Positive January Stocks 11-Month Return	7.42	11.64	-7.21
Mean Negative January Stocks 11-Month Return	6.52	10.54	-7.40
Difference	0.90	1.11	0.19
p-value	0.40	0.39	0.49
Panel B: Median Analysis			
Median Positive January Stocks 11-Month Return	-6.98	-1.31	-16.54
Median Negative January Stocks 11-Month Return	-6.24	-1.37	-17.55
Difference	-0.74	0.06	1.01
p-value	0.95	0.98	0.85

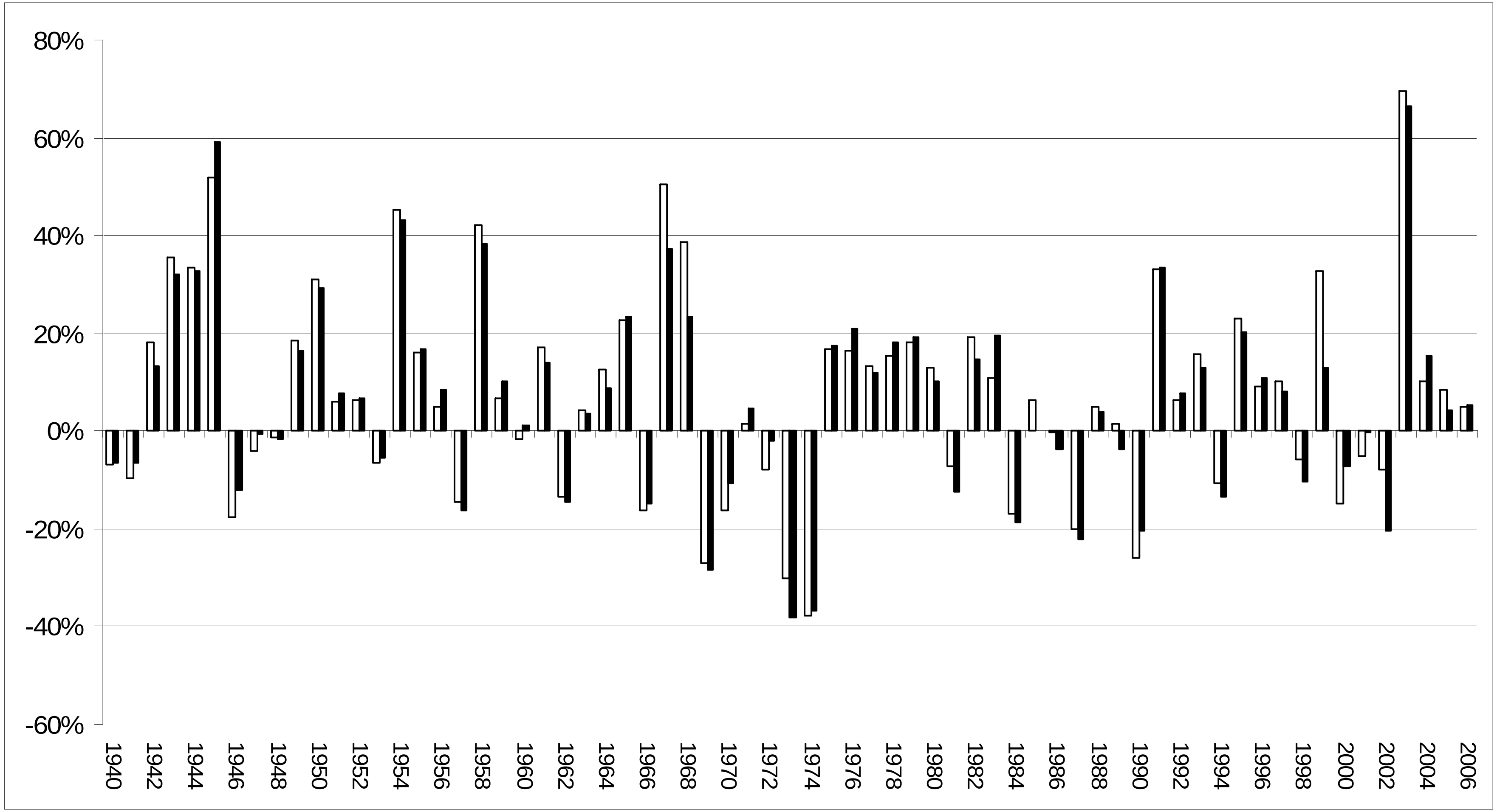
All equity data are sourced from CSRP for the 1940-2006 period. We include all stocks in the CRSP database. Our risk free rate proxy is the 1-month Treasury Bill rate from Global Financial Data. All results are for excess returns, presented as percentages. Results relate to a basic t-test and the non-parametric Wilcoxon test.

