Appendix A: Model

Consider a model where market makers post bid and ask prices. Informed and uninformed investors trade against these prices. The arrival of these investors is governed by Poisson processes with intensity λ per unit of time. Uninformed investors can be of different types including momentum investors, disposition traders, contrarian traders, and tax-motivated traders.

Consider an investor whose trading decision is influenced by returns from a previous period. The Poisson intensity for this investor would vary with past returns. Formally, we specify this investor's arrival intensity as:

$$\lambda_{i,T} = \left| \alpha_i \times S_T + \varepsilon_{i,T} \right|, \tag{1}$$

where S_T is a signal that is a monotonically increasing function of past returns, and α_i is the sensitivity of investor i's propensity to trade to S_T. Factors other than past returns may also influence investors to trade. The supplementary signal ε_i, which is orthogonal to S_T, captures the impact of such factors. We assume that ε_i is an investor specific supplementary signal which is independent and identical across investors. We also assume that ε_i is normally distributed with zero mean and variance σ_ε^2.

The theories we examine have different predictions about how S_T influences the propensity to trade. For example, the propensity of a momentum trader to buy a stock increases with S_T, but that of a contrarian trader or a disposition effect trader decreases with S_T. The factor α_i captures such heterogeneity. Momentum investors and year-end tax loss traders are more likely to buy a stock when S_T increases and to sell a stock when S_T decreases, and for these investors $\alpha_i = +1$. Similarly, for contrarian traders $\alpha_i = -1$. For ease of exposition, we refer to all investors with $\alpha_i = +1$ as momentum investors and all investors with $\alpha_i = -1$ as contrarian investors.[1] For disposition investors, $\alpha_i = -1$ for selling decisions but $\alpha_i = 0$ for buying decisions.

Some investors' trading decisions are entirely unrelated to past returns. An example of such investors would be informed traders who are not influenced by S_T but base their trading decisions on their private signals. Another example would be investors who trade for their own personal liquidity reasons. We refer to these types of investors generically as "Other investors." The arrival intensity of these investors is given by:

$$\lambda_{i,T} = \left| u_{i,T} \right|, \tag{2}$$

where u_i is independently and identically distributed across investors. The signal u_i is normally distributed with zero mean and variance σ_u^2.

The terms $\alpha_i \times S_T + \varepsilon_{i,T}$ in equation (1) and $u_{i,T}$ in equation (2) can be either positive or negative. An investor in our model initiates a buy trade conditional on arrival if the term is positive and a sell trade if the term is negative. Due to the supplementary signal, traders of the same type need not necessarily trade in the same direction. For example, some momentum traders (investors with $\alpha_i = +1$) may sell past winners if the supplementary signal is sufficiently negative while others may buy the same stock. Similarly, contrarian investors (investors with $\alpha_i = -1$) may sell losers if they experience some shock that requires them to liquidate their losers, i.e. if ε_i is sufficiently negative.

There are N_M momentum investors, N_C contrarian traders, N_D disposition traders, and N_O

---

[1] We can also allow for different investors to have different sensitivities to the signal S_T by specifying α_i as a continuous variable rather than a binary variable. However, such heterogeneity does not qualitatively change any of the results and hence we opt for a simple binary specification.

other traders in the economy. We analyze below the expected number of arrivals of buy and sell traders. We can allow for the number of shares that each arriving trader buys or sells to be a random variable that is iid across investors and get qualitatively similar results for the relation between the number of shares that investors trade and past returns, but we present the details only for the orders based on the number of trades.

We first analyze the trading decisions of each type of investors separately and then present the results for aggregate trades. A momentum trader initiates a buy order conditional on arrival if $S_T + \varepsilon_{i,T} > 0$ and a sell order otherwise. We assume that the number of traders of each type is sufficiently large so that we can apply the law of large numbers for investor-specific signals. Therefore, the expected number of buys and sells is given by:

$$\begin{aligned}
&\text{Buys}_{M,T} = N_M \times E\left(S_T + \varepsilon_{i,T} \mid S_T + \varepsilon_{i,T} > 0\right) = N_M \times \left[S_T \Phi(S_T/\sigma_\varepsilon) + \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&-\text{Sells}_{M,T} = N_M \times E\left(S_T + \varepsilon_{i,T} \mid S_T + \varepsilon_{i,T} < 0\right) = N_M \times \left[S_T \Phi(-S_T/\sigma_\varepsilon) - \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&\text{Buys}_{M,T} - \text{Sells}_{M,T} = N_M \times E\left(S_T + \varepsilon_{i,T} \mid S_T\right) = N_M \times S_T \\
&\text{Buys}_{M,T} + \text{Sells}_{M,T} = N_M \times E\left(\left|S_T + \varepsilon_{i,T}\right| \mid S_T\right) = N_M \times \left[2\sigma_\varepsilon \phi(S_T/\sigma_\varepsilon) + 2S_T \Phi(S_T/\sigma_\varepsilon) - S_T\right],
\end{aligned} \tag{3}$$

where ϕ is the standard normal density function and Φ is the standard normal cumulative density function. Thus, the difference between buys and sells is linear function of the signal, S_T. With a large number of traders, the noise, $\varepsilon_{i,T}$, in the trading intensity disappears for the difference between buy and sell trades but manifests itself in the total trading volume of buys and sells.

The trading decisions of contrarian investors are analogous. If $-S_T + \varepsilon_{i,T} > 0$, they initiate a buy trade. On the other hand, if $-S_T + \varepsilon_{i,T} < 0$, they initiate a sell trade. The expected number of trades from contrarian traders conditional on S_T is:

$$\begin{aligned}
&\text{Buys}_{C,T} = N_C \times E\left(-S_T + \varepsilon_{i,T} \mid -S_T + \varepsilon_{i,T} > 0\right) = N_C \times \left[-S_T \Phi(-S_T/\sigma_\varepsilon) + \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&-\text{Sells}_{C,T} = N_C \times E\left(-S_T + \varepsilon_{i,T} \mid -S_T + \varepsilon_{i,T} < 0\right) = N_C \times \left[-S_T \Phi(S_T/\sigma_\varepsilon) - \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&\text{Buys}_{C,T} - \text{Sells}_{C,T} = N_C \times E\left(-S_T + \varepsilon_{i,T} \mid S_T\right) = -N_C \times S_T \\
&\text{Buys}_{C,T} + \text{Sells}_{C,T} = N_C \times E\left(\left|-S_T + \varepsilon_{i,T}\right| \mid S_T\right) = N_C \times \left[2\sigma_\varepsilon \phi(S_T/\sigma_\varepsilon) + 2S_T \Phi(S_T/\sigma_\varepsilon) - S_T\right].
\end{aligned} \tag{4}$$

Disposition traders initiate a sell trade if $-S_T + \varepsilon_{i,T} < 0$. The expected number of trades from disposition traders conditional on S_T is:

$$\begin{aligned}
&\text{Buys}_{D,T} = 0 \\
&-\text{Sells}_{D,T} = N_D \times E\left(-S_T + \varepsilon_{i,T} \mid -S_T + \varepsilon_{i,T} < 0\right) = N_D \times \left[-S_T \Phi(S_T/\sigma_\varepsilon) - \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&\text{Buys}_{D,T} - \text{Sells}_{D,T} = N_D \times \left[-S_T \Phi(S_T/\sigma_\varepsilon) - \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right] \\
&\text{Buys}_{D,T} + \text{Sells}_{D,T} = N_D \times \left[S_T \Phi(S_T/\sigma_\varepsilon) + \sigma_\varepsilon \phi(S_T/\sigma_\varepsilon)\right].
\end{aligned} \tag{5}$$

Other traders initiate buy trades if $u_{i,T} > 0$, and sell trades if $u_{i,T} < 0$. The expected number of buy or sell trades is $\sigma_u \phi(0)$. Since the trading decisions of other traders are unrelated to past returns, the expected difference between the number of buy and sell orders for these investors is equal to zero.

