March 1, 2006

Lagged Order Flows and Returns: A Longer-Term Perspective

Avanidhar Subrahmanyam*

*The Anderson School, University of California at Los Angeles. I thank Tarun Chordia, Laura Frieder, Sheridan Titman, and Ross Valkanov for helpful thoughts on the topic.
Abstract

Lagged Order Flows and Returns: A Longer-Term Perspective

In this paper, we consider the large-sample relation between returns and lagged order flows over horizons of up to two months. The analysis is motivated by work in market microstructure which suggests that the effects of inventory control on stock returns should be discernible over horizons longer than those considered in the literature. We begin our analysis by developing a simple model of inventory effects in the presence of public information. Using mid-quote return data, we then find some evidence of return predictability using order flows, even after controlling for lagged returns, which is consistent with our theoretical setting. The relation is present only for negative imbalances and is stronger in large firms rather than small ones. Overall, the analysis is consistent with the notion that inventory control effects span several weeks.
1 Introduction

Recent years have witnessed an increase in interest on the relation between order flows and both contemporaneous as well as lagged returns (see, for example, Chan and Fong, 2000, and Chordia, Roll, and Subrahmanyam, 2002, 2005). Part of the interest has likely been driven by the theoretical link between returns and order imbalances, manifested in well-known models of market microstructure. For example, the well-known Kyle (1985) model of price formation relates price changes to net (pooled) order flow. It can be argued that the Kyle setting is more naturally applicable in the context of signed order imbalances over a time interval, as opposed to trade-by-trade data, since the theory is not one of sequential trades by individual traders. Similarly, the dynamic inventory models of Ho and Stoll (1983) and Spiegel and Subrahmanyam (1995) also study how market makers accommodate buying and selling pressures from outside investors.

Evidence on the imbalance-return relation indicates that lagged order flows are positively related to future returns (viz, Chordia and Subrahmanyam, 2004, Chordia, Roll, and Subrahmanyam, 2005). This evidence has been attributed to the positive autocorrelations in order flows, which carries over to a predictive relation between returns and lagged order flows when the risk-bearing capacity of market makers is limited. Note that this positive relation is not a steady state possibility when inventory control issues are relevant. As Grossman and Miller (1988) point out, ultimately the phenomenon of inventory control must lead to a negative relation between order flow
and returns, as market makers struggle to offload their excess inventory.

Even though inventory control is a very plausible phenomenon, unambiguous evidence on this issue has been elusive. For example, Glosten and Harris (1988) George, Kaul, and Nimalendran (1991), and Madhavan and Smidt (1991) all find only weak support for inventory-driven return effects, using short-horizon (daily and intraday) data. It is somewhat of a puzzle as to why such an intuitive phenomenon is so difficult to detect in the data.

In this paper, we consider whether clear evidence of inventory control may be found in the cross-section of expected stock returns at horizons longer than those typically considered in the literature. What horizons for investigation of inventory control are reasonable? Hasbrouck and Sofianos (1993) as well as Madhavan and Smidt (1993) indicate that inventory shocks get reversed rather slowly. In fact, Hasbrouck and Sofianos (1993) suggest that inventories are persistent for up to two months, whereas Madhavan and Smidt (1993) document that it takes about 47 days for inventories to be reduced by 50%. Both papers rationalize these results by arguing that working off an inventory imbalance may require a substantially long time-period, ostensibly because of costs involved in monitoring and reacting to inventory imbalances. The papers, using relatively modest samples,¹ suggest that the appropriate horizon at which evidence of inventory control may be found is likely to be longer than days or even weeks. Based on this observation, we consider the large-sample relation between monthly returns

¹Hasbrouck and Sofianos (1993) use data on about 138 stocks while Madhavan and Smidt (1993) have 16 stocks in their sample. The studies span two- and one-year time-periods, respectively.
and lagged order flows at horizons of one and two months.

To perform our analysis, we estimate daily order imbalances for each of a comprehensive sample of NYSE stocks for the period 1988 through 2002. Using data from the Institute for the Study of Security Markets (1988 to 1992) and the Trades and Automated Quotations database provided by the NYSE (1993 through 2002), we sign trades in each stock in our sample using the Lee and Ready (1991) algorithm. We then calculate measures of the monthly order imbalance in each stock using the dollar quantity bought or sold. In the end, we have measures of the monthly order imbalance for each company in our sample.

While analyzing the imbalance-return relation, we are aware that bid-ask bounce in daily returns (Blume and Stambaugh, 1983) is particularly relevant to our study. This is because a high buy order imbalance, for example, would imply a preponderance of trades on the ask side of the market, which would naturally contaminate any attempt to relate the future return to a given day’s order imbalance. We address this issue by relating imbalances to a set of returns calculated from close-to-close bid-ask mid-points. In particular, we pass through the entire transactions database to calculate, for each stock, the mid-point of the quoted bid and ask prices corresponding to the last transaction of each day. We then calculate returns for each stock using the mid-point of the bid and ask prices. Throughout our empirical work, we focus these mid-point return series.

2The data extends the sample of Chordia, Roll, and Subrahmanyam (2002) by four more years.
We motivate our empirical study by a simple model of how prices react to imbalances when market makers have inventory concerns. Our framework explicitly examines how risk averse market makers with inventory concerns accommodate trader demands in the presence of public as well as private information. We find that while in a univariate regression of returns on lagged returns, and in turn, lagged imbalances, one obtains a negative coefficient on either variable, due to inventory-based price reversals. Interestingly, however, in a multiple regression of returns on lagged imbalances and lagged returns, one gets a zero coefficient on lagged returns but a negative coefficient on lagged imbalances. This is because lagged returns incorporate the effect of public information, which is of no help in predicting future returns, since it is completely and rationally impounded into current prices. Lagged imbalances pick up the pure effects of inventory control and thus dominated lagged returns in predicting future returns.

