MUTUAL FUND FLOWS AND SEASONALITIES IN STOCK RETURNS

Moritz Wagner and John Byong-Tek Lee
University of Auckland Business School
Auckland, New Zealand

This Version: February 2015

ABSTRACT

In this paper, we propose a flow-based explanation for a long-standing anomaly in empirical finance – the Sell in May effect. We find that aggregate mutual fund flow exhibits a similar seasonality as stock returns. Given that flow can affect contemporaneous stock returns, the Sell in May effect becomes insignificant in standard statistical tests after controlling for flow. Flow explains about 54% of the variation in excess returns over the winter months. We also find that flow helps explaining the abnormally high returns of small-cap stocks in January.

I Introduction

Numerous seasonalities in stock returns have been documented in previous research. Among the widely cited anomalies are the January or turn-of-the year effect, the turn-of-the-month and the day-of-the week effect. The old saw "Sell in May and go away", known as the Halloween effect, represents probably the most pervasive calendar anomaly. It suggests that stock returns during the winter months should be higher than during the summer months. Bouman and Jacobsen (2002) find higher returns in 36 out of 37 markets in the November-April period than in the May-October period. Jacobsen et al. (2005), Jacobsen and Visaltanachoti (2009), Jacobsen and Zhang (2012) and Andrade et al. (2013) show that this return pattern is also present out of sample, is unrelated to other anomalies, and if anything, has become more pronounced over the recent past. As put forward by Jacobsen and Visaltanachoti, the Sell in May effect has survived all the usual controls and robustness checks up to the present day. It represents a puzzle yet to be explained. The explanations offered in literature such as general investor behaviour and a change in risk aversion due to vacations, Seasonal Affective Disorder (SAD) or temperature can at best partially explain this effect (see Bouman and Jacobsen, 2002; Kamstra et al., 2003, 2009; Cao and Wei, 2005; Jacobsen and Marquering, 2008, 2009; Hong and Yu, 2009). We argue that this empirical pattern is driven by a simple mechanism: mutual fund flows.

The body of literature on mutual fund flows and institutional trading documents that stock returns are contemporaneously correlated with flows into funds (see Chan and Lakonishok, 1993, 1995; Warther, 1995; Edelen and Warner, 2001; Rakowski and Wang, 2009). Coval and Stafford (2007) and Lou (2012) show strong price-pressure effects from flow-induced trading. In addition to the claim that flows cause returns, there are other competing hypotheses to explain the co-movement such as feedback trading, sentiment or simply information revelation. Without favouring one of these hypotheses, given that flow can affect contemporaneous stock returns, it is natural to ask whether flows can also cause seasonalities in stock returns. A visual examination of monthly net flow into US mutual funds reveals a distinct pattern that clearly supports the Sell in May wisdom. On average, realised flow into mutual funds is substantially larger during the winter months compared to the summer months. And in most years over the sample period, stock returns are higher during winter months than during summer months. However, in years with summer flow exceeding winter flow, the Sell in May effect is negative. The usual statistical tests first confirm a Sell in May effect separated from a potential January effect and after controlling for common risk factors. But the seasonal dummy variable drops out when we control for contemporaneous and lagged flow. Average net flow during the winter months in excess over the average flow during the remainder of the year explains about half of the variation of excess returns during the winter months. In addition, we find that flow provides a stronger explanation for the January effect than other explanations discussed in prior research.

The paper is organized as follows. Section II describes the data and methodology. Section III presents the empirical findings and we conclude in section IV.

II Data and Methodology

To estimate the seasonal return pattern, we use CRSP value- and equal-weighted stock market index returns (NYSE + AMEX + NASDAQ), as well as total returns from the S&P 500 index available from Wharton Research Data Services. We obtain monthly net flow for US-based mutual funds that invest in domestic equities and have more than 50 million USD in assets under management from Morningstar. Net flow is estimated from a fund's prior month assets, current month assets and the monthly total return.

$$Flow_{it} = TNA_{it} - TNA_{i,\, t-1}(1+r_{it}) \tag{1}$$

TNA_{it} is a fund's monthly total net asset and r_{it} is the fund's total return. Hence, equation (1) is simply the difference between current and prior month's assets that is not accounted for by monthly total return. The sample period is from January 1995 through December 2014.

We use standard regression analysis to test for seasonalities in stock returns:

$$r_t = \mu + \beta_1 Jan_t + \beta_2 Hal_t + \varepsilon_t \tag{2}$$

where r_t is the return on the stock index for month t, μ is a constant and ε_t is the usual error term. β_1 and β_2 estimate the January and the Halloween effect. Jan_t and Hal_t are seasonal dummy variables that take the value of 1 for January and November to April periods respectively, and 0 otherwise. This equation is equivalent to a simple *t*-test for differences between means. Using this regression however allows us to include other variables, which is vital for the claim of this paper. Table 1 reports summary statistics of the funds in our sample. There is a clear increasing trend visible in the number of funds and the percentage of the stock market held by funds.[1] If flow is truly the underlying force of the Sell in May effect, these trends might provide an explanation why the anomaly has become more rather than less pronounced in recent years.

[1] Percentage of the stock market held by mutual funds is slightly overstated here because we do not consider cash holdings separately. But the figures are mainly comparable to those reported in prior research.

III Results

A. The Sell in May Effect

Panel A of Figure 1 shows average net flow of US mutual funds by month. We first sum net flow over all funds before calculating the monthly average. Hence, Figure 1 presents a market-wide aggregate or the monthly average of one giant fund.