Aggregating orders over all types of traders, we get:

$$\begin{aligned}
&\text{Buys}_T = (N_M + N_C)S_T\Phi\left(S_T/\sigma_\varepsilon\right) + (N_M + N_C)\sigma_\varepsilon\phi\left(S_T/\sigma_\varepsilon\right) - N_C S_T + N_o\sigma_u\phi(0)\\
&-\text{Sells}_T = -(N_M + N_C + N_D)S_T\Phi\left(S_T/\sigma_\varepsilon\right) - (N_M + N_C + N_D)\sigma_\varepsilon\phi\left(S_T/\sigma_\varepsilon\right) + N_M S_T - N_o\sigma_u\phi(0)\\
&\text{Buys}_T - \text{Sells}_T = (N_M - N_C)S_T - N_D S_T\Phi\left(S_T/\sigma_\varepsilon\right) - N_D\sigma_\varepsilon\phi\left(S_T/\sigma_\varepsilon\right)\\
&\text{Buys}_T + \text{Sells}_T = (2N_M + 2N_C + N_D)\left[S_T\Phi\left(S_T/\sigma_\varepsilon\right) + \sigma_\varepsilon\phi\left(S_T/\sigma_\varepsilon\right)\right] - \left(N_M + N_C\right)S_T + 2N_o\sigma_u\phi(0).
\end{aligned} \tag{6}$$

The aggregate buy and sell orders in the economy, thus, depend on relative numbers of momentum, contrarian, and disposition investors. We propose to empirically examine whether investors' aggregate trading decisions are related to past returns and the relative importance of various hypotheses to explain any such relation. The slope coefficient in a regression of, for example, Buys_T on S_T is given by β = cov(Buys_T, S_T)/var(S_T). The sign of the slope depends on $\text{E}\left[\text{Buys}_\text{T} \times S_\text{T}\right]$ where we have assumed that E[S_T]=0 without loss of generality. The proposition below discusses the results from a regression of buys and sells on past returns to test the relative importance of momentum and contrarian investors.

Proposition 1: The sign of the slope coefficient β in a regression of Buys or Sells or Buys−Sells on the signal (and hence past returns) depends on the relative number of momentum, contrarian, and disposition traders. In particular:

$$\begin{aligned}
&\beta_{\text{Buys}} > 0 \text{ iff } N_M > N_C\\
&\beta_{-\text{Sells}} > 0 \text{ iff } N_M > N_C + N_D\\
&\beta_{\text{Buys}-\text{Sells}} > 0 \text{ iff } N_M > N_C + 0.5N_D.
\end{aligned}$$

Proof:

$$\begin{aligned}
E\left[\text{Buys}_T \times S_T\right] &= \left(N_M + N_C\right)E\left[S_T^2\Phi\left(S_T/\sigma_\varepsilon\right)\right] + \left(N_M + N_C\right)\sigma_\varepsilon E\left[S_T\phi\left(S_T/\sigma_\varepsilon\right)\right]\\
&\quad - N_C E\left[S_T^2\right] + N_O\sigma_u\phi(0)E\left[S_T\right]\\
&= 0.5\left(N_M - N_C\right)E\left[S_T^2\right] > 0 \text{ iff } N_M > N_C\\
E\left[-\text{Sells}_T \times S_T\right] &= \left(N_M + N_C + N_D\right)E\left[S_T^2\Phi\left(S_T/\sigma_\varepsilon\right)\right] - \left(N_M + N_C + N_D\right)\sigma_\varepsilon E\left[S_T\phi\left(S_T/\sigma_\varepsilon\right)\right]\\
&\quad + N_M E\left[S_T^2\right] + N_O\sigma_u\phi(0)E\left[S_T\right]\\
&= 0.5\left(N_M - N_C - N_D\right)E\left[S_T^2\right] > 0 \text{ iff } N_M > N_C + N_D\\
E\left[\left(\text{Buys}_T - \text{Sells}_T\right) \times S_T\right] &= \left(N_M - N_C - 0.5N_D\right)E\left[S_T^2\right] > 0 \text{ iff } N_M > N_C + 0.5N_D,
\end{aligned}$$

where we have used the facts that $E\left[X^2\Phi(X)\right] = 0.5\,\text{var}(X)$ and $E\left[X\phi(X)\right] = 0$ for a mean zero normally distributed random variable X.

This result is not totally unexpected in our model because momentum investors' trades are positively related to past returns, contrarian/disposition investors' trades are negatively related to returns, and other investors' trades are unrelated to returns. However, our result shows that we can determine the relative importance of these investors even when they use additional supplementary signals that lead to some investors trading against their types (e.g. some momentum traders sell even when returns are positive). The result also holds even when these investors trade along with other/informed investors.

Our next proposition presents the relation between β and the relative number of Other traders and the importance of S_T relative to the supplementary signal.

Proposition 1a: As the number of Other traders in the economy increases, the regression slope coefficient of Buys or −Sells or Buys−Sells on past returns decreases in magnitude.

$$\frac{\partial|\beta|}{\partial\theta}<0,$$

where θ is the fraction of Other traders in the population. Similarly, the regression slope coefficient decreases in magnitude with an increase in the relative importance of the supplementary signal i.e.

$$\frac{\partial|\beta|}{\partial\sigma_\varepsilon}<0.$$

Proof: The unconditional turnover, Γ, is given by

$$\begin{aligned}\Gamma \equiv E(\text{Buys}+\text{Sells}) &= \left(2N_M+2N_C+N_D\right)E\left[S_T\Phi\left(S_T/\sigma_\varepsilon\right)+\sigma_\varepsilon\phi\left(S_T/\sigma_\varepsilon\right)\right]+2N_O\sigma_\varepsilon\phi(0)\\ &= \left(2N_M+2N_C+N_D\right)\phi(0)\sqrt{\sigma_S^2+\sigma_\varepsilon^2}+2N_O\sigma_\varepsilon\phi(0),\end{aligned}$$

where we have used the fact that $E\left[|X|\right]=\sigma_X\phi(0)$. It is easy to show that $\partial\Gamma/\partial N_O>0$ and $\partial\Gamma/\partial\sigma_\varepsilon>0$. Since scaled Buys and Sales are defined by dividing with Γ, the rest of the proposition follows.

The intuition for this result is that the trades of Other investors and the variability of the supplementary signal dampens the amount of trade initiations related to past returns. Proposition 2 offers reasons why the slope coefficients on past returns may vary over time. For example, the slope coefficients would be smaller in periods when investors rely less on past returns for their trading decisions and more on other signals.

Next, we turn to the impact of the signal and hence past returns on the trading volume as measured by turnover. In the context of our model, turnover is defined simply as the total volume (buys plus sells).

Proposition 2: The sign of the regression slope coefficient of turnover on past returns is independent of the number of momentum and contrarian traders.

Proof:

$$\begin{aligned}E\left[\left(\text{Buys}_T+\text{Sells}_T\right)\times S_T\right] &= \left(2N_M+2N_C+N_D\right)E\left[S_T^2\Phi\left(S_T/\sigma_\varepsilon\right)\right]+\left(2N_M+2N_C+N_D\right)\sigma_\varepsilon E\left[S_T\phi\left(S_T/\sigma_\varepsilon\right)\right]\\ &\quad -\left(N_M+N_C\right)E\left[S_T^2\right]+N_O\sigma_u\phi(0)E\left[S_T\right]\\ &= 0.5N_D E\left[S_T^2\right].\end{aligned}$$

Therefore, the slope coefficient of turnover on S_T does not depend on N_M or N_C.

This proposition implies that we cannot determine whether momentum traders or contrarian traders play a dominant role based on the relation between turnover and past returns. Thus, in order to differentiate among these hypotheses it is important to assess the impact of past returns on Buys, Sells, and order imbalances.

