4-2019

Order flow volatility and equity costs of capital

Tarun CHORDIA
Emory University

Jianfeng HU
Singapore Management University, JIANFENGHU@smu.edu.sg

Avanidhar SUBRAHMANYAM
University of California, Los Angeles

Qing TONG
Singapore Management University, qingtong@smu.edu.sg

Citation

CHORDIA, Tarun; HU, Jianfeng; SUBRAHMANYAM, Avanidhar; and TONG, Qing. Order flow volatility and equity costs of capital. (2019). *Management Science*. 65, (4), 1455-1947.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/4552

October 17, 2016

Order Flow Volatility and Equity Costs of Capital

Tarun Chordia*
Jianfeng Hu**
Avanidhar Subrahmanyam***
Qing Tong**

* Goizueta Business School, Emory University.

**Lee Kong Chian School of Business, Singapore Management University.

***Anderson Graduate School of Management, University of California at Los Angeles.

We are grateful to Gustavo Manso (the editor), and two anonymous referees for insightful and constructive comments. We also thank Gennaro Bernile, Ekkehart Boehmer, Jerry Cao, Fangjian Fu, Dashan Huang, Raymond Kan, Robert Kimmel, Tao Li, Roger Loh, Roni Michaely, Fenghua Song, Yuehua Tang, Jun Tu, Rossen Valkanov, Kam-Ming Wan, Junbo Wang, Xueping Wu, Joe Zhang, Zhong Zhang, and participants at the Fudan University, City University of Hong Kong, Hong Kong Polytechnic University, and Singapore Management University seminar series, the LKCSB Summer Finance Research Camp 2015, and the Macquarie's Global Quantitative Research Conference in Hong Kong, for helpful comments. This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 grant and the Sim Kee Boon Institute for Financial Economics at the Singapore Management University. Jianfeng Hu also acknowledges financial support from the Lee Kong Chian Fund for Excellence.

Order Flow Volatility and Equity Costs of Capital

Abstract

We propose that the volatility of order flow is a proxy for costs of information asymmetry, as order flow volatility varies positively with parameters that also influence adverse selection costs of trading. Empirically, order flow volatility is significantly higher prior to earnings or merger announcements when information asymmetry is likely to be elevated. Levels of and shocks to order flow volatility are positively and significantly correlated with existing illiquidity proxies, and strongly predict stock returns in the cross section. The impact of order imbalance volatility shocks on stock prices is reflected within one month in large, visible stocks, but takes up to three months to be fully reflected in small, "neglected" stocks.

1 Introduction

We examine the relation between returns and both, levels of as well as shocks to the variability of order flows. Both, levels of and shocks to order flow volatility have an impact on required equity returns beyond a comprehensive list of return predictors. Our analysis accords with theoretical arguments that link order flow variability to information asymmetry costs, and also unveils a novel predictor of stock returns, that is statistically and economically significant. To the best of our knowledge, this link between required returns (costs of capital) and order flow volatility has not been previously explored in the literature.

We sign trades as buys and sells to create two measures of order imbalance (OIB), one based on shares traded and the other based on number of trades. OIB, in terms of shares traded, is denoted OIB_SHR and is constructed as the number of shares bought less the number of shares sold as a fraction of the sum of shares bought and sold. Similarly, OIB, in terms of the number of trades is denoted OIB_NUM and is constructed as the number of buy trades less the number of sell trades as a fraction of the sum of the total trades. The OIB volatilities ($VOIB\_SHR$ and $VOIB\_NUM$) are computed each month as the standard deviation of the daily OIB_SHR and OIB_NUM in terms of shares traded and the number of trades, respectively.

We find that both $VOIB\_SHR$ and $VOIB\_NUM$ are cross-sectionally correlated with different measures of liquidity, including turnover, bid-ask spreads and the Amihud measure of illiquidity. Univariate sorts of $VOIB\_SHR$ and $VOIB\_NUM$ into quintile portfolios show that a portfolio that is long the high OIB volatility stocks and short the low OIB volatility stocks, yields a monthly return of about 80 basis points as does the long-short portfolio formed by sorting on the Amihud illiquidity measure. Fama-MacBeth (1973) regressions reveal that lagged values of $VOIB\_SHR$ and $VOIB\_NUM$ positively predict risk-adjusted returns.