Insert Figure 1 here

Average net flow in the period November-April is substantially larger than in the period May-October, with more money being withdrawn on average than invested during the month September. In Panel B we plot net flow together with average monthly returns. As can be seen, average returns tend to be higher in months with higher average flow and vice versa. This plot provides a graphical depiction of the co-movement between fund flows and returns that has been documented in literature. Returns over the period May-October tend to be rather modest between -1% and 1%. In the same months, average flow is hardly over one billion USD. However, during the period November-April returns are approximately 2% plus and average flow is about three and up to seven billion USD. Corresponding with the January effect being predominantly found among small cap stocks, the equally-weighted market index peaks in January which gives more weight to small firms. During the other months of the year the market proxies are fairly close. January, together with April, is also the month with the highest flow measure.

Panel C plots winter excess stock returns and fund flows over time. The solid line is the cumulative return on the CRSP value-weighted market index during November-April minus the return during May-October. This proxy for the Sell in May effect seems to vary from year to year which mainly supports the findings of Jacobsen and Zhang (2013). However, in 15 out of 20 years returns during the winter months are higher than during the summer months. And the same is true for fund flows. The dashed line shows normalised flow over the winter months in excess of the summer months. In most years, flow during winter is higher than during summer, as indicated by a positive value. Remarkably, in all years where this is not the case, i.e. summer flow is higher than winter flow, the Sell in May effect is also negative. The only exception is 2005. The correlation between the two series is 0.73 (p-value = 0.0002).

Turning now to statistical tests, Table 2 reports estimation results from equation (2). As in previous research we find strong seasonalities in stock returns that are statistically and economically significant. The somewhat hefty turn-of-the-year effect is only present in the equally-weighted index, and thus predominantly among small-cap stocks. The dummy variables in columns 5 and 6 treat the January and Sell in May effect as separate seasonalities, i.e. the Sell in May dummy is 1 in the period November-April, except January and 0 otherwise. The last column contains the results of a regression with only the Sell in May dummy defined as 1 in the period November-April including January. In a nutshell, Table 2 resembles the empirical regularity documented in earlier research we are trying to explain in the following section.

Insert Table 2 here

If mutual fund flows can affect contemporaneous stock returns and given that fund flows exhibit a certain pattern (Figure 1), it is natural to ask whether fund flows can help explaining well-known seasonalities in stock returns. To test this possibility, Panel A of Table 3 report estimation results for different variants of equation (2). More specifically, the regression including all explanatory variables is as follows:

$$r_t = \mu + \beta_1 Hal_t + \beta_2 PE_{t-1} + \beta_3 Flow_t + \beta_4 Flow_{t-1} + \beta_5 Flow_{t-2} + \beta_6 Flow_{t-3} + \varepsilon_t \quad (3)$$

where r_t is the monthly return on the S&P 500 index. The first two variables on the right-hand side are the same as described above. PE_{t-1} is the price-earnings ratio of the market index at the end of the previous month. We also use DY_{t-1} below, which is the dividend yield of the market index at the end of the previous month instead of PE_{t-1}. Both have been found to be helpful predicting stock returns but are highly correlated (-0.83). Hence we cannot include both at the same time to avoid multicollinearity issues. This is our attempt to test variables that are related with stock returns but have so far not been considered for the Sell in May effect. $Flow_t$ is the aggregate mutual fund flow in month t estimated by equation (1) and normalised by the value of the stock market (NYSE + AMEX + NASDAQ) at the end of the previous month. In an attempt to capture both effects associated with the price pressure argument, that flow drives stock prices away from their fundamental values and a corresponding but lagged reversal, we also include lagged flow.[2] We do not re-examine temperature and the Onset/Recovery (aka SAD) variable from Kamstra et al. here. Both have been widely debated in literature as a potential cause for the seasonal anomaly in stock returns driven by mood

[2] In unreported tests, regressions of market returns on flow show that contemporaneous flow is positively and the first lag is negatively related to returns (on the one and five percent significance level respectively). We include up to three lags to capture as much as possible of the flow effects.

changes of investors because of the variation in daylight and temperature. However, the evidence in favour for these two explanations is not convincing (see Jacobsen and Marquering, 2008, 2009). This is partly due to their high correlation with the Halloween indicator which is 0.88 and -0.68 for temperature and SAD respectively, which makes it difficult to test the joint effects. By contrast, the correlation between flow and the Halloween indicator is only 0.18.[3] Column two of Panel A in Table 3 shows the strong and positive relation between stock returns and concurrent flow on the macro level that is known since Warther (1995). The coefficient on *Flow* is 1.67 with a *t*-statistic of 3.71. What is new is that the Sell in May dummy becomes insignificant, i.e. after accounting for flow there is no winter-summer seasonality in stock returns left in a statistical sense.

Insert Table 3 here

Furthermore, column three shows that contemporaneous flow is positively related to stock returns (*t*-statistic = 5.29) and consistent with a reversal of the price pressure effect, lagged flow is negatively related to stock returns (*t*-statistic = -2.43 for $Flow_{t-1}$). The estimate on flow is mainly unaffected when we include the price-earnings ratio or the dividend yield as shown in columns four and five. However, only the former is more than two standard errors away from zero in our tests. The coefficient on PE_{t-1} is -0.17 with a *t*-statistic of -2.71 and the coefficient on DY_{t-1} is 1.50 with a *t*-statistic of 1.62.

