DEPARTMENT OF ECONOMICS AND FINANCE

SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Mutual Fund Flows and Seasonalities in Stock Returns

Moritz Wagner
John Byong-Tek Lee
Dimitris Margaritis

WORKING PAPER

No. 17/2018

Department of Economics and Finance
School of Business
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 17/2018

Mutual Fund Flows and Seasonalities in Stock Returns

Moritz Wagner[1†]
John Byong-Tek Lee[2]
Dimitris Margaritis[2]

October 2018

Abstract: We propose a flow-based explanation for two long-standing anomalies in empirical finance – the Sell in May effect and the January effect. We find that the aggregate mutual fund flows exhibit similar seasonal patterns as stock returns. The Sell in May effect becomes insignificant in standard statistical tests after controlling for the impact of mutual fund flows on returns, with flow explaining 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. The Sell in May and January effects appear to be primarily a retail money effect. Similarly, the well-known co-movement between flow and market return is only present in retail fund flow. Overall, the evidence suggests that unanticipated rather than expected flow drives our results.

Keywords: Mutual funds, Fund flows, Return seasonality

JEL Classifications: G12, G14, G23

[1] Department of Economics and Finance, University of Canterbury, NEW ZEALAND

[2] University of Auckland Business School, University of Auckland, NEW ZEALAND

[†] Corresponding author is Moritz Wagner. His email address is moritz.wagner@canterbury.ac.nz.

1. Introduction

Past research has documented various seasonal patterns in stock returns. Among the widely cited calendar month based anomalies at the stock market level are the Halloween effect dubbed as “Sell in May and go away” and the January or turn-of-the year effect. Bouman and Jacobsen (2002) document that stock market returns in November-April are greater than in May-October in 36 out of 37 countries in their sample. Jacobsen et al. (2005), Jacobsen and Visaltanachoti (2009), Jacobsen and Zhang (2012) and Andrade et al. (2013) show that this return seasonality is also present out of sample, is unrelated to other anomalies, and is more pronounced in recent years. A host of behavioural explanations offered in the literature, such as a change in risk aversion due to vacations, a change in investor sentiment due to temperature, and Seasonal Affective Disorder (SAD), can at best only partially explain the puzzle (Bouman and Jacobsen, 2002; Kamstra et al., 2003, 2009; Cao and Wei, 2005; Jacobsen and Marquering, 2008, 2009; Hong and Yu, 2009). We provide a potential explanation that this empirical pattern is driven by a simple mechanism – a similar pattern in mutual fund flows mainly associated with retail investor behaviour.

There is a large body of literature on mutual fund flows and institutional trading documenting that stock returns are contemporaneously correlated with capital flows into funds (Chan and Lakonishok, 1993, 1995; Warther, 1995; Edelen and Warner, 2001; Rakowski and Wang, 2009). Coval and Stafford (2007) and Lou (2012) further show strong price-pressure effects from flow-induced trading. There are also other competing hypotheses explaining the co-movement between flows and returns, such as feedback trading, sentiment, and simply information revelation. Given that mutual fund flow can affect stock returns, we examine whether flows exhibit similar seasonal patterns based on calendar months and whether the return seasonality is independent of observed seasonal patterns exhibited by fund flows.

[INSERT FIGURE 1 AROUND HERE]

Figure 1 illustrates our main findings. Panel A shows that monthly average aggregate net flows into US mutual funds from November to April are substantially larger than those from May to October. Average monthly returns on broad market indices clearly support the Sell in May wisdom documented in the literature and plotted in Panel B; average returns are higher during winter months than during summer months. Consistent with the January effect being predominantly found among small cap stocks, the equally-weighted market index peaks in January. During the other months of the year, the market indices are fairly close to each other. January, together with April, are also the months with the largest flows. The average calendar month returns and flows when plotted together in Panel B show striking resemblance.

Panel C plots the time-series of yearly winter excess returns on the CRSP value-weighted index and winter excess fund flows over their summer counterparts. There is some variation in the Sell in May effect consistent with Zhang and Jacobsen (2013). However, in 15 out of 20 years, returns during the winter months are higher than in the summer months. Again, a very similar pattern emerges for mutual fund flows; in most years, fund flows during winter are higher than during summer. Remarkably, when summer flows are higher than winter flows, the Sell in May effects are also negative with a single exception in 2005. The correlation between the two series is 0.73 (p-value = 0.0002).

Our results show that after accounting for the effect of large fund flows during winter months, the Halloween effect is no longer statistically significant, while average winter net flow in excess over average flow during the remainder of the year explains about half of the variation of excess returns. Since we find very low correlation between fund flow and the Halloween indicator, we surmise that flow provides a stronger explanation of the seasonality in market returns. In addition, we find that mutual fund flow provides a stronger explanation for the January effect than other explanations proposed in previous research. Further analysis reveals that it is the unexpected component of flow driving the results, while expected flow lags return. Results from a VAR model show that flow contains information on returns and vice versa.

We also find that the calendar month pattern is present on capital flows across different investment styles. Distinguishing between flows into retail and institutional funds, however, reveals that both return anomalies are largely driven by retail fund flows. Institutional fund flows do not exhibit a seasonal pattern and cannot help explain either the Sell in May or January effects. Our analysis also shows that the well-known co-movement of aggregate flow and market return is only present in retail fund flow. This is an important finding since it offers a more plausible explanation for the persistence of seasonal market anomalies, namely being the combined effect of (1) unanticipated fund flows, and in particular (2) those driven by overly exuberant or pessimist retail investors.

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.

2. Data and Methodology

We estimate seasonal patterns in returns using CRSP value- and equal-weighted stock market index returns (NYSE + AMEX + NASDAQ), as well as total returns from the S&P 500 index. We obtain monthly net asset values and returns of 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 the fund's net asset values (NAV) in the previous month, current month assets, and the monthly total return:

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

Where TNA_{it} is fund i's total net asset in month t and r_{it} is the fund's total return in that month. In other words, we proxy net flows by simply taking the difference between current and prior month's net asset values that is not accounted for by monthly total return. The sample period is from January 1995 through December 2014.

[INSERT TABLE 1 AROUND HERE]

Table 1 reports summary statistics of the NAVs of the funds in our sample. There is a clear increasing trend for the number of funds and the percentage of the total market capitalization managed by funds.[1] To the extent that price pressure from mutual funds affects stock returns, these trends might be an explanation as to 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. However, the figures are mainly comparable to those reported in prior research.

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

$$r_t = \mu + \beta_1 Season_t + \beta_2 Flow_t + \gamma' Control_t + \varepsilon_t \qquad (2)$$

where r_t is the return on the stock index for month t, μ is a constant and ε_t is the error term. $Season_t$ is a seasonal dummy that assumes the value of one for January and for November to April, to test the January and Sell in May effects, respectively, and zero otherwise, $Flow_t$ is mutual fund flow scaled by the total market capitalization, and $Control_t$ is a vector of other control variables.