Our model thus far does not specify the horizon that investors use to measure past returns.

This feature of the model reflects the fact that the theories we test do not specify a particular horizon over which past returns would influence investors' decisions to trade. For instance, the model of Brennan and Cao (1997) does not address the issue of which return horizon the momentum investors should consider to optimally extract the information that underlies the informed investor trades. Similarly, disposition effect investors who bought shares two periods back would trade based on cumulative returns over the two periods while those who bought the shares one period back would trade based on one period returns. In general, different investors are likely to use returns over different horizons in their trading decisions.

We now generalize the model to examine the empirical implications of such heterogeneity across investors. Specifically, we generalize the model to two signals, S_{T1} and S_{T2}, related to past returns from two non-overlapping periods. Some investors trade based on the cumulative signals and others trade based only on S_{T1}.[2] Let the number of momentum, contrarian, and disposition traders who trade based on the first signal (S_{T1}) be N_{M1}, N_{C1}, and N_{D1}, respectively, and the number these traders who trade based on the cumulative signal (S_{T1}+S_{T2}) be N_{M2}, N_{C2}, and N_{D2}, respectively.

Proposition 3: Suppose we fit the following multivariate regression:

$$\text{Buys}_T = a + \beta_1 S_{T1} + \beta_2 S_{T2} + v_T.$$

Then:

$$\beta_1 > 0 \text{ iff } \left(N_{M1} + N_{M2}\right) > \left(N_{C1} + N_{C2}\right)$$
$$\beta_2 > 0 \text{ iff } N_{M2} > N_{C2}.$$

Analogous relations obtain for −Sells and Buys−Sells.

Proof: The total buys are given by

$$\begin{aligned}\text{Buys}_T &= \left(N_{M1}+N_{C1}\right)S_{T1}\Phi\left(S_{T1}/\sigma_\varepsilon\right) + \left(N_{M1}+N_{C1}\right)\sigma_\varepsilon\phi\left(S_{T1}/\sigma_\varepsilon\right) - N_{C1}S_{T1} \\ &+ \left(N_{M2}+N_{C2}\right)\left(S_{T1}+S_{T2}\right)\Phi\left(\frac{S_{T1}+S_{T2}}{\sigma_\varepsilon}\right) + \left(N_{M2}+N_{C2}\right)\sigma_\varepsilon\phi\left(\frac{S_{T1}+S_{T2}}{\sigma_\varepsilon}\right) \\ &- N_{C2}\left(S_{T1}+S_{T2}\right) + N_O\sigma_u\phi(0).\end{aligned}$$

Consider now the regression slope coefficient β_1 = cov(Buys$_T$, S_{T1})/var(S_{T1}) and β_2 = cov(Buys$_T$, S_{T2})/var(S_{T2}). Since S_{T1} and S_{T2} are signals pertain to past returns from non-overlapping periods, we assume they are orthogonal. Then, β_1 and β_2 are also the regression coefficients in the multivariate regression of OIB on S_{T1} and S_{T2}. The sign of β_1 depends on the sign of E(Buys$_T$ ×S_{T1}). We have:

$$\begin{aligned}\text{E}\left[\text{Buys}_T \times S_{T1}\right] &= 0.5\left(N_{M1}+N_{C1}\right)E\left[S_{T1}^2\right] - N_{C1}E\left[S_{T1}^2\right] + 0.5\left(N_{M2}+N_{C2}\right)E\left[S_{T1}^2\right] - N_{C2}E\left[S_{T1}^2\right] \\ &= 0.5\left[\left(N_{M1}-N_{C1}\right)+\left(N_{M2}-N_{C2}\right)\right]E\left[S_{T1}^2\right] > 0\end{aligned}$$

if $N_{M1} + N_{M2} > N_{C1} + N_{C2}$. Similarly, the sign of β_2 depends on the sign of E(Buys$_T$ ×S_{T2}). We have:

$$\begin{aligned}\text{E}\left[\text{Buys}_T \times S_{T2}\right] &= 0.5\left(N_{M2}+N_{C2}\right)E\left[S_{T1}^2\right] - N_{C2}E\left[S_{T1}^2\right] \\ &= 0.5\left(N_{M2}-N_{C2}\right)E\left[S_{T1}^2\right] > 0 \text{ iff } N_{M2} > N_{C2}.\end{aligned}$$

[2] We can allow for some traders to trade based only on S_{T2} and for traders to trade on cumulative returns over more than two periods and get similar results.

This proposition shows that, if the number of momentum traders who trade based on S_{T1} is greater than the number of contrarian traders, then the slope coefficient on S_{T1} is positive regardless of whether they trade based only on this signal or based on a cumulative signal that includes S_{T1}. Intuitively, even when some investors base their decisions on cumulative signals, the direction of the marginal effect of one-period signals is the same as whether these investors trade only based on the one period signal or based on a cumulative signal. Although we write the slope coefficients in terms of the signals, we get the same results with returns because the signals are monotonically increasing functions of returns. This proposition shows that we can use an additively separable regression specification to assess how past returns affect buys and sells even if some investors use cumulative returns in their trading decisions.

Appendix B: Excluding Lagged Dependent Variables

This appendix reports the results of regression (7) in the paper after excluding the lagged dependent variables from the right-hand side of the regression. Table B1 reports the results of this regression. The main difference between the coefficients on lagged returns in Table B1 and Table 2 is that most of the coefficients in Table B1 are significantly larger in magnitude. Because Y_i is correlated with returns lagged Y_i's in Table 2 absorb some of the effects of lagged returns in Table B1.

In terms of the signs of the coefficient estimates, the one month lagged return coefficient is positive in the Buys regressions in Table B1 while it is negative in Table 2. Also, in the Buys−Sells regressions with shares traded the first two lagged return coefficients are positive albeit insignificantly so in Table B1 while they are significantly negative in Table 2. Note that the economic significance of the impact of lagged returns on Buys, −Sells and Buys−Sells is larger than that in Table 2, probably because in Table 2 the lagged Y_i's absorb some of the effect of the lagged returns.

The likely reason for the difference between the results is the omitted variable bias. The one-month lagged return is essentially uncorrelated with all other explanatory variables in regression (7) except with $Y_{i,t-1}$. When $Y_{i,t-1}$ is excluded from the regression, the slope coefficient on one-month lagged returns partly absorbs the positive relation between $Y_{i,t}$ and $Y_{i,t-1}$ and hence the slope coefficient in Table B1 is different and often larger than that in Table 2. Therefore, it is important to include lagged order imbalances in the regressions.

Table B1: Cross-sectional determinants of buys and sells for NYSE stocks (without lags of dependent variables)

We run the following cross-sectional regression each month:

$$Y_{i,t} = a_t + \sum_{j=1}^{12} b_{t,j} R_{i,t-j} + c_t \ln\left(Size_{i,t-1}\right) + d_t \ln\left(BM_{i,t-1}\right) + e_t Vol_{i,t-1} + u_{i,t},$$

where Y is either Buys, or −Sells, or Buys−Sells. We measure Buys and Sells both in number of trades and number of shares. We scale these variables by the average number of trades / trading volume over the previous 12 months (i.e. months $t-1$ to $t-12$). Return is in percent per month, Size is market capitalization in millions of dollars, BM is book-to-market, and Vol is the standard deviation of returns in percent per month calculated using daily returns within the month. The table reports the time-series averages of the coefficients (multiplied by 100) together with their Newey-West corrected t-statistics (using three lags) within parentheses. Panel B presents joint economic significance of regression coefficients on lagged returns. We first predict buys and sells for each month using the following prediction equations:

$$Y_{i,t}^{\text{predicted,Returns}} = \bar{a} + \sum_{j=1}^{12} \bar{b}_j R_{i,t-j},$$

where bars of top of the coefficients denote average of the time-series coefficients from the above regression. For each month, we rank stocks based on the predicted Ys and form deciles based on this ranking. We judge the economic significance of predictability based on each set of independent variables based on the differences between Ys of the extreme deciles in the following month. The numbers within square brackets in Panel B are the differences divided by corresponding standard deviations. The sample includes all NYSE stocks over the period 1993 to 2010.