Panel B reports results of a regression in which the dependent variable is the six-month return over the period November-April in

[3] The variance inflation factors are 1 or very close to 1 in all regressions that include flow.

excess of the six-month return over May-October. The explanatory variable is the six-month flow during the same winter period in excess of the summer months. This univariate test explains 54% of the variation in the Sell in May effect. The coefficient estimate of 4.21 with a *t*-statistic of 4.89 implies that for an average excess winter flow of 1.39% (normalised) or USD 24 billion (absolute), the six-month excess return is about 5.84%. This estimate is very close to the average difference between November-April and May-October returns reported in Jacobsen and Zhang (2012). Over the past 50 years they find an average difference of 6.25%.

To address the question of causation, the first three columns of Table 4 report results of regressing market returns on expected and unexpected concurrent flow and the Halloween indicator. Expected flow is estimated in a first step using three lags of flow and three lags of returns. Unexpected flow is actual flow minus expected (predicted) flow.[4] The Sell in May dummy becomes only insignificant when we account for the unexpected component of flow. The coefficient on unexpected flow is 3.37 and highly significant with a *t*-statistic of 5.45, while expected flow is not significant in statistical terms. We have hoped to see the opposite as causation is generally accepted if the expected rather the unexpected component of a variable is driving the results. If only the unexpected component seems to matter, doubts are left because both the dependent and independent variable could be affected by an unknown variable causing simple correlation. Hence, more tests are required to clearly separate between the two possibilities.

[4] The results reported in Table 4 are insensitive to variations in the first step, i.e. the number of lags or if we just include flow, e.g. using a simple AR(3) model.

Regressions four and five shed a bit more light on the flow-return relationship by regressing expected and unexpected flow on concurrent and lagged returns. These results highlight why the coefficient on the expected component of flow in the first and third column is statistically insignificant. Expected flow lags return, while concurrent return is only related to unexpected flow. Based on this we can further infer that only the expected component of flow is consistent with the feedback-trader hypothesis which predicts that flows must lag returns.[5]

B. The January Effect

Since flow spikes in January (Figure 1) and January falls into the winter period, we address the obvious question whether flow helps to explain the January effect next. This empirical regularity refers to abnormally high stock returns in January, first documented by Wachtel (1942). Keim (1983), Rozeff and Kinney (1976) and Reinganum (1983) find it to be mainly present among small-cap firms. Schwert (2003) shows that the effect might have become smaller since its discovery, but the effect has not disappeared. And thus, the debate continues to date. Several explanations have been proposed to explain this anomaly, but empirical results are mixed. For example, Rozeff and Kinney (1976), Chang and Pinegar (1988, 1989, 1990), Rogalski and Tinic (1986), Keamer (1994) and Sun and Tong (2010) suggest that the January effect is due to the seasonality in risk or in the compensation for risk. Tinic and West (1984) find the mean-variance trade-off is only present in January. Haugen and

[5] This lag could be anything from picking up the phone or the order of months but there must be a nonzero lag between flows and returns, Warther (1995).

Lakonishok (1987) and Lakonishok et al. (1991) propose a window dressing hypothesis in which institutional investors try to make their portfolios look better by selling stocks with large losses at the end of the year. Branch (1977), Dyl (1977), Reinganum (1983), Jones et al. (1991) and Poterba and Weisenbenner (2001) attribute the effect to tax-loss selling in December and corresponding purchase activities in January. Chen and Singal (2004) demonstrate that tax-loss selling is the main driver behind this anomaly. We do not discuss prior research in more detail here, because the literature is large and surveys can be found elsewhere (e.g. Singal, 2004).

To test our flow-based explanation we begin by repeating the analysis from above but include a January indicator. Since we do not find a January effect in the value-weighted CRSP stock market index, all tests below are based on the equally weighted index. The first column in Table 5 reports an average January effect of 2.73% with a *t*-statistic = 1.95. If we include our flow variables the January indicator is completely subsumed. The coefficient on *Flow* is 4.93 with a *t*-statistic of 6.66. Again, consistent with a lagged reversal of the price pressure effect, lagged flow is negatively related to stock returns. The *t*-statistics on the first, second and third lags of flow are -1.77, -0.76 and -2.63 respectively. The estimates on flow are essentially unaffected when the price-earnings ratio and the dividend yield are included, as shown in columns three and four. Column five shows again that it is the unexpected component of flow that is driving the results with a coefficient of 4.69 and a *t*-statistic of 6.23. The coefficient on expected flow is not statistically significant due to the same reasons as discussed above.

Insert Table 5 here

Next, we provide a more direct test whether flow helps explaining the January regularity alongside other alternatives. For this, we first estimate abnormal return and flow in January with the following regressions:

$$r_t = \mu_t + \beta_1 r_{t-1} + \beta_2 Jan_{1995} + \beta_3 Jan_{1996} + \ldots + \beta_{21} Jan_{2014} + \varepsilon_t \qquad (4)$$

$$Flow_t = \mu_t + \beta_1 Flow_{t-1} + \beta_2 Jan_{1995} + \ldots + \beta_{21} Jan_{2014} + \varepsilon_t \qquad (5)$$

where r_t is the return on the equally-weighted CRSP stock market index in month t. $Flow_t$ is the aggregate net flow of our sample funds standardised by the value of the stock market (NYSE + AMEX + NASDAQ) in the previous month. Lagged values are included to take care of serial correlation. β_2 - β_{21} represent the abnormal return and flow in January estimated for each year over the sample period.