Due to higher specialist participation rates in the smaller capitalization stocks (Hasbrouck and Sofianos, 1993), we would expect to find stronger evidence of inventory control in such stocks. On the other hand, if larger stocks experience greater absolute order imbalances, the reverse could be true. Thus, the prevalence of inventory control could depend on market capitalization. Motivated by these observations, we perform our analysis using three groups of stocks stratified by market capitalization. We find some evidence of return predictability from order flow innovations (estimated using an appropriate time-series model) at horizons of one to two months, especially for stocks other than those belonging to the smallest firm group, suggesting that large
order imbalances in the larger firms cause more inventory problems for market makers.

The predictability of returns from lagged order flows goes beyond the relation between returns and lagged returns. Our results are consistent with our theoretical framework, which suggests that in the presence of public information, lagged order flows are better predictors of future returns than lagged returns. Further, the relation is stronger for negative imbalances than for positive ones. This accords with the notion that short-selling constraints may cause negative sentiment to build up and the accompanying large selling when it crosses a threshold may cause large downward price pressures. Our results survive the Fama and French (1993) factor controls, and are economically significant.

This paper is organized as follows. Section 2 presents a simple theoretical model that motivates our empirical tests. Section 3 describes the data, while Section 4 describes the empirical results. Section 5 concludes.

2 The Theory

Consider a risky security which trades at dates 1 and 2, and pays off a random amount \( \theta + \epsilon \) at date 3, where \( \theta \) and \( \epsilon \) are mutually independent, normally distributed random variables with mean zero and variances \( v_\theta \) and \( v_\epsilon \), respectively. There is a mass unity of risk averse agents (“market makers”) who absorb demand shocks that appear in the market. Each such agent has CARA utility with coefficient \( R \). These agents receive a signal \( \theta + \delta \) at date 2, where \( \delta \) is normally distributed, independent of \( \theta \)
and has variance \( v_\delta \). A liquidity demand shock of \( z \) arrives at the market on date 2; \( z \) is normally distributed with mean 0 and variance \( v_z \), and is independent of all other random variables in the model. Finally, a unit mass of informed agents, also with CARA utility and risk aversion \( R \), observe and trade on a signal, which perfectly reveals \( \theta \). We let \( P_i \) denote the price at date \( i \), with \( P_3 = \theta + \epsilon \).

Given the exponential utility-normal distribution setting, it follows that the demands of the informed agents are given by

\[
x = \frac{\theta - P_2}{R \epsilon},
\]

and those of the market makers are given by

\[
y = \frac{\mu - P_2}{R v},
\]

where \( \mu \equiv E(\theta|\theta + \delta, P_2) \) and \( v \equiv \text{var}(\theta|\theta + \delta, P_2) \). The market clearing condition therefore becomes

\[
\frac{\mu - P_2}{R v} = -\frac{\theta - P_2}{R \epsilon} - z.
\]  

(1)

Note that in equilibrium, from the market makers’ point of view, conditioning on \( P_2 \) is the same as conditioning on \( \tau \equiv \theta + R \epsilon z \). To calculate the equilibrium price, we use the well-known result that if there exist random vectors \( \nu_1 \) and \( \nu_2 \) such that

\[(\nu_1, \nu_2) \sim N \left( (M_1, M_2), \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right) \]

then the conditional distribution of \( \nu_1 \) given \( \nu_2 = U_2 \) is normal with a mean given by the vector

\[E(\nu_1|\nu_2 = U_2) = M_1 + \Sigma_{12} \Sigma_{22}^{-1}(U_2 - M_2),\]

(2)
and variance given by

\[
\text{var}(v_1 | v_2 = U_2) = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}.
\]  

(3)

Using the above result, we have that

\[
\mu = \frac{v_\theta [R \delta v_z + \delta r^2 v_z^2 + \theta (R^2 v_z^2 + v_\theta)]}{R^2 v_\epsilon^2 v_z (v_\delta + v_\theta) + v_\delta v_\theta}
\]

and

\[
v = v_\theta - \frac{v_\theta^2}{v_\delta + v_\theta} - \frac{v_\delta^2 v_\theta^2}{[R^2 v_\epsilon^2 v_z (v_\delta + v_\theta) + v_\delta v_\theta] [v_\delta + v_\theta]}.
\]

Substituting these expressions into the market clearing condition (1) and solving for \(P_2\) yields the following lemma.

**Lemma 1** The equilibrium value of the price at date 2, \(P_2\), is given by

\[
P_2 = a \theta + b z + c \delta,
\]  

(4)

where

\[
a = \frac{v_\theta [R^2 v_\epsilon v_z (v_\delta + v_\epsilon) + v_\delta]}{D},
\]

\[
b = \frac{R v_\delta v_\epsilon (R^2 v_\epsilon v_z + 1)}{D},
\]

\[
c = \frac{R^2 v_\epsilon^2 v_\theta v_z}{D},
\]

and where, in turn,

\[
D \equiv R^2 v_\epsilon v_z [v_\delta (v_\epsilon + v_\theta) + v_\epsilon v_\theta] + v_\delta v_\theta.
\]

The price loads positively on all three random variables including the information \(\theta\), and the liquidity trade as well as the noise in the public signal. It can easily be shown
that the loading on the information is decreasing in the risk aversion coefficient, $R$, as well as the variance of liquidity trading, $v_z$. Note that the date 1 price is nonstochastic, i.e., it does not depend on any of the random variables, $\theta$, $\delta$, $\epsilon$, or $z$.