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys – Sells	Buys	−Sells	Buys − Sells
			Panel A: Regression coefficients			
Constant	42.55 (10.64)	−45.46 (−14.76)	−2.92 (−1.38)	34.92 (8.91)	−53.82 (−16.32)	−18.90 (−7.84)
Return(−1)	12.13 (4.36)	−18.81 (−8.60)	−6.68 (−3.54)	0.85 (0.35)	0.12 (0.05)	0.97 (0.94)
Return(−2)	16.17 (7.56)	−18.57 (−13.01)	−2.40 (−2.12)	7.01 (4.54)	−6.01 (−4.00)	1.00 (1.34)
Return(−3)	18.87 (10.61)	−18.76 (−14.57)	0.10 (0.14)	12.06 (8.43)	−9.30 (−7.36)	2.76 (3.77)
Return(−4)	18.88 (12.99)	−17.06 (−16.63)	1.82 (2.96)	12.34 (8.96)	−9.85 (−8.26)	2.49 (4.15)
Return(−5)	17.80 (14.19)	−14.86 (−15.30)	2.94 (5.59)	12.76 (9.67)	−9.99 (−7.04)	2.76 (4.09)
Return(−6)	20.26 (13.63)	−16.12 (−13.20)	4.14 (6.78)	14.50 (10.36)	−10.99 (−8.63)	3.51 (5.27)
Return(−7)	16.98 (13.47)	−12.73 (−11.87)	4.25 (7.87)	11.34 (8.66)	−7.85 (−6.39)	3.49 (5.69)
Return(−8)	15.92 (13.46)	−12.08 (−10.80)	3.84 (7.86)	12.31 (9.79)	−9.95 (−8.08)	2.36 (4.89)
Return(−9)	14.34 (12.63)	−10.82 (−9.86)	3.52 (6.49)	11.74 (9.75)	−9.02 (−7.40)	2.71 (5.87)
Return(−10)	12.10 (9.46)	−8.90 (−8.03)	3.20 (5.52)	10.13 (7.43)	−8.65 (−6.62)	1.48 (3.07)
Return(−11)	8.90 (7.72)	−5.86 (−5.80)	3.04 (5.41)	6.61 (5.50)	−5.56 (−4.69)	1.05 (1.93)
Return(−12)	9.88 (9.60)	−6.42 (−6.58)	3.46 (6.34)	9.17 (7.90)	−7.06 (−5.69)	2.10 (4.01)
ln(Size(−1))	0.64 (2.37)	−0.34 (−1.42)	0.30 (2.65)	1.02 (3.91)	0.49 (2.07)	1.51 (9.52)
ln(BM(−1))	−1.43 (−5.45)	−0.26 (−1.38)	−1.69 (−8.24)	−0.54 (−2.39)	0.40 (1.75)	−0.14 (−1.70)
Vol(−1)	60.36 (13.64)	−48.05 (−11.75)	12.31 (6.04)	74.98 (18.17)	−64.66 (−17.18)	10.33 (5.70)
avg adj-R^2	11.76	11.93	6.22	6.23	5.09	3.10
avg # stocks	1,329	1,329	1,329	1,329	1,329	1,329
			Panel B: Economic significance			
All return lags	0.216 [0.73]	0.182 [0.76]	0.054 [0.38]	0.133 [0.36]	0.089 [0.25]	0.047 [0.21]

Appendix C: Robustness

A. Large orders

Buys and sells in number of trades equally weight all trades and buys and sells in shares traded weight these measures by the number of shares traded. Hence, shares traded based measures are more representative of large investors while number of trades based measures are more representative of small traders. To more directly examine the impact of large traders, we compute buys and sells using only trades that are larger than $10,000.[3]

Table C1 presents the results for Buys, Sells and Buys−Sells regression using only large orders. The results for Buys-Sells are generally similar to those in Table 2, except that the coefficients on the first two lags of returns for Buys−Sells in shares are not negative for the large trades as they are in Table 2 for all orders. While the regressions with the number of trades suggest the dominance of contrarian trading in the short-run and momentum trading in the long-run, the regressions with shares traded point to the dominance of momentum trading at all horizons, albeit insignificantly so in the short-run.

One concern is about trade cutoff of $10,000 because, in recent years, institutions are increasingly breaking up their orders into small trades. We have examined the pre-decimalization period prior to 2001 when institutions were less likely to break up their trades. The pattern of coefficients on lagged returns for Buys−Sells regressions is quite similar to that in Table C1. When examining trades, the first two lagged return coefficients are negative and the coefficients at longer lags are positive, but in the case of shares traded all the lagged return coefficients are positive. Even the first two lag return coefficients are significantly positive, suggesting that momentum trades overwhelm the contrarian or disposition trades at all lags.

B. Lags of both buys and sells as control variables

The decision to sell shares could be contaminated by short sale constraints. This could lead to an interconnection between current sales and past purchases since investors that are short sale constrained will only be able to sell previously purchased shares. Thus, the decision to sell could be conditional on previous purchase decisions and this could explain why, in the trades regressions, both sells and buys depend positively on short-term lagged returns. To address this we now use lagged buys as well as lagged sells in our Buys and −Sells regressions. Table C2 presents the results.

The coefficient estimates on lagged returns, in general, have the same signs and significance as in Table 2. The only exception is the coefficient on the one month lagged return in the shares traded regression for −Sells which now has a positive and significant coefficient of 7.56 as opposed to an insignificant coefficient of 2.51 in Table 2.

The lagged buy and sell coefficients point to a positive serial correlation in buys/sells as before. In the Buys regressions, the lagged sell coefficients suggest that a decrease in lagged sells is accompanied with an increase in buys. More importantly, as conjectured, the coefficient on the first lag of buys in the −Sells regression is significantly negative suggesting that it is indeed the case that an increase in the previous month's buys leads to an increase in current sells.

---

[3] Chordia, Roll, and Subrahmanyam (2001) use the $10,000 cutoff to delineate large and small traders. In the post-decimalization period, small orders also increasingly come from institutions that break up their total orders into small lots. Although the $10,000 cutoff would miss these trades from institutional traders, the sample of trades above this cutoff are likely to have emanated from large traders, and hence trades by small traders would have a smaller effect on this subsample than in the entire sample.

Panel B presents the economic significance. As in equation (8), we rank stocks based on the predicted Buys or Sells to form the decile rankings and then we judge the economic significance based on the differences between the Buys and Sells of the extreme deciles in the following month. The economic significance is very similar to that in Table 2. For the returns-based prediction, the standardized difference for the number of Buys (Sells) is 0.63 (0.74). For the lagged buys and sell based prediction the standardized difference in terms of the number of shares bought (sold) is 1.29 (1.11).

C. Effect of returns at longer lags

The regression in equation (10) in the paper imposes a linear relation between returns versus buys and sells. However, this relation may be different for positive and negative returns as suggested by Ben-David and Hirshleifer (2012) and Sialm and Starks (2012) for retail and institutional investors, respectively.[4] We now separate the past cumulative returns into positive and negative returns. For instance, Max(Return(−1, −3), 0) refers to positive returns and Min(Return(−1, −3), 0) refers to negative returns over the period $t-1$ to $t-3$. Table C3 presents the results.

The relation between past returns and buys and sells are different in the positive and negative segments over the one- to three-month horizon. For example, for number of Buys, the slope coefficient for positive returns is significantly positive but significantly negative for negative returns. The results indicate that investors are more likely to buy big winners and also big losers. We find a similar V-shaped relation in all regressions for the one- to three-month horizon. Ben-David and Hirshleifer (2012) report that individual investors' buy and sell decision also exhibit a similar V-shaped relation with past returns as we find for aggregate trading decisions.