Insert Table 6 here

Table 6 shows the January effect is positive and statistically significant in 12 out of 20 years. With few exceptions (e.g. 2002 and the GFC), we also see that in years with strong and positive (negative) abnormal flow, the January effect is large and positive (negative). Table 7 reports results of a regression in which the dependent variable is the January excess return estimated with equation (4). The only explanatory variable in the first test is the estimated abnormal flow in January. In this univariate test, flow is positively related to the January effect with a coefficient of 2.87 and a *t*-statistic = 4.97. The diagnostics identify two influential points, the year 2001, which has an extreme January return, and 2009 the height of the global financial crisis. Adjusting for these events would increase the variation explained by abnormal flow to over 27%, but due to the relatively short time period we only report unadjusted results. The conclusion about flow is not affected if we take corrective measures. If anything, the estimate on flow becomes more significant. The second model includes proxies for alternative explanations for the January effect suggested in previous research. PTS_{t-1} is the maximum potential tax-loss selling at the end of a year. It is defined as the percentage decrease in stock price from the highest price during the year to mid-December, usually December 15. If there was no trading on this day, we take the price from the previous trading day. This is close enough to the end of the year and allows sufficient time for tax-related selling. This measure mainly follows prior research and enables us to make meaningful comparisons (e.g. see Reinganum, 1983; Chen and Singal, 2004).

$$PTS = \frac{\sum_{i=1}^{n}\left(\frac{price_{it,Dec.}}{price_{it,High}}\right) - 1}{n} \qquad (6)$$

By design, *PTS* might also pick up window dressing activities from institutional investors. Since we are not interested to disentangle the two competing explanations, we rather appreciate that *PTS* sort of captures both possibilities. We include liquidity and volume as a general source for the January seasonality. Abnormally high volume usually occurs with informed trading and as such, is consistent with the information release hypothesis. However, the entry of noise traders may also affect volume. Another reason to include volume here is to avoid the possibility that flow is just volume in disguise. Standard deviation takes care of the risk argument. Both of these measures are estimated in relative terms, i.e. January dollar volume and January standard deviation relative to volume and standard deviation over the previous six months.

The choice of the time period, whether it is the previous six or eleven months, does not affect the results.

Insert Table 7 here

The results reported in Table 7 show that excess flow helps explaining the January effect alongside alternative explanations discussed in literature. The coefficient of *Flow* is 3.17 with a *t*-statistic of 3.55. As is suggested, *PTS* is positively related to the January effect with a coefficient of -16.72, but it is not significant in statistical terms.[6] The estimates on *Vol* and *Std* are also not more than two standard errors away from zero. However, if we control for the GFC the *t*-statistic of *PTS* is -2.29, while the signs for *Vol* and *Std* change and become more in line with the general risk-return trade-off. Yet, both variables remain insignificant while the amount of variation in the January effect explained increases to 30%. Regardless if we control for potential data issues or not the estimate on flow hardly changes, with a persistent *t*-statistic of more than 3.0. Our results indicate, that flow is related to the January anomaly in stock returns. Even if flow may not be the sole driver behind the effect, it certainly is one element of the explanation.

IV Conclusion

Consistent with prior research we find a statistically and economically significant difference between returns during the winter and the summer months. We provide a flow-based explanation for this long-standing anomaly that challenges basic financial theory. Specifically, we show the Sell in May effect is

[6] PTS is between -1 and 0 by construction.

positive (negative) in years where flow during the winter months is higher (lower) than during the summer months. After controlling for mutual fund flows the Sell in May effect becomes insignificant. Excess fund flow explains almost half of the variation in the Sell in May effect. We also find that flow helps explaining the well-known January effect.

Our results build on the contemporaneous relationship between returns and flow. If flow provides an explanation for seasonalities in stock returns an interesting question remains, what drives seasonalities in fund flows?

REFERENCES

Andrade, Sandro C., Chhaochharia, Vidhi, & Fuerst, Michael E. (2013). "Sell in May and Go Away" Just Won't Go Away. *Financial Analysts Journal, 69*(4), 94-105.

Bouman, Sven, & Jacobsen, Ben. (2002). The Halloween Indicator, "Sell in May and Go Away": Another Puzzle. *American Economic Association, 92*(5), 1618-1635.

Cao, Melanie, & Wei, Jason. (2005). Stock Market Returns: A Note on Temperature Anomaly. *Journal of Banking & Finance, 29*(6), 1559-1573.

Chan, Louis, & Lakonishok, Josef. (1993). Institutional Trades and Intraday Stock Price Behavior. *Journal of Financial Economics*, *33*, 173-199.

Chan, Louis, & Lakonishok, Josef. (1995). The Behavior of Stock Prices Around Institutional Trades. *Journal of Finance*, *50*, 1147-1174.

Chang, Eric, & Pinegar, Michael. (1988a). Does the Market Reward Risk in Non-January Months? *Journal of Portfolio Management*, *15*(1), 55-57.

Chang, Eric, & Pinegar, Michael. (1988b). A Fundamental Study of the Seasonal Risk-Return Relationship: A Note. *Journal of Finance, 43*(4), 1035-1039.

Chang, Eric, & Pinegar, Michael. (1989). Seasonal Fluctuations in Industrial Production and Stock Market Seasonal. *Journal of Financial and Quantitative Analysis, 24*, 59-74.

Chang, Eric, & Pinegar, Michael. (1990). Market Seasonal and Prespecified Multifactor Pricing Relations. *Journal of Financial and Quantitative Analysis, 25*, 517-533.

Chen, Honghui, & Singal, Vijay. (2004). All Things Considered, Taxes Drive the January Effect. *Journal of Financial Research, 27*(3), 351-372.

Coval, Joshua, & Stafford, Erik. (2007). Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics, 86*, 479-512.