3. Results

Table 2 reports the results of statistical tests of the Halloween effect and January effect using (2). As in previous studies, we find strong calendar month 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 it appears to be more pronounced among small-cap stocks. With or without separating January from the rest of winter months (Columns 5 and 6), the Halloween effect is statistically significant. In a nutshell, Table 2 confirms the empirical regularities documented in earlier research.

[INSERT TABLE 2 AROUND HERE]

3.1 The Sell in May Effect

If mutual fund flows affect contemporaneous stock returns, it is natural to ask whether fund flows can help explain the seasonal patterns in stock returns given that we find similar patterns in fund flows and returns as shown in Figure 1. Table 3 reports regression results with fund flows and other control variables added to the set of explanatory variables of (2) – viz., lagged mutual fund flows, P/E ratio, and dividend yield (DY) of the market index at the end of the previous month and the updated investor sentiment index (Sent) from Baker and Wurgler (2006, 2007). Lagged fund flows are included to account for the possibility that fund flows may drive stock prices away from their fundamental values and lead to a reversal as flow-induced price pressure dissipates. The P/E ratio and dividend yield have been found to be positively related to stock returns. We include sentiment mainly for two reasons, (1) to address concerns that investor flows simply represent market sentiment; and (2) to avoid any confounding effects from mutual funds employing strategies that are based on sentiment as shown in Massa and Yadav (2015).

[INSERT TABLE 3 AROUND HERE]

Table 3 shows that with the inclusion of aggregate fund flow, the Halloween dummy is no longer significant (Column 2) suggesting the winter-summer seasonality in stock returns is superseded by fund flow patterns. Contemporaneous inflow is positively and significantly related to stock returns with or without other control variables. Consistent with a reversal of the price pressure effect, lagged flow is negatively related to stock returns (*t*-statistic = -2.43). Similarly, and in line with much evidence in the literature, sentiment is also negatively related to market returns. To explore further the role of sentiment, we estimate fund sentiment betas (FSB) based on a rolling regression of fund excess returns from the previous 36 months on the standard Fama-French-Carhart 4-factors and investor sentiment.[2] The coefficient on the sentiment index indicates whether a particular fund follows a sentiment contrarian (low FSB) or a sentiment catering (high FSB) strategy.[3] Each month, we sort the funds into quintiles based on their FSB and compute aggregate flows for each quintile and month. As before, the Halloween dummy becomes insignificant regardless of which sentiment flow quintile is included in the regression.[4] Overall, our results show that it is fund flow that matters as opposed to sentiment or sentiment-based investing.[5]

Panel B of Table 3 reports the results from a yearly regression in which the dependent variable is the six-month return from November to April in excess of the six-month return from May to October. The explanatory variable is the corresponding six-month flow during the winter period

[2] The results are not reported, but are available from the authors upon request.
[3] Funds in quantile-1 have negative FSBs and are likely to follow a contrarian strategy, whereas funds in quantile-5 have positive FSBs and are likely to follow a sentiment catering strategy. For a discussion of this method, see Massa and Yadav (2015).
[4] Estimating the joint effect of all quintiles in the regressions is difficult due to multicollinearity issues.
[5] We are not suggesting that sentiment does not matter. Similar to Massa and Yadav (2015), we find that investors do not appear to react to FSB directly, i.e. FSB does not directly affect flows, rather works indirectly through fund performance, via fund managers optimally exploiting sentiment.

in excess of the remainder of the year. This simple model explains 54% of the variation in the Sell in May effect. The coefficient estimate of 4.21 (t-statistic = 4.89) implies that for an average excess winter flow of 1.39% of the total market capitalization, the six-month winter excess return is about 5.84%. This estimate is close to the average difference between November-April and May-October returns reported in Jacobsen and Zhang (2012) documenting an average difference of 6.25% over the past 50 years.

We do not include the temperature and the Onset/Recovery (SAD) variables from Kamstra et al. (2003). Both have been widely debated in the literature as potential causes for the seasonal anomaly in stock returns, driven by investor mood changes due to seasonal daylight and temperature changes. However, the evidence in favour of these two explanations is not convincing (see Jacobsen and Marquering, 2008, 2009). This is partly due to their high correlation with the Halloween indicator (0.88 and -0.68) for temperature and SAD, respectively, which makes it difficult to test the joint effects. In contrast, the correlation between flow and the Halloween indicator is only 0.18.[6]

Given that (1) fund flow can explain the seasonal pattern in stock returns and (2) fund flows are predictable, an obvious question is whether expected flow can forecast stock returns. Hence, we proceed decomposing fund flows into expected and unexpected parts. This is important since we would not normally expect these stylistic seasonal return predictability patterns to persist over long periods. More specifically, we estimate expected flow using a seasonal ARIMA model. Based on the ACF and PACF of the differenced flow series, we identify and fit an ARIMA(1,0,1)(0,1,1)$_{12}$ + c model. This model results in the lowest information criterion (AIC) and lowest forecast error (RMSE) compared to a host of competing models, including exponential smoothing, random walk and trend models as reported in Panel A of Table 4. The model specification is:

$$\widehat{Flow}_t = \mu + Flow_{t-12} + \phi_1(Flow_{t-1} - Flow_{t-13}) - \theta_1 e_{t-1} - \Theta_1 e_{t-12} + \theta_1 \Theta_1 e_{t-13} \quad , \qquad (3)$$

where *Flow* is the normalized fund flow, μ is a constant, ϕ denotes an AR(1) coefficient, θ is a non-seasonal MA(1) and Θ a seasonal, SMA(1), coefficient.

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

[INSERT TABLE 4 AROUND HERE]

[INSERT FIGURE 2 AROUND HERE]

Figure 2 plots the time series of normalized flow and its forecast. Besides the identified monthly seasonal, it appears that there is also a long run downward trend present.

Columns 1-3 of Panel B in Table 4 report regression results using expected and unexpected flow to explain market returns. The Sell in May dummy becomes insignificant only when we include the unexpected component of flow. The coefficient on unexpected flow is positive and highly significant (*t*-statistic = 5.56), while expected flow is not statistically significant. The results indicate that the return-flow relation is a contemporaneous one and thus the relation may stem from an unknown common factor.