We view “returns” in this model as synonymous with price changes, as is standard in models that use the exponential-normal setting. We consider $P_3 - P_2$ as the current return and $P_2 - P_1$ as the lagged return. Further, we view the quantity $-x - z$ as the lagged order flow (since it is established prior to the realization of $P_3 - P_2$). One can then derive the following proposition that relates returns to lagged returns as well as lagged order flows (the proof is in Appendix A).

**Proposition 1**

1. In univariate regressions, lagged returns and lagged order flows are both negatively related to future returns.

2. A multiple regression of returns on lagged returns and lagged order flow, yields a coefficient of zero on lagged returns and a negative coefficient of

$$\frac{R^3 v_\delta v_\epsilon^2 v_z}{R^2 v_\epsilon^2 v_z (v_\delta + v_\theta) + v_\delta v_\theta}$$

on order flows.

The intuition of the above result is that returns include the price response to the public signal. Since the reaction to this signal is rational, it offers no predictive power about future returns. Thus, all information about future returns is contained solely in past order flows. Therefore, lagged order imbalances dominate lagged returns in predicting future returns.
We empirically analyze the relation between imbalances and price movements using a comprehensive data set on daily order imbalances which encompasses more than 1100 stocks over the period 1988 to 2002. We use Proposition 1 as the base for our data analysis, though we allow for departures from our simple theoretical setting in our empirical tests. As we pointed out earlier, the model considers price changes. However, as per empirical convention, and to preserve comparability in the cross-section, we analyze proportional returns in our tests to follow. This distinction, of course, is of no material consequence in that the economic forces in the model apply equally to price changes and returns.\textsuperscript{3}

3 Data

The transactions data sources are the Institute for the Study of Securities Markets (ISSM) and the NYSE Trades and Automated Quotations (TAQ) databases. The ISSM data cover 1988-1992 inclusive while the TAQ data are for 1993-2002. We use only NYSE stocks to avoid any possibility of the results being influenced by differences in trading protocols.

3.1 Inclusion Requirements

Stocks are included or excluded depending on the following criteria:

1. To be included in any given year, a stock had to be present at the beginning and

\textsuperscript{3}See for instance, Hong and Stein (1999) who also model price changes but draw implications for returns that are tested in Hong, Lim, and Stein (2000).
at the end of the year in both the Center for Research in Security Prices (CRSP) and the intraday databases.

2. If a firm changed exchanges from Nasdaq to NYSE during the year (no firms switched from the NYSE to the Nasdaq during our sample period), it is dropped from the sample for that year.

3. Since their trading characteristics might differ from those for ordinary equities, assets in the following categories are also expunged: certificates, American Depositary Receipts, shares of beneficial interest, units, companies incorporated outside the U.S., Americus Trust components, closed-end funds, preferred stocks and Real Estate Investment Trusts.

4. To avoid the influence of unduly high-priced stocks, if the price at any month-end during the year was greater than $999, the stock was deleted from the sample for the year.

5. Stock-days on which there are stock splits, reverse splits, stock dividends, repurchases or a secondary offering are eliminated from the sample.

Next, intraday data are purged for one of the following reasons: trades out of sequence, trades recorded before the open or after the closing time, and trades with special settlement conditions (because they might be subject to distinct liquidity considerations). Our preliminary investigation revealed that auto-quotes (passive quotes by secondary market dealers) were eliminated in the ISSM database but not in TAQ.
This caused the quoted spread to be artificially inflated in TAQ. Since there is no reliable way to filter out auto-quotes in TAQ, only BBO (best bid or offer)-eligible primary market (NYSE) quotes are used in calculating imbalances and mid-point returns. Also, quotes established before the opening of the market or after the close were discarded. Negative bid-ask spread quotations, transaction prices, and quoted depths were discarded. Following Lee and Ready (1991), for the period 1988 to 1998, any quote less than five seconds prior to the trade is ignored and the first one at least five seconds prior to the trade is retained. Based on feedback from microstructure scholars, who indicated that reported errors dramatically declined in the 1999-2002 period, this delay was not imposed for the last four years of the sample period.

3.2 Imbalance and Return Data

We sign trades using the Lee and Ready (1991) procedure: if a transaction occurs above the prevailing quote mid-point, it is regarded as a purchase and vice versa. If a transaction occurs exactly at the quote mid-point, it is signed using the previous transaction price according to the tick test (i.e., buys if the sign of the last non-zero price change is positive and vice versa). For each stock we then define OIB, the estimated monthly buyer-initiated minus seller-initiated dollar volume of transactions. We recognize that order imbalance is measured with error and in Appendix B, we explore the likely impact of this error on our inferences.\(^4\)

\(^4\)Odders-White (2000) and Lee and Radhakrishna (2000) argue that the Lee and Ready (1991) algorithm correctly classifies more than between 85%-93% of trades, which informally indicates that the rule may be adequate for our purposes.
We recognize that our algorithm generally allows us to sign only market orders, so that our net imbalance measures the aggregate demand of agents that require immediacy. While this caveat is worth mentioning, we believe that the standard microstructure paradigm is of patient market makers (which include limit order traders) who absorb the demands of traders that have relatively urgent needs to trade. As such, we believe that it is hard to argue that inventory effects, if any, would not manifest themselves in premia required to bear the imbalance caused by submitters of market orders.