Dyl, Edward A. (1977). Capital Gains Taxation and Year-end Stock Market Behaviour. *Journal of Finance, 32*(1), 165-175.

Edelen, Roger M., & Warner, Jerold B. (2001). Aggregate Price Effects of Institutional Trading: A Study of Mutual Fund Flow and Market Returns. *Journal of Financial Economics, 59*, 195-220.

Fama, Eugene F., & French, Kenneth R. (1993). Common Risk Factors in the Return on Bonds and Stocks. *Journal of Financial Economics, 33*, 3-53.

Haugen, Robert A., & Lakonishok, Josef. (1987). *The Incredible January Effect: The Stock Market's Unsolved Mystery*: Dow Jones-Irwin.

Hong, Harrison, & Yu, Jialin. (2009). Gone fishin': Seasonality in Trading Activity and Asset Prices. *Journal of Financial Markets, 12*(4), 672-702.

Jacobsen, Ben, Mamun, Abdullah, & Visaltanachoti, Nuttawat. (2005). Seasonal, Size and Value Anomalies. *Working Paper, Massey University, Auckland New Zealand*.

Jacobsen, Ben, & Marquering, Wessel. (2008). Is it the Weather? *Journal of Banking & Finance, 32*(4), 526-540.

Jacobsen, Ben, & Marquering, Wessel. (2009). Is it the Weather? Response. *Journal of Banking & Finance, 33*(3), 583-587.

Jacobsen, Ben, & Visaltanachoti, Nuttawat. (2009). The Halloween Effect in U.S. Sectors. *The Financial Review, 44*, 437-459.

Jacobsen, Ben, & Zhang, Cherry Y. (2012). The Halloween Indicator: Everywhere and all the Time. *Working Paper, Massey University, Auckland New Zealand*.

Jones, Steven L., Lee, Winson, & Apenbrink, Rudolf. (1991). New Evidence on the January Effect before Personal Income Taxes. *Journal of Finance, 46*(5), 1909-1924.

Kamstra, Mark J., Kramer, Lisa A., & Levi, Maurice D. (2003). Winter Blues: A SAD Stock Market Cycle. *American Economic Review, 93*(1), 324-343.

Kamstra, Mark J., Kramer, Lisa A., & Levi, Maurice D. (2009). Is it the Weather? Comment. *Journal of Banking & Finance, 33*, 578-582.

Keim, Donal B. (1983). Size-related Anomalies and Stock Return Seasonality: Further Empirical Evidence. *Journal of Financial Economics, 12*(1), 13-32.

Lakonishok, Josef, Shleifer, Andrei, Thaler, Richard, & Vishny, Robert. (1991). Window Dressing by Pension Fund Managers. *American Economic Review, 81*, 227-231.

Lou, Dong. (2012). A Flow-Based Explanation for Return Predictability. *The Review of Financial Studies, 25*, 3457-3489.

Poterba, James M., & Weisenbenner, Scott J. (2001). Capital Gains Tax Rules, Tax-loss Trading, and the Turn-of-the-year Returns. *Journal of Finance, 56*(1), 353-368.

Rakowski, David, & Wang, Xiaoxin. (2009). The Dynamics of Short-term Mutual Fund Flows and Returns: A Time-series and Cross-sectional Investigation. *Journal of Banking & Finance, 33*, 2102-2109.

Reinganum, Marc R. (1983). The Anomalous Stock Market Behaviour of Small Firms in January. *Journal of Financial Economics*, *12*, 89-104.

Rogalski, Richard, & Tinic, Seha M. (1986). The January Size Effect: Anomaly or Risk Mismeasurement? *Financial Analysts Journal, 42*(6), 63-70.

Rozeff, Michael S., & Kinney, William R. (1976). Capital Market Seasonality: The Case of Stock Returns. *Journal of Financial Economics, 3*(4), 379-402.

Schwert, G. William, Harris, Milton, & Stulz, Rene M. (2003). Anomalies and Market Efficiency. In George Constantinides (Ed.), *Handbook of the Economics of Finance* (pp. 937-972): Elsevier North-Holland.

Singal, Vijay. (2004). *Beyond the Random Walk: A Guide to Stock Market Anomalies and Low-Risk Investing*. New York: Oxford University Press, Inc.

Sun, Qian, & Tong, Wilson H.S. (2010). Risk and the January Effect. *Journal of Banking & Finance, 34*(5), 965-974.

Tinic, Seha M., & West, Richard R. (1984). Risk and Return: January vs. the Rest of the Year. *Journal of Financial Economics*, *13*(4), 561-574.

Wachtel, Sidney. (1942). Certain Observations on Seasonal Movement in Stock Prices. *Journal of Business, 15*, 184-193.

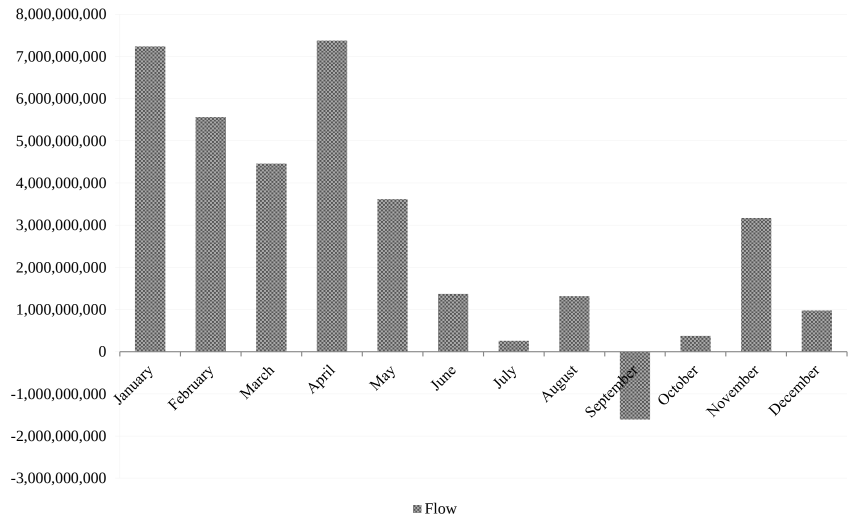
Warther, Vincent A. (1995). Aggregate Mutual Fund Flows and Security Returns. *Journal of Financial Economics, 39*, 209-235.