Columns 4 and 5 shed more light on the relationship by regressing expected and unexpected flow, respectively, 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 is affected by lagged return, while concurrent return is only related to unexpected flow. Based on the findings, 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.[7]

3.2 Dynamics between Flow and Market Returns and the Sell in May Effect

As discussed earlier, there are competing explanations for the co-movement of fund flow and market return. To investigate the co-movement between flow and market returns in more detail, we test the relation within a vector autoregressive framework that also allows us to address the question of causality using the concept of Granger (1969). Furthermore, we partially address potential endogeneity issues as VAR specifications allows for both variables to affect each other. Coval and Stafford (2007) and Lou (2012) provide strong evidence for the price pressure hypothesis. But causality could go the other way in that investors buy (sell) fund shares in response to good (poor) stock market performance. Edelen and Warner (2001) find some evidence in favour

[7] 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).

of this feedback hypothesis. Their findings are also consistent with a third explanation, the information-response hypothesis, which suggests that both returns and flows react to new information. Jank (2012) documents supporting evidence for the information-response hypothesis.

Using a bivariate VAR specification allows us to test the feedback hypothesis in the equation with flow as the dependent variable, while the market return equation can be used to test the price-pressure hypothesis. By including, deterministic variables, such as the Sell in May indicator (HAL) and other exogenous variables that proxy for news, we are able to test for seasonalities and the information-response hypothesis. For the latter, we follow Jank (2012) and include predictive variables as proxies for macroeconomic news. Specifically, we include changes in the dividend yield (ΔDY), default spread (ΔDefault), term spread (ΔTerm) and the three months T-bill (ΔT-bl) rate. These variables have been found to be related to the present state of the economy but also to predict economic activity and the equity premium (e.g. Shiller et al., 1984; Campbell, 1987; Campbell and Shiller, 1988; Fama and French, 1988, 1989; Campbell, 1991; Chen, 1991; Ferson and Harvey, 1991; Hodrick, 1992; Lettau and Ludvigson, 2005).

News about the economy is reflected in changes of these variables. We also include changes in the VIX index, the so-called fear index, since our sample period contains several crises. Monthly dividend yields are obtained from Thomson Reuters Datastream. Default spread is calculated as the difference between Moody's Baa and Aaa corporate bond yields, and the term spread is defined as the difference between 10-year and one-year maturity Treasury rates. Data on corporate bonds, Treasury rates and the VIX are obtained from the FRED database of the Federal Reserve Bank of St. Louis.

Table 5 presents the correlation matrix of these variables, along with fund flow and market return. The co-movement of flow and returns is indicated by a correlation coefficient of 0.29. Flow also has a significant negative correlation with the change of dividend yield and the VIX. Apart from Δ*Term*, most of the other variables are correlated with each other with coefficients above (below) 0.10 (-0.10).

[INSERT TABLE 5 AROUND HERE]

A general specification of our bivariate VAR-X model, where flow and market returns depend on different combinations of the previous p values of both variables, the exogenous variables (X) mentioned above and error terms is given by:

$$y_t = v + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B_0 x_t + \varepsilon_t \quad , \tag{4}$$

where, y_t is a 2x1 vector containing the time series of flow and market returns, A_1 through A_p are 2x2 matrices of parameters, x_t is a 6x1 vector of exogenous variables, B_0 is a 2x6 matrix of coefficients, v is a 2x1 vector of intercepts, and ε_t are white noise error terms. In order to carry out joint significance tests on the lags of flow and returns, we require that both time series are stationary. Phillips-Perron tests reported in Table 6 reject the unit root hypotheses at the one percent level for both flow and returns.

[INSERT TABLE 6 AROUND HERE]

We determine the optimal lag length of (4) based on Akaike's information criterion (AIC), which is minimized by including 9 lags out of a maximum of 12 lags. We also assess the validity of the VAR system by testing its dynamic stability and residual autocorrelation. None of the eigenvalues lies outside the unit root and there is no serial correlation left in the residuals. Panel A of Table 7 shows results from the VAR estimation.

[INSERT TABLE 7 AROUND HERE]

Since several lags of the variables are included in each of the equations of the system, the degrees of significance and signs vary across lags, making interpretation of individual lag coefficients inherently difficult. Both series are described by several of their own lags but also by lags of the other variable. There is some evidence of flow persistence consistent with the notion that flow is largely predictable. Flow also appears to have an effect on returns at various lags. However, there is much less evidence of return persistence in the dynamics of the VAR system.

To overcome the difficulty of drawing inferences based on individual coefficients, we turn next to test jointly flow and market return predictability. Panel B of Table 7 reports results from Granger causality tests. We find strong evidence of two-way causality. The *p*-values of both joint significance tests are below one percent. Hence, we conclude that flow contains information about market returns and vice versa. Interestingly, the coefficient on the Sell in May dummy is only significant in the flow equation but not in the return equation. After accounting for lagged flow there is no seasonality left in stock returns, which is consistent with the results discussed in the previous section. Finally, it appears that returns react more to news about the economy than flow. All coefficients on the news proxies in the return equation are statistically significant, while only news reflected in ΔDY seems to be important in the flow equation. Overall, the results provide evidence in favour of both the price-pressure and feedback hypotheses after accounting for news proxies as exogenous variables.

3.3. The January Effect

Since flow spikes in January (see Figure 1) and January falls in the winter period, we turn next to examine whether flow helps to explain the January effect. This long standing empirical regularity refers to abnormally high stock returns in January, first documented by Wachtel (1942), with more recent evidence, viz. Keim (1983), Rozeff and Kinney (1976) and Reinganum (1983), suggesting it is 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. Several explanations have been proposed for this anomaly, albeit 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 the compensation for risk. Tinic and West (1984) find the mean-variance trade-off is only present in January. Haugen and Lakonishok (1987) and Lakonishok et al. (1991) propose a window dressing hypothesis in which institutional investors improve the appearance of their portfolios 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 followed by excessive purchasing activity in January. Chen and Singal (2004) demonstrate that tax-loss selling is the main driver behind this anomaly.

To test our flow-based explanation we begin by repeating the analyses above using a January indicator. Since we do not find a January effect in the value-weighted CRSP stock market index, all the tests below are based on the CRSP equally weighted index.

[INSERT TABLE 8 AROUND HERE]

The first column in Table 8 reports an average January effect of 2.73% with a *t*-statistic of 1.95. After accounting for the effect of flow, column 2 shows the January indicator is completely subsumed.[8] The coefficient on *Flow* is 4.93, larger in magnitude than the equivalent in Table 3, and statistically significant (*t*-statistic = 6.66). Again, consistent with a lagged reversal of the price pressure effect, lagged flow is negatively related to stock returns. The estimates on flow are essentially unaffected when the price-earnings ratio and the dividend yield are included, as shown in columns 3 and 4. Column 5 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 for the same reasons as discussed above.

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:

$$r_t = \mu_t + \beta_1 r_{t-1} + \beta_2 Jan_{1995} + \beta_3 Jan_{1996} + \ldots + \beta_{21} Jan_{2014} + \varepsilon_t \qquad (5)$$

$$Flow_t = \mu_t + \beta_1 Flow_{t-1} + \beta_2 Jan_{1995} + \ldots + \beta_{21} Jan_{2014} + \varepsilon_t \qquad (6)$$

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 account for serial correlation. β_2 - β_{21} represent the abnormal return and flow in January estimated for each year over the sample period.