For the four- to 12-month horizon all slope coefficients have the same sign. For the 13- to 36-month horizon, we find some differences between the coefficients for positive and negative returns but all coefficients with a sign different from the corresponding coefficient in Panel A are insignificant. For the longest lags, none of the slope coefficients are significant. Overall, for longer lags, we do not the find distinct V-shaped relation that we find for the one- to three-month horizon.

D. Subsample analysis

We provide results for different sub-periods in Table C4. We first examine the results for sub-periods demarcated by decimalization from 1993 to 2001 and from 2002 to 2010. The results are qualitatively similar in both sub-periods. The only differences pertain to the slope coefficients on returns at shorter lags. The coefficients on first two lags of returns for the Buys regressions are strongly negative during the first sub-period but only the coefficient on the first lag of returns for the Buys regressions in shares is negative in the second sub-period. This suggests that contrarian trading has become weaker in second half of the sample period. The other main difference is the strong positive coefficient on the first lag of returns in shares regression in the first sub-period. This is inconsistent with either contrarian trading or the disposition effect. The results in the first sub-period are generally stronger than in the second sub-period. In the internet appendix A we show that the slope coefficients decrease with an

[4] Ben-David and Hirshleifer (2012) examine a sample of retail trades to document a discontinuity in trading behavior around a return of zero. Sialm and Starks (2012) argue that mutual funds with different tax clienteles behave differently around the zero return in terms of their capital gain realization behavior.

increase in the relative strength of supplementary signals and in the number of traders who use signals other than past returns in their trading decisions. Therefore, the results here indicate that these factors were more important in the second sub-period than in the first sub-period.

We also consider the crisis period separately by breaking up the second sub-period above into the period January 2002-June 2007 and July 2007-December 2010. The latter period also coincides with an increase in high frequency trading. Once again the main differences in the two sub-periods obtain in the impact of the short-term lagged returns. The evidence of short-term contrarian buying during January 2002-June 2007 is somewhat weaker but during the recent financial crisis there is strong evidence of short-term contrarian buying in the Buys regressions with shares traded. Also, in the −Sells regression with shares traded, there is some evidence of disposition / contrarian selling of winners during January 2002-June 2007 but during the recent financial crisis in the period July 2007-December 2010 there is strong evidence of selling of losers.

We urge some caution in interpreting the subperiod results because the subperiods are relatively short and hence the coefficients are estimates less precisely. Additionally, the Lee-Ready algorithm may not be as accurate at classifying the active side of each trade during recent years because high frequency traders use sophisticated algorithms to mask active trades. In spite of these potential issues, our results are generally similar across subperiods.

Table C1: Cross-sectional determinants of buys and sells for large orders in NYSE stocks

We run the following cross-sectional regression each month:

$$Y_{i,t} = a_t + \sum_{j=1}^{12} b_{t,j} R_{i,t-j} + c_t \ln\left(Size_{i,t-1}\right) + d_t \ln\left(BM_{i,t-1}\right) + e_t Vol_{i,t-1} + \sum_{j=1}^{3} f_{t,j} Y_{i,t-j} + u_{i,t},$$

where Y is either Buys, or −Sells, or Buys−Sells. We measure Buys and Sells both in number of trades and number of shares. We scale these variables by the average number of trades / trading volume over the previous 12 months (i.e. months t−1 to t−12). We use only trades that are larger than $10,000 in these computations. The slope coefficients for Buys−Sells regression are not the sum of the corresponding slope coefficients for Buys and −Sells regressions because the independent variables include respective lagged dependent variables, which differ across the three specifications. Return is in percent per month, Size is market capitalization in millions of dollars, BM is book-to-market, and Vol is the standard deviation of returns in percent per month calculated using daily returns within the month. The table reports the time-series averages of the coefficients (multiplied by 100) together with their Newey-West corrected t-statistics (using three lags) within parentheses. Panel B presents joint economic significance of regression coefficients on lagged returns and lagged dependent variables. We first predict buys and sells for each month using the following prediction equations:

$$Y_{i,t}^{\text{predicted,Returns}} = \bar{a} + \sum_{j=1}^{12} \bar{b}_j R_{i,t-j} \text{ and } Y_{i,t}^{\text{predicted,Lags}} = \bar{a} + \sum_{j=1}^{3} \bar{f}_j Y_{i,t-j},$$

where bars of top of the coefficients denote average of the time-series coefficients from the above regression. For each month, we rank stocks based on the predicted Ys and form deciles based on this ranking. We judge the economic significance of predictability based on each set of independent variables based on the differences between Ys of the extreme deciles in the following month. The numbers within square brackets in Panel B are the differences divided by corresponding standard deviations. The sample includes all NYSE stocks over the period 1993 to 2010.

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys − Sells	Buys	−Sells	Buys − Sells
Panel A: Regression coefficients						
Constant	35.32 (12.39)	−38.37 (−15.91)	−3.12 (−2.37)	36.56 (10.43)	−51.44 (−16.96)	−16.54 (−6.55)
Return(−1)	31.40 (9.52)	−44.14 (−15.36)	−6.67 (−5.58)	3.79 (1.16)	−7.57 (−2.11)	1.05 (0.72)
Return(−2)	20.36 (8.94)	−25.61 (−13.93)	−2.27 (−3.25)	5.93 (3.91)	−7.32 (−4.33)	0.94 (0.85)
Return(−3)	22.83 (11.87)	−24.74 (−14.15)	0.14 (0.22)	11.86 (8.06)	−11.15 (−7.64)	1.94 (1.78)
Return(−4)	18.09 (10.62)	−18.88 (−12.93)	2.01 (3.35)	10.75 (6.92)	−10.10 (−6.34)	2.54 (2.95)
Return(−5)	14.13 (8.96)	−14.47 (−10.72)	2.24 (3.82)	13.50 (4.99)	−12.14 (−3.85)	2.70 (2.31)
Return(−6)	17.47 (10.48)	−16.36 (−10.93)	3.42 (6.55)	12.08 (8.45)	−9.52 (−6.83)	4.11 (4.39)
Return(−7)	10.41 (8.97)	−10.22 (−9.09)	2.41 (4.98)	6.50 (3.84)	−5.01 (−2.46)	2.94 (3.01)
Return(−8)	10.27 (7.39)	−10.25 (−7.74)	1.79 (3.59)	9.67 (6.43)	−10.83 (−7.30)	0.09 (0.08)
Return(−9)	7.46 (5.53)	−7.76 (−5.47)	1.40 (3.16)	7.78 (5.32)	−6.04 (−3.86)	2.52 (2.66)
Return(−10)	5.38 (4.72)	−5.17 (−4.46)	1.50 (3.14)	6.73 (5.82)	−6.49 (−4.63)	1.11 (1.40)
Return(−11)	1.18 (1.08)	−1.31 (−1.23)	1.03 (1.90)	3.40 (2.73)	−3.28 (−2.88)	1.12 (1.29)
Return(−12)	2.03 (1.89)	−0.96 (−0.89)	1.69 (3.88)	6.10 (5.37)	−4.09 (−3.29)	2.66 (3.45)
ln(Size(−1))	−0.40 (−2.14)	0.55 (3.34)	0.31 (3.81)	−0.01 (−0.05)	1.05 (5.33)	1.33 (8.04)
ln(BM(−1))	−0.51 (−2.95)	−0.12 (−0.78)	−0.72 (−8.11)	−0.26 (−1.20)	0.26 (1.11)	−0.01 (−0.12)
Vol (−1)	−16.05 (−4.08)	9.31 (2.64)	3.26 (2.15)	8.05 (2.00)	−15.52 (−3.67)	3.26 (1.42)
Y(−1)	35.42 (35.79)	31.12 (32.13)	17.49 (16.78)	21.97 (33.34)	17.64 (26.82)	8.21 (9.38)
Y(−2)	8.30 (9.44)	6.96 (9.13)	8.61 (10.51)	6.99 (13.63)	5.66 (11.20)	4.72 (6.81)
Y(−3)	4.79 (10.67)	4.53 (10.08)	6.74 (11.39)	5.19 (15.07)	4.50 (12.94)	4.59 (7.61)
avg adj-R^2	27.97	25.56	10.52	10.70	7.36	4.14
avg # stocks	1,293	1,293	1,293	1,293	1,293	1,293
Panel B: Economic significance						
All return lags	0.471 [1.11]	0.436 [1.14]	0.028 [0.15]	0.195 [0.40]	0.143 [0.29]	0.054 [0.14]
All Y lags	0.633 [1.49]	0.543 [1.41]	0.172 [0.91]	0.443 [0.92]	0.370 [0.75]	0.177 [0.47]