Zhang, Cherry Y., & Jacobsen, Ben. (2013). Are Monthly Seasonals Real? A Three Century Perspective. *Review of Finance, 17*, 1743-1785.

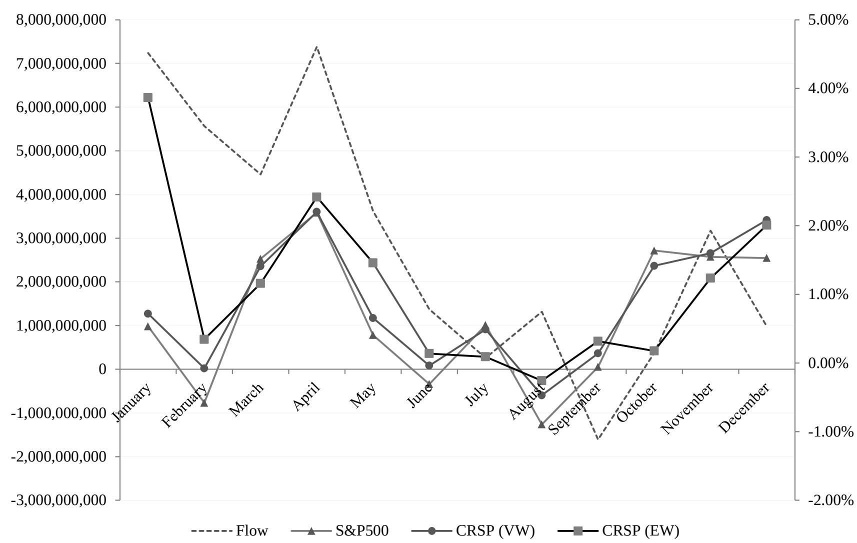
Figure 1
Mutual Fund Flows and Stock Returns

Panel A of this figure reports the average monthly flow of our sample funds (market-wide aggregate) by months. Panel B plots the same flow measure together with average monthly returns on the CRSP value- and equally-weighted stock market indices (NYSE + AMEX + NASDAQ) and the S&P 500 index. Panel C reports six-month returns on the CRSP value-weighted stock market index of the period November-April in excess over May-October and the same for mutual fund flows, normalised by the value of the market (NYSE + AMEX + NASDAQ) and scaled by 1000. The sample period is January 1995 to December 2014.

Panel A – Average monthly flow



Panel B – Average monthly flow and average monthly returns



Panel C – Winter excess fund flows and market returns

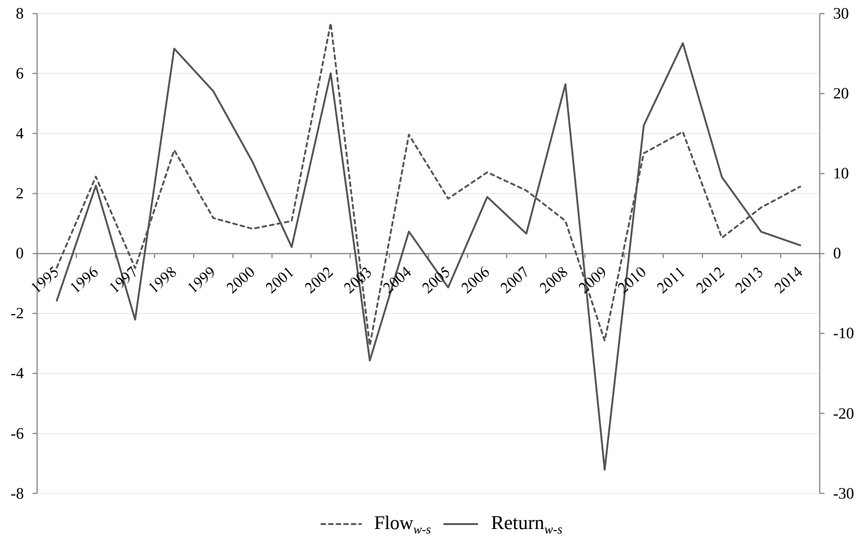


Table 1
Summary Statistics of US Equity Mutual Funds

This table reports summary statistics of all US-based mutual funds that invest in domestic equities as of the end of December in each year. The only filter we apply is that they have more than 50 million dollars in assets under management based on the most recent portfolio date. The number of funds is given in share classes. We calculate percent of market value as total net assets divided by total value of the stock market (NYSE + AMEX + NASDAQ). The sample period is from 1995 to 2014.