Further, shocks to $VOIB\_SHR$ and $VOIB\_NUM$ (denoted respectively by $SVOIB\_SHR$ and $SVOIB\_NUM$) are highly positively correlated with shocks to the bid-ask spread and shocks to the Amihud measure of illiquidity and negatively correlated with shocks to turnover. A positive shock reduces the contemporaneous and next month's returns. Quintile portfolios with the largest shocks to volatility of OIB underperform those with the smallest shocks by

2.68% (2.00%) in the current month and by 0.79% (0.95%) for $SVOIB\_SHR$ ($SVOIB\_NUM$) in the following month. Thus, shocks to the volatility of order flow have contemporaneous as well as delayed effects on illiquidity and required rates of return, consistent with the notion that agents' reaction to liquidity shocks is not instantaneous. This part of the analysis also accords with Bali, Peng, Shen, and Tang (2014) (BPST), who show that shocks to the Amihud (2002) measure of illiquidity lead to lower future returns. However, the effect of order flow volatility shocks survives controls for shocks to the Amihud (2002) measure of illiquidity.

Our motivation for studying order flow volatility stems from the recognition that adverse selection costs are an important source of premiums in asset prices (Easley and O'Hara, 2004). There is a large and growing literature that examines the impact of adverse selection on asset returns. Easley, Hvidkjaer and O'Hara (2002) estimate the probability of informed trading (PIN) from order imbalances to show that PIN impacts expected returns.[1] Easley, Lopez de Prado and O'Hara (2012) compute a modified measure of PIN denoted $VPIN$ to show that it can predict stressful events in the market such as the flash crash. Kelly and Ljungqvist (2012) use exogenous shocks to analyst following of firms to show that prices and uninformed demand fall as information asymmetry increases and the channel that links asymmetry to prices is liquidity. Akbas, Armstrong and Petkova (2011) document a positive relation between the volatility of liquidity and expected returns. Using an information risk measure based on the price discovery of large trades, Hwang and Qian (2011) find that their measure is priced in the cross-section of stock returns. Johnson and So (2015) use the option-to-stock volume ratio as a measure of information asymmetry. Choi, Jin and Yan (2016) find a positive relation between institutional ownership volatility and expected returns. Yang, Zhang and Zhang (2015) use abnormal idiosyncratic volatility as a measure of information asymmetry to show that it is positively related to future returns. Collin-Dufresne and Fos (2015) study how measures of adverse selection respond to informed trading given the possibility that informed traders with long-lived, monopolistic private information may submit limit orders.[2]

Given the many and diverse measures of information asymmetry proposed in the literature,

[1] Back, Crotty and Li (2016) show that a hybrid measure of informed trading based on PIN and the Kyle (1985) model explains more cross-sectional variation in price impacts than PIN.

[2] We discuss this in more detail in Section 9.

we motivate our measure on the basis of a model (presented in the appendix), which shows that order flow volatility and adverse selection costs are related positively to the variables capturing information in financial markets. Therefore, order flow volatility, which can be readily measured from transactions data, can proxy for the adverse selection costs. We test this by comparing order flow volatility prior to earnings or merger and acquisition announcements (when information asymmetry is likely to be higher) to other periods. Order flow volatility is indeed significantly higher prior to these events.

We find that the predictability of subsequent months' returns from order flow volatility and its shocks survives a long list of control variables including those used in the literature listed above. These variables include firm characteristics such as momentum, monthly reversals, idiosyncratic volatility, profitability, analyst forecast dispersion, illiquidity, turnover, asset growth, accruals, new issues, PIN, return volatility, turnover volatility, illiquidity volatility and shocks to turnover, illiquidity, institutional holdings and return volatility. Importantly, as mentioned earlier, this return forecastability also survives the measure of illiquidity shocks, i.e., innovations to the Amihud measure of liquidity, as developed by BPST, suggesting that shocks to order flow volatility measure something more than shocks to illiquidity and are possibly more closely related to adverse selection measures.

The effect of lagged shocks to the volatility of OIB is robust to different return definitions including risk-adjusted returns, raw returns, and open-to-close midquote returns. The finding is not driven by the recent financial crisis and it is robust to alternative order flow calculation based on dollars traded. The results also survive across different subperiods. The impact of shocks to volatility of OIB is stronger for firms with small market capitalization, low institutional holdings, and high idiosyncratic volatility. We also find that for small stocks it takes over three months for $VOIB$ shocks to be incorporated into prices, but for large companies, the adjustment happens in one month. We also find that $SVOIB$ has a negative impact on prices initially, but the shock's impact turns positive after three months, consistent with the notion that order flow volatility positively impacts firms' required returns.