Table C2: Cross-sectional determinants of buys and sells for NYSE stocks with lags of buys and sells

We run the following cross-sectional regression each month:

$$Y_{i,t} = a_t + \sum_{j=1}^{12} b_{t,j} R_{i,t-j} + c_t \ln\left(Size_{i,t-1}\right) + d_t \ln\left(BM_{i,t-1}\right) + e_t Vol_{i,t-1} + \sum_{j=1}^{3} f_{t,j} Buys_{i,t-j} - \sum_{j=1}^{3} g_{t,j} Sells_{i,t-j} + u_{i,t},$$

where Y is either Buys or −Sells. We measure Buys and Sells both in number of trades and number of shares. We scale these variables by the average number of trades / trading volume over the previous 12 months (i.e. months $t-1$ to $t-12$). We use only trades that are larger than \$10,000 in these computations. Return is in percent per month, Size is market capitalization in millions of dollars, BM is book-to-market, and Vol is the standard deviation of returns in percent per month calculated using daily returns within the month. The table reports the time-series averages of the coefficients (multiplied by 100) together with their Newey-West corrected *t*-statistics (using three lags) within parentheses. Panel B presents joint economic significance of regression coefficients on lagged returns and lagged dependent variables. We first predict buys and sells for each month using the following prediction equations:

$$Y_{i,t}^{\text{predicted,Returns}} = \bar{a} + \sum_{j=1}^{12} \bar{b}_j R_{i,t-j} \text{ and } Y_{i,t}^{\text{predicted,Lags}} = \bar{a} + \sum_{j=1}^{3} \bar{f}_j Buys_{i,t-j} - \sum_{j=1}^{3} \bar{g}_j Sells_{i,t-j},,$$

where bars of top of the coefficients denote average of the time-series coefficients from the above regression. For each month, we rank stocks based on the predicted Ys and form deciles based on this ranking. We judge the economic significance of predictability based on each set of independent variables based on the differences between *Y*s of the extreme deciles in the following month. The numbers within square brackets in Panel B are the differences divided by corresponding standard deviations. The sample includes all NYSE stocks over the period 1993 to 2010.

	Numbers of trades		Shares traded	
	Buys	−Sells	Buys	−Sells
Panel A: Regression coefficients				
Constant	31.33 (12.71)	−31.11 (−15.68)	33.57 (11.47)	−45.99 (−17.88)
Return(−1)	−3.46 (−1.61)	−7.60 (−5.82)	−11.43 (−6.46)	7.56 (3.95)
Return(−2)	3.08 (2.53)	−5.59 (−6.51)	−2.05 (−1.72)	−0.33 (−0.27)
Return(−3)	7.25 (7.59)	−7.33 (−9.10)	3.54 (3.39)	−4.39 (−4.32)
Return(−4)	7.67 (7.35)	−5.47 (−7.12)	6.27 (5.31)	−5.02 (−4.58)
Return(−5)	6.12 (7.11)	−3.79 (−5.72)	6.01 (5.69)	−4.33 (−3.42)
Return(−6)	9.07 (8.58)	−6.32 (−7.60)	7.94 (7.21)	−5.49 (−5.66)
Return(−7)	5.29 (5.56)	−3.13 (−4.02)	4.37 (4.24)	−2.24 (−2.16)
Return(−8)	5.17 (5.96)	−3.81 (−5.01)	6.40 (6.32)	−5.39 (−5.51)
Return(−9)	4.41 (6.22)	−3.28 (−4.92)	5.89 (6.79)	−4.22 (−4.82)
Return(−10)	3.54 (4.09)	−2.38 (−3.14)	4.95 (4.58)	−4.54 (−4.44)
Return(−11)	2.06 (2.71)	−0.78 (−1.16)	2.29 (2.44)	−1.92 (−2.20)
Return(−12)	3.89 (4.94)	−2.29 (−3.32)	5.36 (5.46)	−3.99 (−3.83)
ln(Size(−1))	−0.28 (−1.70)	0.33 (2.41)	−0.13 (−0.70)	1.08 (6.64)
ln(BM(−1))	−0.41 (−2.74)	−0.28 (−2.46)	−0.38 (−2.41)	0.37 (2.16)
Vol (−1)	−25.91 (−8.89)	20.32 (9.17)	−10.21 (−3.11)	8.52 (2.89)
Buys(−1)	53.56 (25.83)	−20.72 (−11.27)	34.70 (23.20)	−17.34 (−12.14)
Buys(−2)	7.53 (5.12)	1.96 (1.80)	7.71 (6.57)	0.03 (0.03)
Buys(−3)	6.61 (7.49)	2.48 (2.88)	8.22 (15.22)	−0.37 (−0.46)
−Sells(−1)	7.52 (4.05)	23.20 (16.35)	0.27 (0.18)	15.54 (11.30)
−Sells(−2)	3.86 (2.87)	6.81 (6.12)	2.52 (2.27)	5.21 (5.51)
−Sells(−3)	0.86 (0.90)	8.31 (8.72)	2.80 (4.75)	4.72 (6.10)
avg adj-R^2	35.75	32.53	19.53	16.35
avg # stocks	1,304	1,304	1,304	1,304

Panel B: Economic significance				
All return lags	0.187 [0.63]	0.176 [0.74]	0.101 [0.28]	0.080 [0.22]
All Buys/Sells lags	0.530 [1.79]	0.392 [1.65]	0.473 [1.29]	0.396 [1.11]

Table C3: Cross-sectional determinants of buys and sells for NYSE stocks: Effect of Returns at longer lags

We run the following cross-sectional regression each month:

$$Y_{i,t} = a_t + b_{t,1}R_{i,t-1:t-3} + b_{t,2}R_{i,t-4:t-12} + b_{t,3}R_{i,t-13:t-36} + b_{t,4}R_{i,t-37:t-60} + c_t \ln\left(Size_{i,t-1}\right) + d_t \ln\left(BM_{i,t-1}\right) + e_t Vol_{i,t-1} + \sum_{j=1}^{60} f_{t,j} Y_{i,t-j} + u_{i,t},$$

where Y is either Buys, or −Sells, or Buys−Sells. We measure Buys and Sells both in number of trades and number of shares. We scale these variables by the average number of trades / trading volume over the previous 12 months (i.e. months t−1 to t−12). The slope coefficients for Buys−Sells regression are not the sum of the corresponding slope coefficients for Buys and −Sells regressions because the independent variables include respective lagged dependent variables, which differ across the three specifications. Returns (separated into positive and negative) are in percent per month, Size is market capitalization in millions of dollars, BM is book-to-market, and Vol is the standard deviation of returns in percent per month calculated using daily returns within the month. The table reports the time-series averages of the coefficients (multiplied by 100) together with their Newey-West corrected t-statistics (using 60 lags) within parentheses. Coefficients on lags greater than one of the dependent variable are not shown for simplicity. The sample includes all NYSE stocks over the period 1993 to 2010.