Year	Number of Funds	Total Net Assets ($M)	% of Market Value
1995	908	778,391	11.47%
1996	1,071	1,058,615	12.75%
1997	1,268	1,473,680	13.65%
1998	1,425	1,842,371	13.86%
1999	1,640	2,452,570	14.41%
2000	1,938	2,375,616	15.20%
2001	2,240	2,184,445	15.78%
2002	2,414	1,634,410	14.82%
2003	2,737	2,450,488	16.81%
2004	2,989	2,953,683	17.95%
2005	3,284	3,226,563	18.57%
2006	3,505	3,745,166	19.10%
2007	3,728	4,095,311	20.28%
2008	3,998	2,447,566	20.18%
2009	4,160	3,237,890	20.49%
2010	4,325	3,751,258	20.29%
2011	4,434	3,593,521	20.09%
2012	4,498	4,013,517	19.72%
2013	4,495	5,378,888	20.47%
2014	4,364	5,739,679	19.82%

Table 2
The Sell in May and January Effect

This table reports estimation results of the Sell in May or Halloween effect and the January effect. The first two rows are the value- and equally-weighted indices of all stocks listed on the NYSE, AMEX and NASDAQ. The S&P 500 index represents the 500 largest publicly traded corporations in the US. The January dummy is 1 for returns that fall into January and 0 otherwise. The Sell in May dummy (not) adjusted for January is 1 for the period November-April (including) excluding January and 0 otherwise. Mean and adjusted R^2 are reported from the regression including the January and the adjusted Sell in May dummy. The sample period is January 1995 to December 2014. *t*-statistics are reported in parentheses based on Newey-West corrected standard errors.

Market Index	Adj-R^2	Obs.	Mean	January Dummy	Sell in May Dummy (adjusted for January)	Sell in May Dummy (not adjusted for January)
CRSP (VW)	0.01	240	0.37	0.04	**1.22**	**1.03**
			(0.82)	(0.04)	**(2.19)**	**(1.88)**
CRSP (EW)	0.02	240	0.31	**3.30**	**1.26**	**1.60**
			(0.54)	**(2.38)**	**(1.76)**	**(2.36)**
S&P 500	0.01	240	0.24	-0.05	**1.17**	**0.96**
			(0.57)	(-0.05)	**(2.23)**	**(1.85)**

Table 3
Mutual Fund Flows and the Halloween Seasonality in Stock Returns

Panel A reports estimation results for different variants of equation (2) with *t*-statistics in parentheses based on Newey-West corrected standard errors. The dependent variable is the monthly return on the S&P 500 index. The Sell in May dummy, Hal_t, is 1 for the period November-April and 0 otherwise. PE_{t-1} and DY_{t-1} are the price-earnings ratio and the dividend yield of the market index at the end of the previous month. $Flow_t$ is the estimated monthly net flow (market-wide) of our sample funds normalised by the value of the market (NYSE + AMEX + NASDAQ) and scaled by 1000. The dependent variable in Panel B is the six-month return over the period November-April minus the six-month return May-October, a proxy for the Sell in May effect. The explanatory variable is the half-year flow over the winter months November-April minus the flow during the remainder of the year ($Flow_{w-s}$), normalised by the value of the market at the previous month and scaled by 1000. The sample period is January 1995 to December 2014.

Panel A

	(1)	(2)	(3)	(4)	(5)
Obs.	240	240	240	240	240
Adj-R^2	0.01	0.08	0.13	0.16	0.15
Intercept	0.24	0.00	0.41	**3.69**	-2.45
	(0.57)	(0.01)	(0.99)	**(2.77)**	(-1.44)
Hal	**0.96**	0.50	0.14	0.14	0.12
	(1.85)	(1.03)	(0.30)	(0.30)	(0.27)
PE_{t-1}				**-0.17**	
				(-2.71)	
DY_{t-1}					1.50
					(1.62)
Flow		**1.67**	**3.30**	**3.37**	**3.31**
		(3.71)	**(5.29)**	**(5.50)**	**(5.38)**
$Flow_{t-1}$			**-1.59**	**-1.41**	**-1.46**
			(-2.43)	**(-2.24)**	**(-2.33)**
$Flow_{t-2}$			-0.06	0.09	0.04
			(-0.07)	(0.12)	(0.06)
$Flow_{t-3}$			-0.73	-0.53	-0.58
			(-1.20)	(-0.88)	(-0.96)

Panel B

	Dependent Variable: Market $Return_{w-s}$			
	Obs.	R^2	Intercept	$Flow_{w-s}$
Coef.	20	0.54	-1.02	**4.21**
(*t* -stat.)			-0.46	**(4.89)**

Table 4

Expected and Unexpected Mutual Fund Flows and Stock Returns

The first three columns of this table reports estimation results from regressing market returns on expected and unexpected fund flow and the Halloween indicator, *Hal*. This is based on a two-step estimation procedure where expected and unexpected flow are generated from estimates of a first-step regression. In the first step we regress flow on three lags of flow and three lags of returns. Expected flow is the fitted value, while unexpected flow is the residual. In columns four and five we regress expected and unexpected flow on lagged market returns. The sample period is January 1995 to December 2014. *t*-statistics are reported in parentheses based on Newey-West corrected standard errors.