Of course, return computations may be subject to the well-known bid-ask bounce bias. We therefore do not use returns obtained from CRSP data in our empirical analysis. Instead, we use a series which calculates the daily returns using quote midpoints associated with the last transaction on a particular day (which are not necessarily the same as closing quote midpoints). Throughout, these midpoint returns are used in the analysis.

Some of our regressions require the use of imbalance innovations obtained from past return and imbalance data. For this purpose, in order to allow for reliable estimation of coefficients in the first stage innovation regressions, we require 48 months of return and imbalance observations to be available for stocks to be included in the sample.

\(^5\)Monthly returns are calculated by cumulating the daily mid-point returns.
4 Empirical Results

Our theoretical model, for tractability, and to succinctly bring out the intuition, considers the case of a single security. In performing our analysis on individual securities, however, we have the choice of performing the time-series aggregation of estimated coefficients from cross-sectional regressions, or the cross-sectional aggregation of time-series regression estimates. We show in Appendix B that under certain plausible conditions, the sign of the coefficient estimate obtained from the two aggregation methods will coincide. Aggregation of the estimates obtained from time-series regressions presents econometric problems because the residuals are likely to be cross-correlated due to a systematic component in the independent variable (i.e., imbalance), which contaminates inferences from simple t statistics associated with the average estimates. We choose to focus on time-series aggregation of cross-sectional regression estimates because they represent the well-accepted Fama and MacBeth (1973) technique.\(^6\)

Our goal is to explore the relation between monthly returns and lagged returns as well as lagged order flow. We choose the monthly horizon because of its consideration in earlier papers which explore the inventory and overreaction explanations for short-horizon return predictability (e.g., Jegadeesh and Titman, 1995, Conrad, Hameed, and Niden, 1994, Cooper, 1999, and Mase, 1999).\(^7\)

\(^6\)In unreported analysis, we have aggregated time-series regression estimates cross-sectionally (using the approximation of Chordia, Roll, and Subrahmanyam, 2005 to correct for standard errors), and have obtained similar results.

\(^7\)Hvidkjaer (2005) considers how longer-horizon predictability (specifically, momentum at six- to twelve-month horizons) is related to order imbalances. His focus is on linking momentum to the behavioral models of Daniel, Hirshleifer, and Subrahmanyam (1998), and Barberis, Shleifer, and
4.1 Summary Statistics and Choice of Variables

Table 1 presents the summary statistics for the pooled time-series, cross-sectional sample of 273,585 firm-months. The average monthly return in the sample is 1.3%, whereas the average imbalance is $21.5 million. The mean proportional imbalance is about 2.2%. The average of the absolute level of the proportional imbalance is quite high, about 16.2%. The standard deviations indicate adequate variations relative to the mean to allow us to capture relationships between returns and order flow.

In our empirical analysis, we consider three subsamples of stocks stratified by market capitalization. The motivation for this exercise is that specialist participation rates are lower for the relatively less active issues (small firms), and inventory phenomena are more likely to be evident in such firms.\(^8\)

4.2 Regression Evidence

Of course, one problem in regressing returns on multiple lags of imbalances is the multicollinearity caused by autocorrelations in imbalance (see Chordia, Roll, and Subrahmanyam, 2002).\(^9\) To address this issue, we calculate imbalance innovations by regressing imbalances on contemporaneous as well as twelve lags each of past returns.

---

\(^8\)Madhavan and Sofianos (1998) show that specialist participation rates vary from 52% of the least frequently traded firms to 17% for the most actively traded ones.

\(^9\)For our sample, the cross-sectional averages of the first order autocorrelations range from 0.20 for the large firm sample to 0.06 for the small firm sample. As Chordia and Subrahmanyam (2004) point out, imbalance autocorrelations are greater for larger firms.
and twelve lags of past imbalances.\textsuperscript{10}

In addition, we consider three subsamples of stocks stratified by market capitalization. The motivation for this exercise is twofold. First, specialist participation rates are lower for the relatively less active issues (small firms), and this may influence inventory phenomena.\textsuperscript{11} Specifically, if the specialist faces greater competition from floor traders to provide liquidity, it may take longer to work off an order imbalance since he has less flexibility in price setting. Another argument is that large firms may have greater dollar imbalances due to greater institutional participation, and these may cause greater inventory control problems for market makers.\textsuperscript{12}

Table 2 presents the simplest possible test. Specifically, we report the results of performing Fama-MacBeth regressions which involve regressing the current month’s returns on the past two months’ returns and two lagged imbalance innovations.\textsuperscript{13} We find mixed support for the inventory hypothesis. For the full sample, the second lag of imbalance is significant at the 10\% level; the first lag is not significant. For the sample stratified by size, we find that the first lag of imbalance is significant for the midcap size tercile, and no imbalance lag is significant for the other two terciles.

\textsuperscript{10}We performed identification checks on the time series of imbalances using the Bayesian information criterion for a random sample of fifty stocks. In no case was the indicated autoregressive lag length greater than twelve. We thus settle for a twelve lag-model as being a reasonable one for obtaining imbalance innovations.

\textsuperscript{11}Madhavan and Sofianos (1998) show that specialist participation rates vary from 52\% of the least frequently traded firms to 17\% for the most actively traded ones.

\textsuperscript{12}Confirming this intuition, within our sample, the average absolute order imbalance as high as $74.5 million for the large cap group, but only about $6.6 million for the other two groups combined.

\textsuperscript{13}The choice of lags is based on the analyses of Hasbrouck and Sofianos (1993), and Madhavan and Smidt (1993), which are discussed in the introduction. In unreported analyses, we found that longer lags of order flows (up to twelve) were not significant.
Next, in Table 3, we cumulate returns and imbalances over two months to provide a more complete picture of how lagged returns and imbalances affect future returns. Specifically, we calculate the compounded return over the previous two months, and the total imbalance innovation over the same period. We then perform a Fama-MacBeth regression of monthly returns on lagged values of these cumulated returns and imbalance innovations.