[8] Since the January effect is size-related, we also tested our flow-based explanation using only flow of small-cap and medium-cap funds. All results reported in this section are confirmed and, if anything, the effect of flow becomes stronger. For example, the coefficients on concurrent flow are larger in magnitude than those shown in Table 8, and have a *t*-statistic in excess of 7.3. These results can be obtained from the authors.

[INSERT TABLE 9 AROUND HERE]

Table 9 shows the January effect is positive and statistically significant in 12 out of 20 years. With a few exceptions (e.g. 2002 and the GFC), we also observe that in years with strong and positive (negative) abnormal flow, the January effect is large and positive (negative).

[INSERT TABLE 10 AROUND HERE]

Table 10 reports results of a regression in which the dependent variable is the January excess return estimated using equation (5). The only explanatory variable in the first column 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 second specification 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. Following previous studies in the literature, we define it 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.

$$PTS = \frac{\sum_{i=1}^{n}\left(\frac{price_{it,Dec.}}{price_{it,High}}\right) - 1}{n} \tag{7}$$

We include volume (*Vol*) as another potential 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. The standard deviation of market returns (*Std*) is also included to account for risk. 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.[9]

The results reported in Table 10 show that excess flow helps explain the January effect alongside alternative explanations discussed in literature. The coefficient of *Flow* is 3.17 with a *t*-statistic of 3.55. Consistent with previous studies, *PTS* is negatively related to the January effect but it is not significant in statistical terms.[10] The estimates on *Vol* and *Std* are also not statistically significant. However, if we exclude the GFC period (not reported) 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%. In any case, the coefficients on flow are consistently significant with *t*-statistics of more than 3.0. Overall, our results confirm that flow is related to the January anomaly in stock returns.

3.4. Retail versus Institutional Funds

In this section, we distinguish between flows into retail and institutional funds as they cater to very different clienteles. Funds with a minimum investment of $100,000 or more are usually classified as institutional and generally target corporations, pension funds, endowments, foundations and other large, including high net-worth, investors. On the other hand, retail funds focus on individual investors. Prior research has shown that portfolio choice, investor behaviour and the flow-performance relation between both groups are different. For example, clients of institutional funds tend to use more sophisticated performance measures such as risk-adjusted return measures or tracking error and do not chase returns in the same way as their retail counterparts (e.g. Del Guercio and Tkac, 2002; James and Karceski, 2006; Salganik and Schreiber, 2013). As a result, net flows of retail and institutional funds are likely to differ. Indeed, Figure 3 shows different patterns in flows between these two groups.

[INSERT FIGURE 3 AROUND HERE]

[9] The choice of the time period, whether it is the previous six or eleven months, does not affect the results.
[10] PTS is between -1 and 0 by construction.

Figure 3 shows that net flows of retail funds display a clear winter-summer seasonal, while those of institutional funds do not exhibit any pattern except for being positive throughout the year. Retail fund flows, by contrast, are mainly positive during winter months and negative, on average, during most summer months.

[INSERT TABLE 11 AROUND HERE]

In Table 11, we repeat our analyses for both groups separately. Consistent with Figure 3, Panel A shows that the Sell in May dummy becomes insignificant only when we account for flows of retail funds. The coefficient on the seasonal dummy in column 2 drops to 0.11% and has a *t*-statistic of 0.24. However, the Sell in May effect remains at about 1% when we account for institutional fund flows with a *t*-statistic of 1.92. This is largely due to the lack of a clear flow-performance relation in institutional funds, which is in line with James and Karceski (2006). They find that institutional fund flow is less sensitive to performance than retail fund flow, explained by the more sophisticated performance measures that this investor group implements. Accordingly, the coefficients on flow and lagged flow in column 3 are all less than two standard errors away from zero. Hence, we can further conclude that the well-known positive relation between aggregate fund flow and concurrent market return seems to stem mainly from retail flows. The coefficient on *Flow* in column one is 3.60 with a *t*-statistic of 5.25. The estimate on lagged flow is again negative, -1.70 with a *t*-statistic of -2.30.

Panel B shows essentially the same results for the January effect. After accounting for retail investor flow, the estimate drops from 2.73% to an insignificant 0.73% with a *t*-statistic of 0.53. The coefficients on lagged flow are -1.68 and -1.90 with *t*-statistics of -1.67 and -2.90, respectively. After accounting for institutional fund flow, the coefficient on the seasonal dummy is with 2.35% not much different compared to column 1 but with a *t*-statistic of 1.64. Again, none of the flow variables are related to market returns.

3.5. Seasonality in Fund Flow by Fund Style

To shed more light on the seasonality of fund flows, we calculate average monthly flows by fund style in order to examine if the documented patterns in fund flow are largely driven by investor's appetite for certain stocks that varies over the seasons. We group all the funds in our sample based on the nine Morningstar categories which are assigned based on the underlying securities in each portfolio. The categories are built along two dimensions - market capitalisation (large, mid-cap and small) and valuation ratios (growth, blend and value). Figure 4 shows that essentially all investment styles exhibit a similar winter-summer seasonal as the overall market.

[INSERT FIGURE 4 AROUND HERE]

Flow during the winter months is on average larger than during the summer months. This pattern is particularly pronounced for funds that predominantly invest in large and mid-cap growth stocks, with very low or even net outflow in most of the summer months. However, value funds also exhibit a similarly strong winter-summer seasonal. There is not much difference in the pattern between large-cap and small-cap oriented funds, or between those that predominantly invest in stocks for which neither growth nor value characteristics are dominating (blend stocks). Overall, we cannot identify a clear preference for a certain style characteristic during the winter months. Instead, the pattern appears to be strikingly robust across all categories.

4. 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 tenets. Specifically, we first document a strong winter-summer seasonality in mutual fund flows that is present across all investment styles. We then show that the Sell in May effect is positive (negative) in years when fund flows during the winter months are higher (lower) than those during the summer months. After controlling for mutual fund flows the Sell in May effect becomes insignificant. Excess fund flow in winter months explains almost half of the variation in the Sell in May effect. We also find that fund flow helps to explain the January effect. Both of these calendar month effects are mainly driven by flows of retail funds, and in particular unanticipated flows. This is an important finding since it offers a plausible explanation for the persistence of calendar month anomalies. Similarly, the well-known contemporaneous relation between fund flow and market returns is only exhibited in retail fund flow. Results from an unrestricted VAR model shows that flows contain information on return and vice versa. Our findings are therefore consistent with both the price pressure and the positive feedback trading hypothesis.

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.

Baker, Malcolm, & Wurgler, Jeffrey. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance, 61*(4), 1645-1680.