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys − Sells	Buys	−Sells	Buys − Sells
Constant	33.93 (8.44)	−35.66 (−7.43)	3.31 (1.70)	39.59 (10.56)	−49.81 (−7.14)	−1.01 (−0.46)
Max(Return(−1: −3), 0)	13.79 (5.38)	−15.12 (−6.23)	1.70 (2.03)	16.00 (4.82)	−16.16 (−6.41)	1.73 (1.21)
Min(Return(−1: −3), 0)	−8.73 (−4.11)	3.46 (2.35)	−4.76 (−2.54)	−25.84 (−5.81)	19.65 (6.52)	−6.41 (−2.90)
Max(Return(−4: −12), 0)	3.09 (5.50)	−2.77 (−6.14)	0.43 (1.04)	2.98 (4.61)	−1.86 (−3.47)	0.15 (0.43)
Min(Return(−4: −12), 0)	7.52 (5.61)	−8.23 (−5.78)	2.25 (2.94)	9.09 (3.89)	−9.67 (−4.18)	2.68 (2.68)
Max(Return(−13: −36), 0)	−0.06 (−0.40)	0.17 (1.50)	−0.02 (−0.27)	0.24 (1.85)	0.06 (0.43)	−0.07 (−1.10)
Min(Return(−13: −36), 0)	2.08 (3.59)	−2.89 (−4.74)	0.64 (1.82)	5.08 (4.68)	−5.77 (−4.21)	0.71 (1.65)
Max(Return(−37: −60), 0)	0.08 (0.64)	−0.05 (−0.57)	0.00 (0.02)	0.14 (0.73)	0.08 (0.63)	0.02 (0.27)
Min(Return(−37: −60), 0)	0.13 (0.20)	−0.26 (−0.50)	0.12 (0.35)	1.65 (1.51)	−2.08 (−1.76)	1.41 (1.62)
ln(Size(−1))	−0.11 (−0.42)	−0.02 (−0.05)	−0.10 (−1.15)	−0.06 (−0.34)	0.69 (2.06)	0.19 (1.11)
ln(BM(−1))	−0.24 (−0.83)	−0.12 (−0.52)	−0.11 (−2.76)	−0.02 (−0.06)	0.10 (0.23)	0.12 (2.14)
Vol(−1)	−30.94 (−5.47)	18.55 (5.60)	−7.07 (−4.84)	−16.89 (−6.03)	7.36 (3.89)	−4.07 (−4.26)
Y(−1)	0.50 (51.82)	0.46 (76.41)	0.28 (6.81)	0.35 (20.27)	0.32 (11.97)	0.16 (5.41)
Y(−2) to Y(−60)			Not reported for brevity			
avg adj-R^2	43.22	41.06	29.71	25.51	22.97	19.50
avg # stocks	880	880	880	880	880	880

Table C4: Cross-sectional determinants of buys and sells for NYSE stocks: Subsample analysis

We run the following cross-sectional regression each month:

$$Y_{i,t} = a_t + \sum_{j=1}^{12} b_{t,j} R_{i,t-j} + c_t \ln\left(Size_{i,t-1}\right) + d_t \ln\left(BM_{i,t-1}\right) + e_t Vol_{i,t-1} + \sum_{j=1}^{3} f_{t,j} Y_{i,t-j} + u_{i,t},$$

where Y is either Buys, or −Sells, or Buys−Sells. We measure Buys and Sells both in number of trades and number of shares. We scale these variables by the average number of trades / trading volume over the previous 12 months (i.e. months $t-1$ to $t-12$). The slope coefficients for Buys−Sells regression are not the sum of the corresponding slope coefficients for Buys and −Sells regressions because the independent variables include respective lagged dependent variables, which differ across the three specifications. Return is in percent per month, Size is market capitalization in millions of dollars, BM is book-to-market, and Vol is the standard deviation of returns in percent per month calculated using daily returns within the month. The table reports the time-series averages of the coefficients (multiplied by 100) together with their Newey-West corrected t-statistics (using three lags) within parentheses. The sample includes all NYSE stocks over the period 1993 to 2010. The three panels show the results for three different subsamples.

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys − Sells	Buys	−Sells	Buys − Sells
			Panel A: Sample January 1993 to December 2001			
Constant	27.03 (10.67)	−33.59 (−16.94)	−4.28 (−3.38)	23.80 (9.45)	−46.47 (−18.25)	−24.43 (−16.57)
Return(−1)	−15.23 (−5.33)	−10.22 (−4.45)	−21.54 (−9.78)	−17.19 (−7.05)	7.03 (2.92)	−2.06 (−1.51)
Return(−2)	−4.26 (−2.45)	−1.49 (−1.28)	−4.41 (−4.46)	−4.05 (−2.46)	1.10 (0.65)	−0.26 (−0.25)
Return(−3)	1.95 (1.60)	−3.78 (−3.65)	0.03 (0.04)	3.02 (2.37)	−3.85 (−3.35)	0.56 (0.58)
Return(−4)	6.69 (3.82)	−5.43 (−4.02)	4.13 (4.89)	8.41 (4.23)	−8.18 (−4.36)	2.24 (2.60)
Return(−5)	4.74 (3.54)	−3.41 (−3.37)	3.80 (5.05)	6.64 (4.90)	−5.62 (−3.60)	2.90 (3.69)
Return(−6)	8.31 (4.66)	−5.95 (−4.47)	4.41 (5.86)	9.12 (6.27)	−7.01 (−5.17)	4.10 (4.46)
Return(−7)	6.08 (4.23)	−4.06 (−3.41)	3.83 (5.31)	6.42 (4.22)	−5.32 (−3.86)	3.10 (3.81)
Return(−8)	4.37 (3.62)	−3.61 (−3.07)	2.34 (3.91)	6.02 (4.29)	−6.70 (−4.79)	1.09 (1.37)
Return(−9)	5.22 (5.44)	−4.02 (−4.18)	2.67 (3.80)	8.15 (6.94)	−6.54 (−4.43)	2.88 (4.23)
Return(−10)	3.64 (3.61)	−2.84 (−2.66)	2.17 (3.25)	5.48 (4.60)	−6.54 (−4.86)	0.52 (0.90)
Return(−11)	2.48 (2.20)	−1.30 (−1.46)	2.33 (3.51)	2.99 (2.29)	−3.77 (−2.93)	0.67 (0.68)
Return(−12)	4.39 (3.34)	−2.04 (−1.76)	3.38 (6.10)	6.62 (4.33)	−5.41 (−2.98)	2.43 (2.90)
ln(Size(−1))	−0.23 (−1.43)	0.39 (3.15)	0.31 (3.81)	0.51 (2.97)	0.86 (5.40)	1.78 (18.34)
ln(BM(−1))	−1.50 (−8.81)	0.21 (1.86)	−1.25 (−9.04)	−1.03 (−7.26)	1.19 (7.24)	0.09 (0.64)
Vol(−1)	−25.84 (−6.27)	14.68 (4.83)	0.05 (0.02)	−0.07 (−0.02)	−7.58 (−1.97)	7.02 (2.80)
Y(−1)	47.27 (24.17)	39.09 (21.35)	36.15 (40.40)	31.46 (24.11)	24.00 (21.26)	14.27 (15.04)
Y(−2)	7.23 (5.39)	5.83 (5.37)	11.50 (13.80)	6.00 (7.98)	5.13 (7.34)	6.47 (11.25)
Y(−3)	5.72 (9.51)	4.81 (8.88)	9.78 (14.95)	5.85 (11.80)	4.15 (8.45)	7.66 (12.37)
avg adj-R^2	32.37	25.54	28.15	16.79	10.50	7.75
avg # stocks	1,376	1,376	1,376	1,376	1,376	1,376