The results in the table indicate that the lagged imbalance variable is now negative and strongly significant for the full sample, supporting the inventory hypothesis, whereas the lagged return variable is significant only at the 10% level. This supports our theoretical model. For the small cap group, however, lagged imbalance is not significant. The point estimate of this variable is negative and strongly significant for the midcap group, and significant at the 10% level for the largecap group.

Why the midcap results are the strongest among the three size-based groups presents an interesting conundrum. Perhaps this results from two different types of effects: on the one hand specialists of largecap firms have greater competition from floor traders which may limit their ability to work off their inventory, but on the other the trading activity for such firms is larger, making it easier to work off the inventory. The midcap results may balance these two effects. The lack of evidence for inventory control in the smallcap firms may be because specialists, facing little competition from competing liquidity sources, may be able to work off imbalances in a more timely manner. Overall, the results provide qualified support for the inventory hypothesis, especially
for the midcap and, to a lesser extent, the large firms.

In our next set of tests, we separately consider positive and negative imbalances. Thus, we consider two separate imbalance-related variables in our regressions, specifically, \( \text{OIBP}=\max(\text{OIB},0) \) and \( \text{OIBM}=\min(\text{OIB},0) \). One motivation is that intense buying by outsiders may require market makers to go short and thus, due to short-selling constraints, may result in stronger evidence of inventory control. Alternatively, short-selling constraints may cause negative sentiment (or information-based demand) to build up to very high levels, and the excessive selling that is caused by the demand crossing a threshold may cause a severe price drop and a greater rebound than the corresponding one on the buy side.

Table 4 presents the results of performing Fama-MacBeth regressions which involve regressing the current month’s returns on the past two months’ returns and two lagged positive and negative imbalance innovations,\(^{14}\) together with twelve lags of returns as controls (the coefficients on the returns are not reported for brevity). Note that controlling an year’s worth of lagged returns in both the innovation calculations and regressions provides insight on whether imbalances continue to predict returns after accounting for the effects of return momentum (Jegadeesh and Titman, 1993).

We find that for the full sample, the second lags of negative imbalance innovations are significant in the manner suggested by the inventory control theory. For the small

\(^{14}\)The choice of lags is based on the analyses of Hasbrouck and Sofianos (1993), and Madhavan and Smidt (1993), which are discussed in the introduction. In unreported analyses, we found that longer lags of order flows (up to twelve) were not significant.
cap firms, the second lag of positive imbalances are significant at the 10% level. For the mid cap and large firms, the first and second lag of negative imbalance innovations are significant. These results are broadly consistent with inventory control. From the perspective of economic significance, for the full sample, perturbation of the 2nd lag of the imbalance from its mean to zero causes an excess return of 0.099% per month, which is substantive.

The results in Table 4 also suggest that the first lag of monthly returns is significant for all the terciles. This phenomenon appears in Jegadeesh (1990) and has been analyzed further in Cooper (1999) as well as Gutierrez and Kelley (2005). The latter paper suggests that the monthly reversal arises from investor overreaction. Since the phenomenon has already been considered and analyzed in the literature, we do not focus on it within our paper, but instead focus on order flows. What is noteworthy for our purposes is that the second lag of return is nowhere significant, but the second lag of order flows often is so.\footnote{We have not split lagged returns into positive and negative ones in the results we report. In unreported regressions, however, we have ascertained that such a split does not cause the second lag of the return variables to become significant, though the first lags continue to be significant. These results are available upon request.}

We next control for risk by regressing each stock’s monthly return on the Fama and French (1993) factors (obtained from Wharton Research Data Services) and using the residuals (i.e., the risk-adjusted returns) as our dependent variable. The results are presented in Table 5. As can be seen, while there is a reduction in the point estimate of the imbalance innovation coefficients, the results are qualitatively unaltered. The
results are broadly unchanged when the sample is stratified by size and hence are not presented for brevity.

Overall, our evidence, to a large extent, is consistent with inventory control phenomena at longer horizons of up to two months, especially for stocks other than those belonging to the smallest firm group. The results accord with the analysis of Hasbrouck and Sofianos (1993) as well as Madhavan and Smidt (1993), who suggest that inventory mean-reverts rather slowly, and has a half life of more than a month. Our finding that the results are stronger for negative imbalances than for positive ones is consistent with the notion that short-selling constraints may cause negative sentiment to build up to very high levels. The excessive selling that is caused by the sentiment crossing a threshold may cause a severe price drop and a greater rebound due to inventory control issues than the corresponding one on the buy side.

5 Conclusion

In this paper, we argue that inventory effects, which have been hard to identify in financial market data, may be present at horizons longer than those suspected. We begin with our theoretical model, which predicts a negative relation between lagged order flows and returns, even after controlling for returns. Our empirical analysis provides some evidence that longer horizon (monthly and bimonthly) order flow innovations are negatively related to future returns in a manner consistent with inventory control by market makers. The pattern is more evident in the larger firms. This conforms
to intuition that imbalances may be larger for large firms, causing greater inventory control effects in such firms. In addition, the results are stronger for negative imbalances. This result is consistent with the notion that short-selling constraints may cause negative sentiment to build up to high levels, and the spurt of selling accompanying the sentiment crossing a threshold may cause significant inventory problems for market makers.