Baker, Malcolm, & Wurgler, Jeffrey. (2007). Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives, 21*(2), 129-151.

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

Campbell, John Y. (1987). Stock Returns and the Term Structure. *Journal of Financial Economics, 18*(2), 373-399.

Campbell, John Y. (1991). A Variance Decomposition for Stock Returns. *The Economic Journal, 101*, 157-179.

Campbell, John Y., & Shiller, Robert. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies, 1*(3), 195-228.

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, Nai-Fu. (1991). Financial Investment Opportunities and the Macroeconomy *The Journal of Finance, 46*(2), 529-554.

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. (1988). Dividend Yields and Expected Stock Returns. *Journal of Financial Economics, 22*(1), 3-25.

Fama, Eugene F., & French, Kenneth R. (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics, 25*(1), 23-49.

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

Ferson, Wayne E., & Harvey, Campbell R. (1991). The Variation of Economic Risk Premium. *Journal of Political Economy, 99*(2), 385-415.

Granger, Clive W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica, 37*(3), 424-438.

Guercio, Diane Del, & Tkac, Paula A. (2002). The Determinants of the Flow of Funds on Managed Portfolios: Mutual Funds vs. Pension Funds. *Journal of Financial and Quantitative Analysis, 37*(04), 523-557.

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

Hodrick, Robert J. (1992). Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. *Review of Financial Studies, 5*(3), 357-386.

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*.

James, Christopher, & Karceski, Jason. (2006). Investor Monitoring and Differences in Mutual Fund Performance. *Journal of Banking & Finance, 30*(10), 2787-2808.

Jank, Stephan. (2012). Mutual Fund Flows, Expected Returns, and the real Economy. *Journal of Banking & Finance, 36,* 3060-3070.

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.

Lettau, Martin, & Ludvigson, Sydney C. (2005). Expected Returns and Dividend Growth. *Journal of Financial Economics, 76,* 583-626.

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

Massa, Massimo, & Yadav, Vijay. (2015). Investor Sentiment and Mutual Fund Strategies. *Journal of Financial and Quantitative Analysis, 50*(4), 699-727.

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.

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.

Salganik, Galla, & Schreiber, Ammon. (2013). The Determinants of Investment Flows: Retail versus Institutional Mutual Funds. *Working Paper, Ben-Gurion University of the Negev*.

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.

Shiller, Robert, Fischer, Stanley, & Friedman, Benjamin M. (1984). Stock Prices and Social Dynamics. *Brookings Paper on Economic Activity, 1984*(2), 457-510.

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.

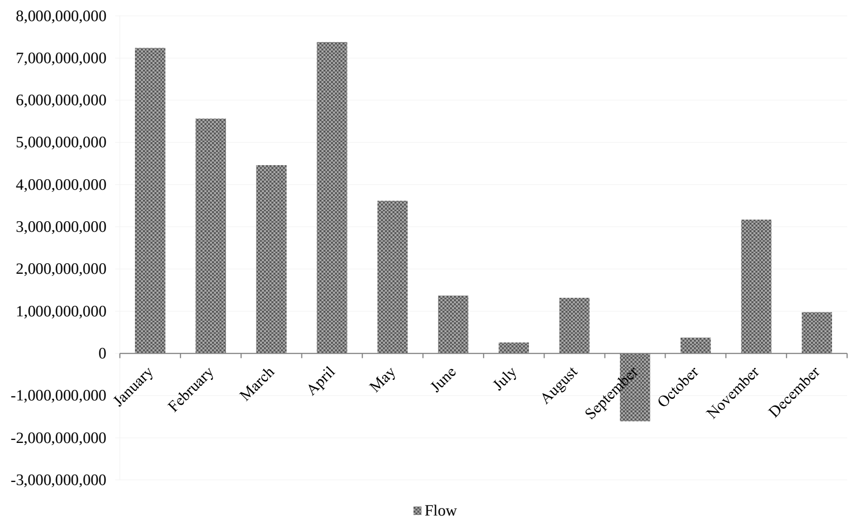
Warther, Vincent A. (1995). Aggregate Mutual Fund Flows and Security Returns. *Journal of Financial Economics, 39*, 209-235.

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

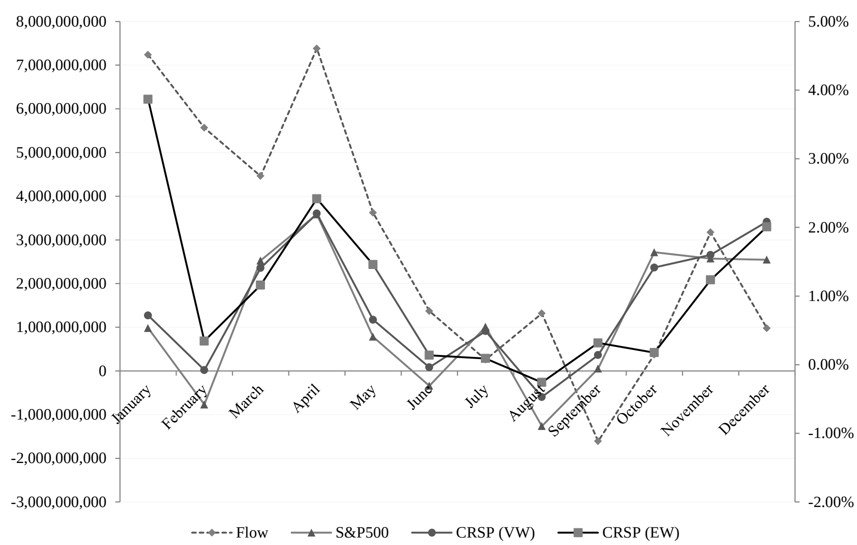
Figure 1
Mutual Fund Flows and Stock Returns

Panel A of this figure reports the average monthly flow of our sample funds (market-wide aggregate). 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