Table 9. Commodity Results

	Start	Positive Jan Return	N	Negative Jan Return	N	Spread	P-Value OLS	P-Value NW	P-Value BS
Aluminium	1990	4.22	34	-6.55	26	10.77	0.12	0.06	0.08
Cocoa	1984	-1.17	37	-5.56	22	4.39	0.40	0.41	0.24
Corn	1947	-3.18	17	2.56	27	-5.74	0.51	0.49	0.76
Cotton	1980	-4.95	15	-0.49	18	-4.46	0.29	0.30	0.89
Gold	1970	-11.37	19	-6.52	7	-4.85	0.52	0.49	0.72
Heating Oil	1987	14.82	16	-2.81	21	17.63	0.04	0.04	0.18
Copper	1968	11.31	21	-4.40	18	15.71	0.34	0.24	0.38
Live Cattle	1974	12.34	20	-4.09	19	16.43	0.34	0.26	0.29
Live Hogs	1981	3.65	7	-6.83	16	10.48	0.30	0.30	0.12
Natural Gas	1998	-5.81	6	-1.86	6	-3.95	0.85	0.82	0.43
Nickel	1990	-1.19	12	0.35	15	-1.54	0.91	0.91	0.48
Platinum	1986	-0.94	18	-13.63	9	12.69	0.12	0.16	0.16
Silver	1968	0.92	12	0.31	9	0.61	0.93	0.91	0.57
Soybean Oil	1950	-0.47	10	-1.91	7	1.44	0.90	0.88	0.48
Soybeans	1948	3.17	14	1.98	9	1.19	0.93	0.93	0.41
Sugar	1980	14.76	9	6.01	11	8.75	0.63	0.61	0.45
Coffee	1995	70.25	4	11.02	5	59.23	0.43	0.45	0.23
Unleaded Gas	1987	21.08	11	-7.42	6	28.50	0.27	0.15	0.20
Crude Oil	1984	1.75	30	0.24	27	1.51	0.86	0.86	0.50
Wheat	1963	4.17	11	2.42	9	1.75	0.90	0.91	0.40
Zinc	1990	11.98	10	-6.72	7	18.70	0.17	0.13	0.04

All our commodity series are sourced from Global Financial Data. Our risk free rate proxy is the 1-month Treasury Bill rate from Global Financial Data. All commodity results are for excess returns. We calculate the 11-month return from February to December following positive (negative) January returns and then calculate the spread, which is the difference in the 11-month returns. The statistical significance of the spread (null hypothesis is spread equals 0) is determined using OLS t-statistics (OLS), Newey West (1987) t-statistics (NW) and bootstrapped t-statistics (BS), which are calculated using the method advocated by Cooper, McConnell, and Ovtchinnikov (2006). All returns are presented as percentages.

Figure 1. 11-Month Returns for Stocks with Positive and Negative January Returns



We use all stocks in CRSP for the 1940-2006 period and report excess returns. We graph the mean 11-month returns for stocks that had a positive return in January (light colored bars) and the mean 11-month returns for stocks that had a negative return in January (dark colored bars).