A puzzle that is raised by our work, and that of Hasbrouck and Sofianos (1993) as well as Madhavan and Smidt (1993), is exactly why it may take so long for inventory control effects to manifest themselves in financial market data. A possibility is cognitive constraints of market makers that preclude them from frequent monitoring of inventory. Further analysis of this issue is left for future research.
References


Gutierrez, R., and E. Kelley, 2005, Evidence to the contrary: Extreme weekly returns are not overreactions, working paper, University of Oregon.


Madhavan, A., and S. Smidt, 1993, An analysis of changes in specialist inventories and


Appendix A

**Proof of Proposition 1:** Note that the lagged return is simply $P_2 - P_1$, and $P_1$ is non-stochastic. Also, the lagged order flow is simply $-x - z$. Further, the covariance between $P_3 - P_2$ and $P_2 - P_1$ can be written as

$$a(1 - a)v_{\theta} - b^2v_z - c^2v_\delta.$$

Substituting for the coefficients $a$, $b$, and $c$ from Lemma 1 yields the result that the covariance is

$$-\frac{R^4v_\delta v_\epsilon^3v_\theta^2(\epsilon^2v_z + 1)}{[R^2v_\epsilon v_z(v_\delta v_\epsilon + v_\delta v_{\theta} + v_\epsilon v_{\theta}) + v_\delta v_{\theta}]^2},$$

which is negative. The covariance between $P_3 - P_2$ and the order flow $(-x - z)$ is

$$\frac{(1 - a)^2v_{\theta}}{Rv_\epsilon} - b\left[1 - \frac{b}{Rv_\epsilon}\right]v_z - \frac{c^2v_\delta}{Rv_\epsilon}.$$ 

Again, substituting for the coefficients from Lemma 1, the covariance is found to be

$$-\frac{R^3v_\delta v_\epsilon^3v_\theta^2[R^2v_\epsilon v_z(v_\delta + v_{\theta}) + 2v_{\theta}]}{[R^2v_\epsilon v_z(v_\delta v_\epsilon + v_\delta v_\theta + v_\epsilon v_\theta) + v_\delta v_\theta]^2},$$

which also is negative. This proves Part 1 of Proposition 1.

For proving Part 2 of the proposition, we use the expression for conditional expectation given in (2). In our case, $v_1 = [P_3 - P_2]$, and $v_2 = [P_2 - P_1, -x - z]$. Substituting for the various components and explicitly calculating the conditional expectation yields the desired result in Part 2 of Proposition 1. □
Appendix B

In this appendix, we briefly discuss two issues. The first relates to alternative methods of estimation: Fama-Macbeth cross-sectional regressions with aggregation of the estimates through time, and individual stock time-series regressions with cross-sectional aggregation of the time-series estimates. The second discusses the likely impact of measurement error in order imbalances.

Cross-Sectional and Time-Series Estimation Methods

Letting \( c \) denote a cross-sectional expectation and \( t \) denote a time-series one, the Fama-Macbeth approach produces the following (population) expression for the coefficient from a regression of a random variable \( y \) on another random variable \( x \) (we omit subscripts on the variables for brevity):

\[
E_t \left[ \frac{E_c(xy) - E_c(x)E_c(y)}{E_c(x - E_c(x))^2} \right].
\]

The time-series approach produces the expression

\[
E_c \left[ \frac{E_t(xy) - E_t(x)E_t(y)}{E_t(x - E_t(x))^2} \right].
\]

Comparing the two approaches we find that whether the signs of the two expressions coincide depends, in part on the cross-sectional variation in \( E_t(x - E_t(x))^2 \), and the time-series variation in \( E_c(x - E_c(x))^2 \). If these are vanishingly small, for example, then the denominator in each case can be taken out of the expectation and signs of the two estimates will coincide. While no general results are available beyond this, we initially choose to adopt the Fama-Macbeth approach because it allows us to report the
simple $t$ statistic that assumes independence. The time-series estimates, on the other hand, are correlated across stocks so this approach is not possible, and we employ the standard error correction described in the main text. We also use a panel approach that explicitly corrects for cross-correlation in the error terms generated by the time-series regression.

In matrix notation, when a random variable $Y$ is regressed on multiple variables represented by a random vector $X$, then the two centered expressions above can be written in matrix notation as

$$E_t \left[ \Sigma_{cYX} \Sigma_{cXX}^{-1} \right],$$

and

$$E_c \left[ \Sigma_{tYX} \Sigma_{tXX}^{-1} \right],$$

where $\Sigma_{AB}$ denotes the covariance vector of the variable $A$ with the vector $B$, $\Sigma_{AA}$ is the variance-covariance matrix of the vector $A$, and the subscripts $t$ and $c$ are defined as before. A similar intuition holds in this case. If the elements of $\Sigma_{tXX}^{-1}$ and $\Sigma_{cXX}^{-1}$ exhibit vanishingly small variation in the cross-section and time-series, respectively, the expectation operator will have a negligible impact on these terms, and the signs of the two estimate vectors will coincide on an element-by-element basis.