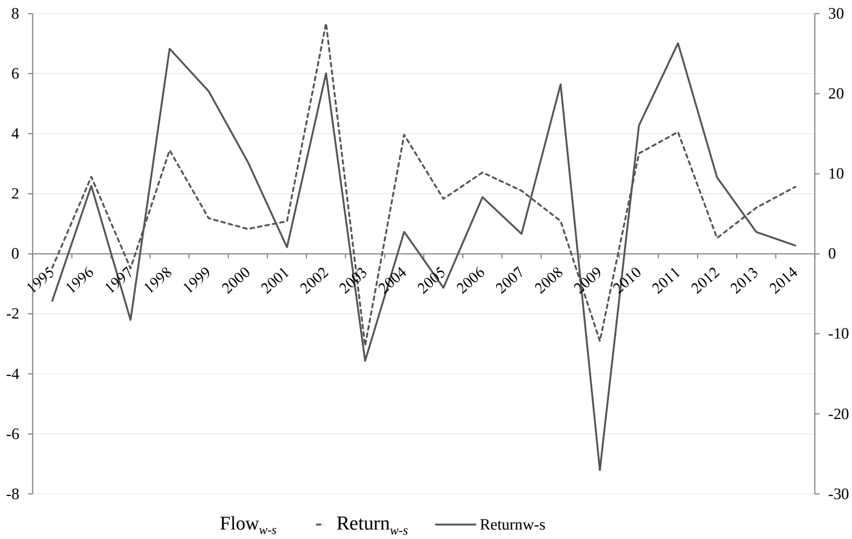


Figure 2
Time Series Plot of Flow and Predicted Flow

This figure shows the time series of aggregate mutual fund flow normalised by the value of the stock market (NYSE + AMEX + NASDAQ) in the previous month scaled by 1,000 and predicted flow. The blue line shows the predicted flow estimated using an ARIMA(1,0,1)(0,1,1)$_{12}$+c model. The time period is January 1996 to January 2015.

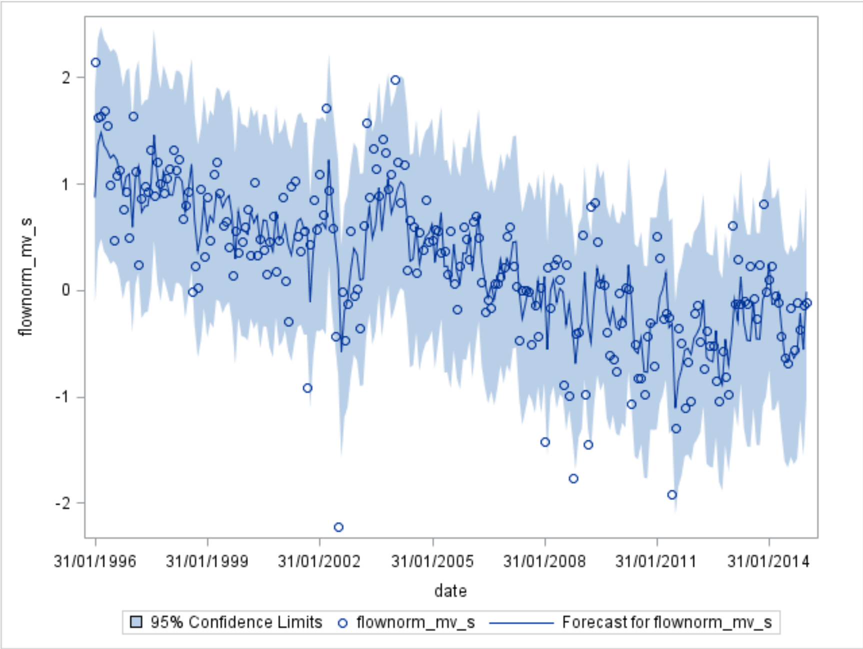


Figure 3
Retail Flow versus Institutional Flow

This figure shows the average monthly flow of retail and institutional funds separately. Funds with a minimum investment of USD 100,000 or more are categorised as institutional, all other funds are retail funds. The sample period is January 1995 to December 2014.

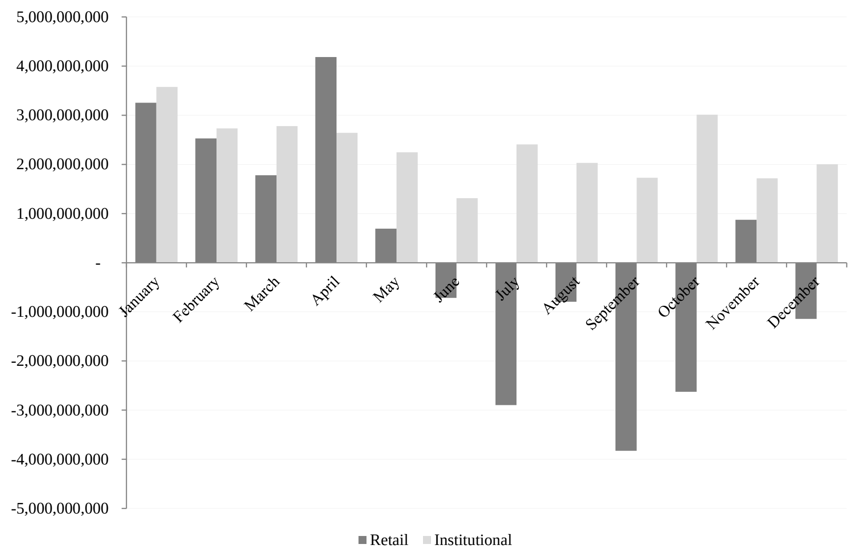


Figure 4
Fund Flow and Investment Style

This figure shows the average monthly fund flow of all funds categorised by investment style. The categories are taken from Morningstar which are assigned based on the underlying securities in each fund portfolio. The nine categories are large growth, large blend, large value, mid-cap growth, mid-cap blend, mid-cap value, small growth, small blend and small value. The time period is January 1995 to December 2014.