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys − Sells	Buys	−Sells	Buys − Sells
Panel B: Sample January 2002 to June 2007						
Constant	40.10 (8.82)	−31.31 (−8.35)	8.91 (6.07)	48.56 (9.58)	−50.57 (−12.84)	−1.12 (−0.48)
Return(−1)	11.88 (5.40)	−16.50 (−7.58)	−0.56 (−0.72)	−2.24 (−0.68)	−6.91 (−1.95)	−6.55 (−5.11)
Return(−2)	8.19 (4.93)	−8.96 (−6.98)	−0.40 (−0.50)	0.35 (0.16)	−4.19 (−1.93)	−4.49 (−5.13)
Return(−3)	10.74 (6.96)	−10.62 (−8.55)	−0.42 (−0.75)	6.51 (3.12)	−7.68 (−4.07)	−2.31 (−2.28)
Return(−4)	7.25 (4.81)	−6.63 (−5.45)	1.28 (1.97)	5.93 (3.57)	−4.62 (−3.01)	1.78 (1.93)
Return(−5)	7.03 (5.31)	−6.02 (−5.55)	1.44 (2.52)	8.41 (3.92)	−7.67 (−2.64)	0.98 (0.71)
Return(−6)	10.89 (7.19)	−9.38 (−7.43)	1.97 (2.57)	10.54 (5.90)	−8.99 (−5.45)	2.37 (2.15)
Return(−7)	5.10 (3.76)	−5.14 (−4.31)	0.97 (1.77)	4.01 (2.21)	−2.63 (−1.19)	2.24 (1.70)
Return(−8)	5.42 (3.69)	−5.46 (−4.36)	0.34 (0.69)	7.77 (3.98)	−6.79 (−3.23)	1.23 (1.51)
Return(−9)	3.97 (3.03)	−4.62 (−3.58)	−0.25 (−0.43)	4.30 (2.72)	−4.25 (−2.84)	0.48 (0.64)
Return(−10)	4.00 (2.07)	−4.11 (−2.44)	0.36 (0.63)	7.22 (2.98)	−6.28 (−2.71)	1.43 (2.15)
Return(−11)	1.90 (1.43)	−1.77 (−1.33)	0.48 (1.16)	2.75 (1.59)	−1.79 (−1.18)	0.95 (1.20)
Return(−12)	2.82 (2.63)	−2.82 (−3.46)	0.18 (0.41)	4.49 (3.21)	−2.82 (−2.06)	1.64 (2.42)
ln(Size(−1))	−0.78 (−2.45)	0.35 (1.31)	−0.33 (−3.41)	−0.90 (−2.61)	1.30 (4.92)	0.42 (2.61)
ln(BM(−1))	0.06 (0.21)	−0.39 (−1.54)	−0.37 (−5.58)	0.03 (0.11)	−0.25 (−0.82)	−0.20 (−2.62)
Vol(−1)	−41.11 (−8.91)	20.97 (5.00)	−11.48 (−7.05)	−22.73 (−3.90)	7.11 (1.01)	−3.37 (−1.51)
Y(−1)	52.63 (22.97)	47.65 (29.43)	38.01 (21.61)	35.17 (28.31)	30.81 (20.52)	25.60 (18.04)
Y(−2)	2.98 (2.27)	5.12 (3.78)	12.35 (8.89)	5.27 (7.87)	5.81 (6.62)	11.84 (9.41)
Y(−3)	6.16 (5.58)	5.09 (4.77)	11.90 (12.66)	5.30 (9.81)	4.82 (7.26)	10.53 (11.39)
avg adj-R^2	36.45	33.95	30.48	17.60	15.66	15.59
avg # stocks	1,279	1,279	1,279	1,279	1,279	1,279

	Numbers of trades			Shares traded		
	Buys	−Sells	Buys − Sells	Buys	−Sells	Buys − Sells
Panel C: Sample July 2007 to December 2010						
Constant	17.65 (3.60)	−19.67 (−4.17)	−3.13 (−3.53)	22.87 (4.00)	−24.16 (−4.43)	−2.88 (−2.27)
Return(−1)	−2.23 (−1.09)	−2.79 (−1.31)	−2.64 (−6.44)	−12.18 (−4.64)	7.31 (2.71)	−2.89 (−4.87)
Return(−2)	4.85 (2.92)	−5.40 (−3.42)	−1.01 (−3.39)	0.37 (0.18)	−1.31 (−0.67)	−1.92 (−3.31)
Return(−3)	5.41 (3.12)	−5.83 (−3.66)	−0.23 (−0.67)	3.54 (1.62)	−3.22 (−1.65)	0.31 (0.49)
Return(−4)	3.55 (3.55)	−3.64 (−4.03)	−0.13 (−0.31)	1.57 (1.03)	−1.77 (−1.14)	−0.24 (−0.44)
Return(−5)	2.79 (2.19)	−2.06 (−1.53)	0.69 (1.78)	1.63 (1.20)	−0.39 (−0.26)	0.92 (1.67)
Return(−6)	4.37 (3.11)	−4.31 (−3.50)	0.50 (2.34)	1.18 (0.55)	−1.51 (−0.79)	0.12 (0.12)
Return(−7)	0.96 (0.80)	−0.17 (−0.14)	1.07 (3.68)	0.40 (0.28)	0.53 (0.38)	1.15 (4.14)
Return(−8)	4.93 (2.82)	−4.17 (−2.75)	1.43 (3.38)	5.21 (2.82)	−4.36 (−2.57)	1.54 (3.01)
Return(−9)	1.20 (1.84)	−1.46 (−1.86)	0.40 (1.16)	3.14 (2.16)	−2.06 (−1.44)	1.68 (2.71)
Return(−10)	0.40 (0.38)	−0.41 (−0.43)	0.26 (0.85)	0.06 (0.03)	−0.99 (−0.67)	−0.28 (−0.31)
Return(−11)	0.17 (0.15)	−0.06 (−0.05)	0.57 (2.20)	−0.20 (−0.13)	−0.40 (−0.24)	−0.11 (−0.15)
Return(−12)	3.64 (3.23)	−3.41 (−2.81)	0.57 (2.29)	4.24 (2.41)	−4.61 (−2.72)	−0.24 (−0.31)
ln(Size(−1))	0.33 (0.80)	−0.22 (−0.56)	0.17 (3.08)	0.14 (0.33)	−0.12 (−0.29)	0.16 (2.35)
ln(BM(−1))	0.14 (0.68)	−0.14 (−0.70)	−0.02 (−0.50)	0.30 (1.25)	−0.34 (−1.35)	−0.07 (−1.13)
Vol(−1)	−15.04 (−3.69)	13.97 (3.52)	−0.78 (−0.76)	−6.74 (−1.44)	1.54 (0.32)	−3.09 (−1.86)
Y(−1)	49.75 (22.35)	48.88 (24.55)	13.98 (10.07)	39.96 (18.34)	37.92 (19.36)	7.37 (3.95)
Y(−2)	2.97 (1.87)	3.53 (2.92)	1.52 (0.91)	5.75 (4.28)	6.44 (6.51)	2.62 (1.02)
Y(−3)	8.02 (8.89)	8.07 (8.74)	2.46 (2.44)	7.48 (8.30)	7.17 (7.38)	2.37 (1.78)
avg adj-R^2	38.62	37.73	6.78	25.53	23.94	4.29
avg # stocks	1,182	1,182	1,182	1,182	1,182	1,182

References

Ben-David, I., and D. Hirshleifer. “Are Investors Really Reluctant to Realize Their Losses? Trading Responses to Past Returns and the Disposition Effect.” *Review of Financial Studies*, 25 (2012), 2485–2532.

Brennan, M. J., and H. Cao. “International Portfolio Investment Flows.” *Journal of Finance*, 52 (1997), 1851–1880.

Chordia, T.; R. Roll; and A. Subrahmanyam. “Market Liquidity and Trading Activity.” *Journal of Finance*, 56 (2001), 501–530.

Sialm, C. and L. Starks, “Mutual Fund Tax Clienteles,” *Journal of Finance*, 67 (2012), 1397–1422.