**Measurement Error in Order Imbalance**

We now attempt to understand how the error-in-variable problem could affect our regression estimation. We conduct the analysis in number of transactions, rather than dollars, because this allows a tractable handling of the issue; however, this does not
change the basic nature of the conclusions. Let \( O \) be the true imbalance, measured as

\[
O = \rho V - (1 - \rho)V = (2\rho - 1)V,
\]

where \( V \) represents the total number of transactions, and \( \rho \) is the fraction of trades signed as buys. Suppose a fraction \( r \) of the trades are signed with error. Then, we can write the measured imbalance, \( \hat{O} \), as

\[
\hat{O} = (1 - r)(2\rho - 1)V + rV(2\hat{\rho} - 1),
\]

where \( \hat{\rho} \) is the assigned fraction of buys in that part of the total number of transactions which are signed incorrectly. Let \( 2\rho - 1 \equiv k \) and \( 2\hat{\rho} - 1 = k' \). Then, we have

\[
\hat{O} = kV + rV(k - k').
\]

Suppose a dependent variable \( Y \) is regressed on \( \hat{O} \) whereas its true relationship with imbalance is expressed as \( bO + e \), where \( e \) is the usual OLS error term. Assume for simplicity that \( k \) and \( k' \) are independent of each other and of \( V \) and have the same variance \( \sigma_k^2 \). Let \( \sigma_V^2 \) denote the variance of \( V \). This implies that

\[
\text{cov}(Y, \hat{O}) = (1 + r)b\sigma_k^2\sigma_V^2,
\]

and

\[
\text{var}(\hat{O}) = \sigma^2\sigma_V^2(1 + r)^2 + r^2\sigma_k^2\sigma_V^2.
\]

Thus, the probability limit of the measured regression coefficient, \( \hat{b} \), can be written as

\[
\text{plim} \ \hat{b} = \frac{(1 + r)b}{(1 + r)^2 + r^2},
\]
so that the estimated coefficient is inconsistent. In papers by Lee and Radhakrishna (2002), Odders-White (2002), and Ellis, Michaely, and O’Hara (2000), the estimate of $r$ ranges from about 7% to about 15%. The corresponding range for the inconsistency factor $(1 + r)/[(1 + r)^2 + r^2]$ is 0.93 to 0.86, with implied inflation factors (reciprocals) of 1.08 and 1.17. Such a range does not appear sufficient to alter our conclusions since most of our conclusions are based on highly significant $t$-statistics.
Table 1: Summary Statistics

This table presents the summary statistics associated with the pooled cross-section and monthly time-series of NYSE stocks used in the analysis. The time-period is January 1988 to December 2002. Returns are cumulated using daily mid-points of the quoted spreads associated with the last transaction of each trading day. Order imbalance is measured by the difference in dollar buys and dollar sells (in millions of dollars). Proportional order imbalance is obtained from dividing order imbalance by the total dollar values of buys and sells.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>0.0133</td>
<td>0.0044</td>
<td>0.6146</td>
</tr>
<tr>
<td>Absolute return</td>
<td>0.0970</td>
<td>0.0615</td>
<td>0.6070</td>
</tr>
<tr>
<td>Order imbalance</td>
<td>21.52</td>
<td>0.3702</td>
<td>117.17</td>
</tr>
<tr>
<td>Absolute value of order imbalance</td>
<td>27.44</td>
<td>2.71</td>
<td>115.92</td>
</tr>
<tr>
<td>Proportional imbalance</td>
<td>0.0222</td>
<td>0.0472</td>
<td>0.2169</td>
</tr>
<tr>
<td>Absolute value of proportional imbalance</td>
<td>0.1624</td>
<td>0.1259</td>
<td>0.6146</td>
</tr>
</tbody>
</table>
Table 2: Monthly Cross-Sectional Regressions for Lagged Returns and Imbalance Innovations

This table presents the results of individual stock time-series regressions for monthly returns of NYSE stocks on two lags of monthly returns and two lags of imbalance innovations. The time-period is January 1988 to December 2002. Returns are cumulated using daily mid-points of the quoted spreads associated with the last transaction of each trading day. LRET and L2RET denote the first two lags of the monthly return. LROIB and L2ROIB represent the first two lags of imbalance innovations obtained from regressing imbalance on current returns and twelve lags each of returns and imbalances. Size terciles are formed by sorting all stocks into groups every month by their market capitalization as of the end of the previous month. All return and imbalance coefficients are multiplied by factors of $10^2$ and $10^4$, respectively.

Panel A: Full sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>$t$ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRET</td>
<td>−2.65</td>
<td>−3.44</td>
</tr>
<tr>
<td>L2RET</td>
<td>−0.211</td>
<td>−0.24</td>
</tr>
<tr>
<td>LROIB</td>
<td>−0.077</td>
<td>−0.71</td>
</tr>
<tr>
<td>L2ROIB</td>
<td>−0.139</td>
<td>−1.81</td>
</tr>
</tbody>
</table>

Panel B: By size tercile

<table>
<thead>
<tr>
<th>Variable</th>
<th>Small firm tercile</th>
<th>Mid-cap firm tercile</th>
<th>Large firm tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>$t$ statistic</td>
<td>Coefficient</td>
</tr>
<tr>
<td>LRET</td>
<td>−4.13</td>
<td>−3.04</td>
<td>−3.35</td>
</tr>
<tr>
<td>L2RET</td>
<td>−0.117</td>
<td>−0.07</td>
<td>−3.28</td>
</tr>
<tr>
<td>LROIB</td>
<td>7.23</td>
<td>1.25</td>
<td>−3.20</td>
</tr>
<tr>
<td>L2ROIB</td>
<td>−5.31</td>
<td>−1.07</td>
<td>−1.58</td>
</tr>
</tbody>
</table>
Table 3: Monthly Cross-Sectional Regressions for Lagged Returns and Lagged Imbalances

This table presents the results of cross-sectional Fama-Macbeth type regressions for monthly returns of NYSE stocks on the two-month cumulative lagged return and the lagged cumulative imbalance over two months. The time-period is January 1988 to December 2002. Returns are cumulated using daily mid-points of the quoted spreads associated with the last transaction of each trading day. L2CRET is the cumulated return over the past two months. Order imbalance innovations (obtained from regressing imbalance on current returns and twelve lags each of returns and imbalances) are cumulated over the past two months (L2COIB). Size terciles are formed by sorting all stocks into groups every month by their market capitalization as of the end of the previous month. All return and imbalance coefficients are multiplied by factors of $10^2$ and $10^4$, respectively.