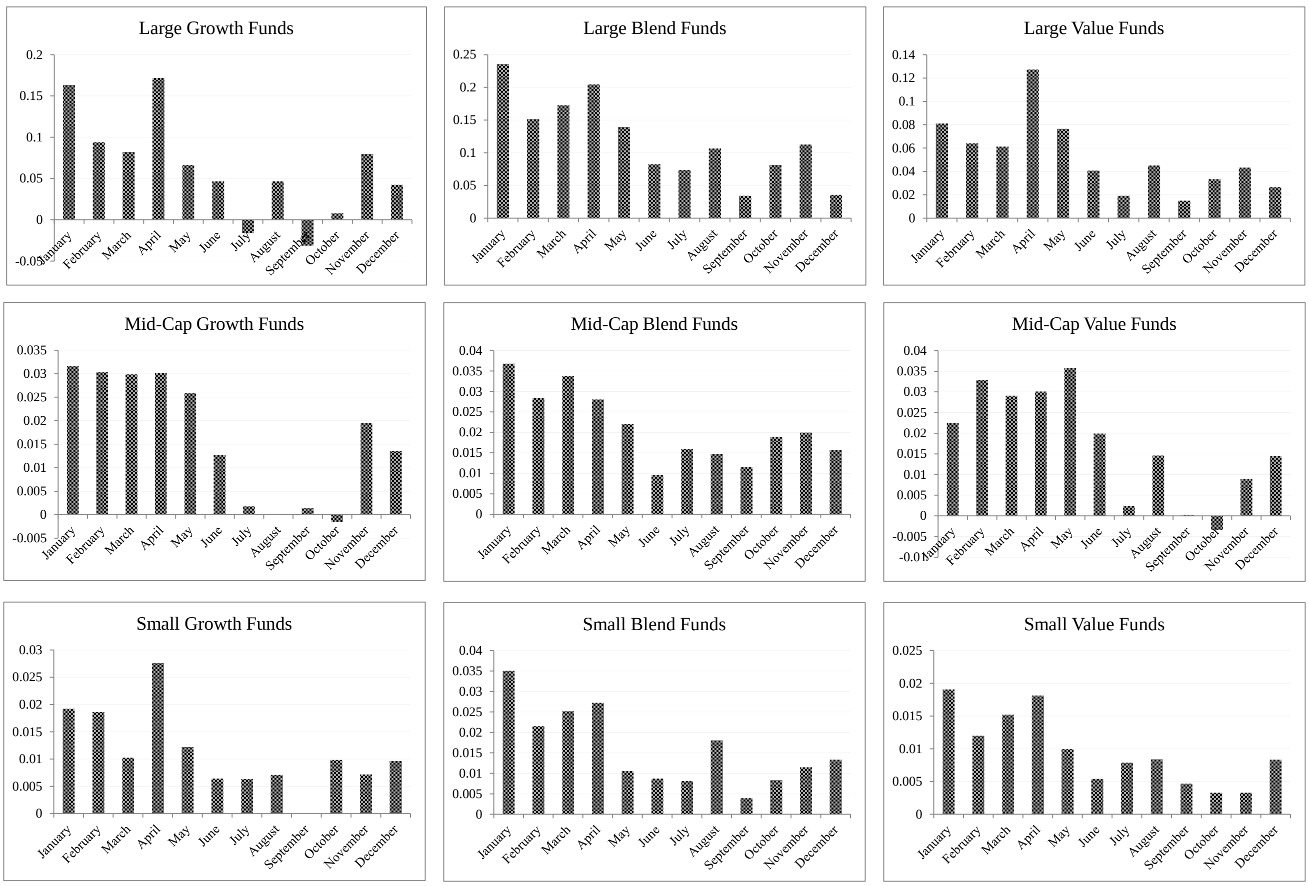


Table 1
Summary Statistics of US Equity Mutual Funds

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
1994	779	501,506	10.03%
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 Effects

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. $Sent_t$ is a proxy of market-wide sentiment. 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)	(6)
Obs.	240	240	240	240	240	240
Adj-R^2	0.01	0.08	0.13	0.16	0.15	0.15
Intercept	0.24	0.00	-1.26	**3.69**	-2.45	0.42
	(0.57)	(0.01)	(-0.64)	**(2.77)**	(-1.44)	(1.06)
Hal	**0.96**	0.50	0.10	0.14	0.12	0.14
	(1.85)	(1.03)	(0.22)	(0.30)	(0.27)	(0.32)
$Flow_t$		**1.67**	**3.45**	**3.37**	**3.31**	**3.33**
		(3.71)	**(5.21)**	**(5.50)**	**(5.38)**	**(5.44)**
$Flow_{t-1}$			**-1.50**	**-1.41**	**-1.46**	**-1.53**
			(-2.31)	**(-2.24)**	**(-2.33)**	**(-2.38)**
$Flow_{t-2}$			0.01	0.09	0.04	0.06
			(0.01)	(0.12)	(0.06)	(0.07)
$Flow_{t-3}$			-0.61	-0.53	-0.58	-0.61
			(-0.98)	(-0.88)	(-0.96)	(-1.00)
PE_{t-1}				**-0.17**		
				(-2.71)		
DY_{t-1}					1.50	
					(1.62)	
$Sent_t$						**-1.03**
						(-3.24)

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

Panel A of this table reports statistics for different models to estimate expected and unexpected flow. The first three columns of Panel B report 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 a seasonal ARIMA $(1, 0, 1)(0, 1, 1)_{12}$+c model (model no. 1 in Panel A). 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.

Panel A - Statistical Models for Expected and Unexpected Flow

No.	Model		AIC	RMSE
1	Seasonal ARIMA with white noise errors incl. constant	$ARIMA(1,0,1)(0,1,1)_{12}$+c	344.63	0.51
2	Equivalent to simple exponential smoothing	ARIMA(0,1,1)	354.75	0.51
3	Seasonal ARIMA with white noise errors without constant	$ARIMA(1,0,1)(0,1,1)_{12}$	355.25	0.52
4	Seasonal random trend model with MA(1) and SMA(1) terms	$ARIMA(0,1,1)(0,1,1)_{12}$	360.58	0.53
5	Equivalent to seasonal exponential smoothing	$ARIMA(0,1,13)(0,1,0)_{12}$	364.19	0.52
6	Equivalent to Winters method	$ARIMA(0,1,13)(0,1,0)_{12}$+c	365.71	0.52
7	Equivalent to double exponential smoothing	ARIMA(0,2,2)	376.06	0.53
8	Seasonal random walk model with AR(1) term	$ARIMA(1,0,0)(0,1,0)_{12}$+c	472.42	0.68
9	Seasonal random walk model	$ARIMA(0,0,0)(0,1,0)_{12}$+c	516.78	0.75
10	Seasonal random trend model	$ARIMA(0,1,0)(0,1,0)_{12}$	541.63	0.79

Panel B - Regressions of Return (Flow) on Flow (Return)

	Dependent Variable				
	Market Return			Expected Flow	Unexpected Flow
	(1)	(2)	(3)	(4)	(5)
Obs.	228	228	228	228	228
Adj-R^2	0.00	0.15	0.15	0.05	0.18
Intercept	0.15	**0.41**	0.37	**0.21**	-0.04
	(0.32)	**(1.15)**	(0.97)	**(2.84)**	(-1.52)
Hal	**0.98**	0.50	0.46		
	(1.86)	(1.11)	(1.00)		
Expected Flow	(-0.06)		(0.27)		
	(-0.10)		(0.51)		
Unexpected Flow		**3.44**	**3.47**		
		(5.66)	**(5.56)**		
Return				0.00	**0.04**
				(-0.17)	**(5.29)**
Return_{t-1}				**0.02**	**0.02**
				(2.87)	**(2.69)**
Return_{t-2}				**0.02**	0.00
				(3.46)	(-0.49)
Return_{t-3}				**0.01**	-0.01
				(2.36)	(-0.77)

Table 5
Correlation between Flow, Return and News Proxies

This table reports correlation coefficients between normalised mutual fund flow, market return and predictive variables. ΔDY is the change in the dividend-price ratio of the S&P 500 index (in log terms), ΔDefault is the change in the default spread, ΔTerm is the change in the term spread, ΔT-bill is the change in the 3-months T-bill rate and ΔVIX is the change in the Chicago Board Options Exchange (CBOE) Volatility index. The sample period is January 1995 to December 2014.