	Dependent Variable				
	Market Return			Expected Flow	Unexpected Flow
	(1)	(2)	(3)	(4)	(5)
Obs.	240	240	240	240	240
Adj-R^2	0.01	0.14	0.14	0.17	0.13
Intercept	0.13	**0.65**	0.53	**0.22**	-0.03
	(0.25)	**(1.79)**	(1.20)	**(3.39)**	(-0.90)
Hal	**0.95**	0.15	0.14		
	(1.90)	(0.33)	(0.29)		
Expected Flow	(0.44)		(0.46)		
	(0.73)		(0.77)		
Unexpected Flow		**3.37**	**3.38**		
		(5.45)	**(5.42)**		
Return				0.00	**0.04**
				(0.32)	**(4.99)**
$Return_{t-1}$				**0.05**	0.00
				(7.34)	(-0.50)
$Return_{t-2}$				**0.02**	0.00
				(3.10)	(0.23)
$Return_{t-3}$				**0.02**	0.00
				(2.77)	(-0.62)

Table 5
Mutual Fund Flows and the January Effect

This table reports estimation results of abnormal returns in January. The dependent variable is the monthly return on the EW CRSP stock market index (NYSE + AMEX + NASDAQ). The January dummy, Jan_t, is 1 for returns that fall into January and 0 otherwise. $Flow_t$ is the estimated monthly net flow (market-wide) of our sample funds. SMB_t and HML_t are Fama and French's (1993) firm size (small minus big) and value (high minus low book-to-market ratio) factors. MOM_t is Carhart's (1997) momentum (winner minus loser) factor. Standard errors are corrected for heteroskedasticity and autocorrelation. *t*-statistics are reported in parentheses. The sample period is January 1995 to December 2014.

	(1)	(2)	(3)	(4)	(5)
Obs.	240	240	240	240	240
Adj-R^2	0.02	0.20	0.21	0.20	0.18
Intercept	**0.89**	0.74	2.86	**-1.86**	0.79
	(2.02)	(1.63)	(1.52)	**(-0.82)**	(1.56)
Jan	**2.73**	0.42	0.47	0.44	0.74
	(1.95)	(0.29)	(0.33)	(0.30)	(0.53)
Flow		**4.93**	**4.97**	**4.94**	
		(6.66)	**(6.76)**	**(6.70)**	
$Flow_{t-1}$		**-1.52**	**-1.40**	**-1.40**	
		(-1.77)	**(-1.72)**	**(-1.77)**	
$Flow_{t-2}$		-0.58	-0.49	-0.49	
		(-0.76)	(-0.63)	(-0.63)	
$Flow_{t-3}$		**-1.56**	**-1.42**	**-1.42**	
		(-2.63)	**(-2.26)**	**(-2.20)**	
PE_{t-1}			**-0.11**		
			(-1.70)		
DY_{t-1}				1.35	
				(1.08)	
Expected Flow					0.92
					(1.23)
Unexpected Flow					**4.69**
					(6.23)

Table 6
Abnormal Return and Flow in January

Column two of the table reports abnormal return on the EW CRSP stock market index in January estimated with equation (4). Abnormal flow of the sample funds in January based on equation (5) is reported in column four. Standard errors are corrected for heteroskedasticity and *t*-statistics are reported in parentheses. The sample period is January 1995 to December 2014.

	January Excess Return		January Excess Flow	
	Coef.	*(t-stat.)*	Coef.	*(t-stat.)*
Obs.	240		240	
Adj-R^2	0.07		0.54	
1995	**2.41**	**5.28**	**0.59**	**19.14**
1996	**2.61**	**7.21**	**1.16**	**20.73**
1997	**5.78**	**14.40**	**1.24**	**40.56**
1998	**1.39**	**2.65**	**0.36**	**8.27**
1999	**5.41**	**15.43**	**0.61**	**18.56**
2000	**2.56**	**3.66**	**0.23**	**7.38**
2001	**21.99**	**52.10**	**0.50**	**16.09**
2002	-0.10	-0.19	**0.65**	**21.06**
2003	0.81	1.26	0.00	-0.02
2004	**5.07**	**12.78**	**1.17**	**26.23**
2005	**-4.74**	**-10.07**	**0.10**	**3.13**
2006	**6.68**	**18.88**	**-0.09**	**-2.90**
2007	**1.24**	**3.53**	**0.31**	**8.71**
2008	**-4.93**	**-11.13**	**-1.48**	**-34.51**
2009	**-3.98**	**-9.67**	**0.76**	**12.26**
2010	**-2.99**	**-6.19**	**0.48**	**5.83**
2011	-0.71	-1.23	**0.96**	**12.24**
2012	**7.95**	**19.61**	**0.47**	**4.89**
2013	**5.21**	**14.81**	**1.26**	**13.42**
2014	**-1.34**	**-3.76**	**0.08**	**1.73**

Table 7

Relationship between *$Flow_t$*, *PTS_t*, *Vol_t*, *Std_t* and the January Effect

This table reports results of regressions in which the dependent variable is the January effect estimated with equation (4). $Flow_t$ is the abnormal flow in January estimated with equation (5). PTS_t is the equally weighted year's end potential tax-loss selling over all stocks listed on the NYSE, AMEX and NASDAQ. It is defined as the percentage decrease from the highest price attained during a year to December 15. If there was no trading on December 15, we take the price of the previous day. Vol_t is the natural logarithm of dollar volume in January relative to the average monthly dollar volume over the previous six months (July – December). Monthly volume is calculated from the numbers of shares traded on day *t* times closing price of day *t* of each stock listed on the NYSE, AMEX and NASDAQ. Data on prices and number of shares are obtained from CRSP via WRDS. Std_t is the natural logarithm of the standard deviation of the EW CRSP stock market index in January relative to the standard deviation of the index over the previous six months (July – December). Standard deviation is calculated from daily returns. Standard errors are corrected for heteroskedasticity and autocorrelation. *t*-statistics are reported in parentheses. The sample period is 1995 to 2014.

	Flow	PTS	Vol	Std	R^2
Coef.	**2.87**				0.09
(t-stat.)	**(4.97)**				
Coef.	**3.17**	-16.72	4.48	-2.35	0.20
(t-stat.)	**(3.55)**	(-0.93)	(1.28)	(-0.63)	