Panel A: Full sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>$t$ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2CRET</td>
<td>$-1.08$</td>
<td>$-1.73$</td>
</tr>
<tr>
<td>L2COIB</td>
<td>$-0.149$</td>
<td>$-2.87$</td>
</tr>
</tbody>
</table>

Panel B: By size tercile

<table>
<thead>
<tr>
<th>Variable</th>
<th>Small firm tercile</th>
<th>Mid-cap firm tercile</th>
<th>Large firm tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>$t$ statistic</td>
<td>Coefficient</td>
</tr>
<tr>
<td>L2CRET</td>
<td>$-2.28$</td>
<td>$-2.36$</td>
<td>$-3.08$</td>
</tr>
<tr>
<td>L2COIB</td>
<td>$1.54$</td>
<td>$0.54$</td>
<td>$-2.15$</td>
</tr>
</tbody>
</table>
Table 4: Monthly Cross-Sectional Regressions for Lagged Returns and Imbalance Innovations

This table presents the results of individual stock time-series regressions for monthly returns of NYSE stocks on two lags of monthly returns and two lags of imbalance innovations. The time-period is January 1988 to December 2002. Returns are cumulated using daily mid-points of the quoted spreads associated with the last transaction of each trading day. LRET and L2RET denote the first two lags of the monthly return. LROIB and L2ROIB represent the first two lags of imbalance innovations obtained from regressing imbalance on current returns and twelve lags each of returns and imbalances. We define LROIBP=max(LROIB,0) L2ROIBP=max(L2ROIB,0), LROIBM=min(LROIB,0) and L2ROIBM=min(L2ROIB,0). Size terciles are formed by sorting all stocks into groups every month by their market capitalization as of the end of the previous month. Twelve lags of the monthly return are included in the regressions, but their coefficients are not reported. All return and imbalance coefficients are multiplied by factors of $10^2$ and $10^4$, respectively.

Panel A: Full sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRET</td>
<td>$-3.29$</td>
<td>$-4.60$</td>
</tr>
<tr>
<td>L2RET</td>
<td>$-0.435$</td>
<td>$-0.53$</td>
</tr>
<tr>
<td>LROIBP</td>
<td>$0.177$</td>
<td>$1.57$</td>
</tr>
<tr>
<td>L2ROIBP</td>
<td>$0.003$</td>
<td>$0.18$</td>
</tr>
<tr>
<td>LROIBM</td>
<td>$-0.256$</td>
<td>$-1.26$</td>
</tr>
<tr>
<td>L2ROIBM</td>
<td>$-0.461$</td>
<td>$-3.47$</td>
</tr>
</tbody>
</table>

Panel B: By size tercile

<table>
<thead>
<tr>
<th>Variable</th>
<th>Small firm tercile</th>
<th>Mid-cap firm tercile</th>
<th>Large firm tercile</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRET</td>
<td>$-4.85$</td>
<td>$-3.41$</td>
<td>$-3.24$</td>
</tr>
<tr>
<td>L2RET</td>
<td>$-0.464$</td>
<td>$-0.30$</td>
<td>$-1.71$</td>
</tr>
<tr>
<td>LROIBP</td>
<td>$-4.09$</td>
<td>$-0.27$</td>
<td>$0.188$</td>
</tr>
<tr>
<td>L2ROIBP</td>
<td>$-23.53$</td>
<td>$-1.83$</td>
<td>$-0.146$</td>
</tr>
<tr>
<td>LROIBM</td>
<td>$-4.31$</td>
<td>$-0.40$</td>
<td>$0.003$</td>
</tr>
<tr>
<td>L2ROIBM</td>
<td>$34.93$</td>
<td>$1.36$</td>
<td>$-0.360$</td>
</tr>
</tbody>
</table>
Table 5: Monthly Cross-Sectional Regressions for Lagged Returns and Lagged Imbalances, using Fama and French (1993)-adjusted returns

This table presents the results of cross-sectional Fama-Macbeth type regressions for monthly Fama and French (1993)-adjusted returns of NYSE stocks on two lags of monthly returns and two lags of imbalance innovations. The time-period is January 1988 to December 2002. Returns are cumulated using daily mid-points of the quoted spreads associated with the last transaction of each trading day. LRET and L2RET denote the first two lags of the monthly return. LROIB and L2ROIB represent the first two lags of imbalance innovations obtained from regressing imbalance on current returns and twelve lags each of returns and imbalances. We define LROIBP=max(LROIB,0) L2ROIBP=max(L2ROIB,0), LROIBM=min(LROIB,0) and L2ROIBM=min(L2ROIB,0). All return and imbalance coefficients are multiplied by factors of $10^2$ and $10^4$, respectively.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRET</td>
<td>-2.23</td>
<td>-2.82</td>
</tr>
<tr>
<td>L2RET</td>
<td>-0.272</td>
<td>-0.31</td>
</tr>
<tr>
<td>LROIBP</td>
<td>0.193</td>
<td>1.71</td>
</tr>
<tr>
<td>L2ROIBP</td>
<td>0.001</td>
<td>0.06</td>
</tr>
<tr>
<td>LROIBM</td>
<td>-0.178</td>
<td>-0.78</td>
</tr>
<tr>
<td>L2ROIBM</td>
<td>-0.397</td>
<td>-2.85</td>
</tr>
</tbody>
</table>