	Flow	Ret	ΔDY	ΔDefault	ΔTerm	ΔT-bill	ΔVIX
Flow	1						
Ret	0.29	1					
ΔDY	-0.29	-0.96	1				
ΔDefault	-0.11	-0.14	0.15	1			
ΔTerm	0.05	0.03	-0.06	0.07	1		
ΔT-bill	0.14	0.11	-0.09	-0.06	-0.47	1	
ΔVIX	-0.17	-0.71	0.72	0.12	-0.05	-0.05	1

Table 6
Unit Root Test of Fund Flow and Market Returns

This table reports estimation results for the Phillips-Perron unit root test with automatic Newey-West bandwidth using Bartlett kernel. The null hypothesis is that both flow and return have a unit root. The sample period is January 1995 to December 2014.

		Flow		Return	
		Adj. *t* -stat.	Prob.*	Adj. *t* -stat.	Prob.*
Phillips-Perron test statistic		-6.49	0.000	-14.20	0.000
Test critical values:	1% level	-3.46		-3.46	
	5% level	-2.87		-2.87	
	10% level	-2.57		-2.57	

*MacKinnon (1996) one-sided p-values.

	Flow	Return
Residual variance (no correction)	0.28	18.92
HAC corrected variance (Bartlett kernel)	0.28	20.75

Table 7
Vector Autoregression Analysis of Fund Flow and Market Returns

Panel A of this table reports estimation results from a vector autoregressive regression of normalised fund flow and market returns. *Hal* is the Sell in May dummy and is 1 for the period November-April and 0 otherwise. ΔDY is the change in the dividend yield of the S&P 500 index (logged), ΔDefault is the change in the default spread, ΔTerm is the change in the term spread, ΔT-bill is the change in the 3-months T-bill rate and ΔVIX is the change in the CBOE Volatility index. Panel B reports results of Granger Causality Wald tests. The sample period is January 1995 to December 2014.

Panel A - Vector Autoregression of Flow and Market Returns

	Flow		Return	
	Coef.	*t* -stat.	Coef.	*t* -stat.
N	231		231	
Adj. R^2	0.68		0.95	
Constant	**-0.09**	**(-2.01)**	**0.54**	**(4.75)**
$Flow_{t-1}$	**0.22**	**(3.36)**	0.14	(0.85)
$Flow_{t-2}$	**0.14**	**(2.03)**	-0.07	(-0.40)
$Flow_{t-3}$	**0.16**	**(2.38)**	**0.29**	**(1.78)**
$Flow_{t-4}$	0.10	(1.38)	-0.26	(-1.51)
$Flow_{t-5}$	**0.14**	**(2.01)**	**-0.50**	**(-2.96)**
$Flow_{t-6}$	0.07	(1.01)	0.20	(1.20)
$Flow_{t-7}$	-0.01	(-0.14)	-0.12	(-0.71)
$Flow_{t-8}$	0.04	(0.66)	-0.07	(-0.42)
$Flow_{t-9}$	0.07	(1.09)	0.02	(0.10)
$Return_{t-1}$	**0.02**	**(3.09)**	0.02	(1.26)
$Return_{t-2}$	0.00	(0.02)	0.04	(1.88)
$Return_{t-3}$	-0.01	(-1.30)	0.03	(1.42)
$Return_{t-4}$	0.01	(1.20)	0.03	(1.84)
$Return_{t-5}$	-0.01	(-1.22)	**0.06**	**(3.46)**
$Return_{t-6}$	-0.01	(-1.10)	**0.03**	**(1.85)**
$Return_{t-7}$	**-0.01**	**(-1.71)**	0.02	(1.18)
$Return_{t-8}$	-0.01	(-1.08)	**0.04**	**(2.18)**
$Return_{t-9}$	-0.01	(-1.34)	**0.04**	**(2.51)**
HAL	**0.25**	**(4.04)**	0.06	(0.42)
ΔDY	**-2.40**	**(-2.53)**	**-91.49**	**(-39.26)**
ΔDefault	0.02	(0.14)	**0.72**	**(2.06)**
ΔTerm	0.27	(1.48)	**-0.78**	**(-1.75)**
ΔT-bill	0.26	(1.39)	**-0.79**	**(-1.69)**
ΔVIX	-0.01	(-1.29)	**-0.06**	**(-2.54)**

Continued: Panel B - Granger Causality Wald Tests

Flow			Return		
Excluded	Chi-sq	Prob.	Excluded	Chi-sq	Prob.
Return	22.49	0.008	Flow	23.12	0.006
All	22.49	0.008	All	23.12	0.006

Table 8
Mutual Fund Flows and the January Effect

This table reports estimation results for different variants of equation (2). 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. 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 9
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 (5). Abnormal flow of the sample funds in January based on equation (6) 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 10

Relationship between $\boldsymbol{Flow_t}$, $\boldsymbol{PTS_t}$, $\boldsymbol{Vol_t}$, $\boldsymbol{Std_t}$ 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)	

Table 11
Retail versus Institutional Funds and Seasonalities in Stock Returns

This table reports estimation results for different variants of equation (2). The dependent variable in Panel A 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. The dependent variable in Panel B 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 retail and institutional funds. The latter are fund with a minimum investment of USD 100,000 or more. Standard errors are corrected for heteroskedasticity and autocorrelation. *t*-statistics are reported in parentheses. The sample period is January 1995 to December 2014.

Panel A - Sell in May Effect

		Fund Types	
		Retail Funds	Institutional Funds
	(1)	(2)	(3)
N	240	240	240
Adj-R^2	0.01	0.14	0.01
Intercept	0.24	0.57	0.82
	(0.57)	(1.56)	(1.35)
Hal	**0.96**	0.11	**1.03**
	(1.85)	(0.24)	**(1.92)**
Flow		**3.60**	1.52
		(5.25)	(0.68)
$Flow_{t-1}$		**-1.70**	-3.38
		(-2.30)	(-1.41)
$Flow_{t-2}$		0.01	-1.85
		(0.01)	(-1.10)
$Flow_{t-3}$		-1.02	-0.16
		(-1.50)	(-0.07)

Continued: Panel B - January Effect

		Fund Types	
		Retail Funds	Institutional Funds
	(1)	(2)	(3)
N	240	240	240
Adj-R^2	0.01	0.20	0.01
Intercept	**0.89**	**0.94**	**1.31**
	(2.08)	**(2.44)**	**(1.68)**
Jan	**2.73**	0.73	2.35
	(1.99)	(0.53)	(1.64)
Flow		**5.30**	3.40
		(6.17)	(1.11)
$Flow_{t-1}$		**-1.68**	-1.32
		(-1.67)	(-0.40)
$Flow_{t-2}$		-0.54	-2.27
		(-0.59)	(-1.04)
$Flow_{t-3}$		**-1.90**	-2.31
		(-2.90)	(-0.74)