Appendix 1: The Ability of Different Months of the Year to Predict the next 11-Month Returns

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
						Panel A: US						
EW	**19.15**	-1.93	-5.03	3.46	1.34	7.52	2.97	1.95	5.77	-2.79	1.84	0.93
VW	**14.17**	2.64	-3.77	0.98	0.48	5.23	0.28	1.61	1.70	-1.01	0.53	-2.75
MSCI	**11.81**	7.11	-3.38	**12.72**	5.66	0.40	-3.45	-1.20	-0.96	6.02	-1.00	-2.85
						Panel B: International Markets						
Australia	-1.16	6.15	7.45	4.44	-1.82	-3.05	-7.85	7.09	7.25	1.52	-0.92	5.35
Austria	5.89	15.93	13.10	8.46	13.64	13.73	7.21	6.30	**20.93**	12.32	4.67	0.89
Belgium	7.52	**10.84**	-1.76	-1.63	10.12	1.21	4.41	11.22	5.39	-3.59	3.10	-0.81
Canada	-1.63	4.62	-0.50	7.47	-2.18	-1.79	-1.78	7.53	6.96	6.39	7.55	**16.23**
Denmark	12.23	13.75	-2.09	**23.76**	1.08	**13.01**	4.83	**16.88**	**29.19**	9.18	0.45	3.52
Finland	34.21	-1.09	13.13	34.28	7.32	**36.07**	-2.51	-26.16	**61.37**	24.58	14.02	1.65
France	-6.79	3.70	5.55	2.29	10.48	13.39	-5.41	-4.11	2.30	-1.48	5.39	7.16
Germany	0.88	2.92	-6.87	11.37	10.36	11.24	-11.14	3.01	9.57	-5.13	8.86	6.61
Greece	22.70	12.51	-34.30	35.19	-1.59	28.18	14.57	-27.87	16.70	-7.31	10.25	24.09
Hong Kong	0.76	**28.17**	**41.52**	14.81	-8.90	-11.55	-8.95	8.07	2.48	2.69	-9.04	4.51
Ireland	11.12	12.74	-1.18	13.74	2.30	12.33	1.56	5.22	6.08	-2.53	-3.23	1.92
Italy	8.69	**16.74**	-16.48	**21.84**	13.80	**32.67**	7.04	6.08	**24.04**	-8.95	2.84	4.64
Japan	2.01	**16.43**	6.26	-1.04	-5.29	**12.73**	0.53	8.46	4.77	-3.15	6.86	2.75
Netherlands	8.48	**13.92**	-4.68	12.62	1.48	3.13	-1.09	1.72	**13.60**	-3.16	-0.30	9.53
New Zealand	2.83	4.78	7.47	9.66	4.92	3.59	**-19.75**	1.91	-8.65	-0.52	7.65	-9.18
Norway	19.84	18.91	15.63	9.57	1.47	10.72	0.92	**19.41**	0.68	0.94	15.57	4.17
Portugal	15.11	15.23	12.12	6.72	**44.75**	**24.16**	-6.89	5.26	5.21	-6.55	13.18	14.02
Singapore	0.17	18.59	**23.54**	10.87	8.32	10.61	**18.63**	2.93	-0.02	3.96	0.36	1.54
Spain	**17.92**	11.10	0.08	-4.54	9.37	8.28	4.50	-1.16	1.79	**20.80**	**14.36**	12.81
Sweden	-4.67	9.77	-13.76	-3.69	10.50	-7.79	-5.99	-18.26	18.25	**18.05**	**13.79**	-0.54
Switzerland	**13.34**	**14.72**	-3.49	**13.85**	2.52	5.37	-1.22	-9.42	6.26	-1.45	-1.13	7.94
United Kingdom	6.59	2.64	-2.73	3.30	7.89	-0.67	-3.84	-5.08	-8.69	-5.01	-3.94	10.10
World	**9.22**	8.13	-5.25	4.20	-1.06	-0.28	**-13.01**	-2.26	2.57	-3.38	2.56	4.84

EW and VW are the CRSP equally weighted and value weighted indices respectively. MSCI is the US total return index calculated by Morgan Stanley. All our international country indices are also based on MSCI USD total return data. These data are sourced from DataStream. We calculate the 11-month return from the end of each conditioning month. For example, for a May conditioning month, the 11-month return is June – April. As per Table 1, we then calculate the spread in 11-month returns following positive (negative) conditioning month returns. 11 month spreads are presented as percentages. Spreads that are statistically significant at the 10% level are highlighted in bold. All returns are presented as percentages. It is important to note that each Barometer in this table is calculated for the last time in 2005. Given that our data series end in December 2006 it is possible to calculate a January Barometer in 2006 (which we do in our other tables) but not a barometer for other months. We therefore use December 2005 as our last conditioning month.