The rest of the paper is organized as follows. Section 2 presents the motivation in the context of a model presented in the appendix. Section 3 describes the sample selection and variable construction in empirical analysis. Sections 4 and 5 examine the pricing effect of the proposed volatility of order flow as a measure of illiquidity and the shocks to the volatility of order flow. Section 6 presents the shocks to the volatility of order flow results in different informational environments. Section 7 studies the return dynamics of the liquidity shocks. Sections 8 further validates our measure by considering order flow volatility around corporate events. Section 9 contrasts our measure to one obtained from using the limit order book. Section 10 concludes. A large number of robustness checks appear in an Internet Appendix, and are referenced within the paper.

2 Motivation

Adverse selection costs of trading have been recognized as a determinant of equity cost of capital in Easley and O'Hara (2004). However, these costs are hard to measure. We take the approach that such costs are linked to other endogenous constructs in equilibrium and measuring these other constructs might be easier than measuring adverse selection costs directly. Moreover, this measurement may lead to additional insights on the relation between information asymmetry and asset prices. Specifically, we note that in the celebrated Kyle (1984, 1985) model, order flow from traders who submit market orders plays a key role as a signal from which the market maker attempts to extract information. Indeed, we analytically show in the appendix that from an ex ante standpoint, the standard deviation of order flow is directly linked to the exogenous parameters that drive adverse selection costs. Intuitively, high order flow volatility indicates that informed traders are more active, which is associated with higher adverse selection costs. Since this standard deviation of order flow can be measured using transactions data, the dynamics of the second moment of order flow can potentially be linked to the dynamics of the (unobserved) true adverse selection costs, and, in turn, to required equity returns, even when the exogenous parameters are not observable.

Previous efforts to link trading costs to equity prices in the cross section mainly focus on the relation between the level of such costs and future stock returns; see, for example, Amihud and

Mendelson (1986), Brennan and Subrahmanyam (1996), Jacoby, Fowler, and Gottesman (2000), Jones (2002), and Amihud (2002). In a recent paper, BPST examine the relation between stock returns and *shocks* to illiquidity, where illiquidity is measured by the procedure suggested by Amihud (2002), and by bid-ask spreads. BPST explore the idea that investor reactions to illiquidity shocks might not be immediate owing to limitations in information processing, so that such shocks might have a persistent impact on stock prices.

Motivated by the above papers, we examine the relation between returns and *both* levels of as well as shocks to the variability of order flows. We show that levels of and shocks to order flow volatility have an impact on required returns beyond a comprehensive list of return predictors, including those documented by BPST. Our analysis unveils a novel predictor of stock returns, that is statistically and economically significant. To the best of our knowledge, this link between required returns and order flow volatility, has not been previously explored in the literature.

3 Sample Selection and Hypotheses

Our sample includes common stocks listed on the NYSE, AMEX and Nasdaq in the period from January 1983 to December 2012. To be included in the monthly analysis, a stock must have the following data available: (i) its returns in the current month and the past twelve months from CRSP, (ii) sufficient data to calculate market capitalization and turnover, (iii) data on the Compustat tapes to calculate the book-to-market ratio as of December of the previous year, and (iv) data in the NYSE Trade and Quote (TAQ) database or the Institute for Study of Security Markets (ISSM) dataset to calculate the order imbalance. To avoid extremely illiquid stocks, we eliminate stock-month observations with month-end stock prices below one dollar. The following securities are also eliminated from the sample since their trading characteristics can differ from ordinary equities: ADRs, shares of beneficial interest, units, companies incorporated outside the U.S., Americus Trust components, closed-end funds, preferred stocks and REITs.

Transactions data are obtained from ISSM (1983-1992) and from TAQ (1993-2012). To eliminate data errors, we exclude trades with non-positive prices and trades mapped to crossed (the bid price greater than the ask price) or locked (the bid price equal to the ask price) quotes.

We also exclude trades in the first fifteen minutes and the last five minutes of trading on each day to increase the accuracy of the trade signing algorithm.[3] We require that all stock-month observations have at least 14 daily trading records in a month.

TAQ and ISSM data do not contain information on whether a trade is initiated by the buyer or the seller. We use the Lee and Ready (1991) algorithm to classify transactions as either a buy or a sell as follows: if a trade is executed at a price above (below) the quote midpoint, we classify it as a buy (sell); if a trade occurs exactly at the quote midpoint, we sign it using the previous transaction price according to the tick test (i.e., a buy if the sign of the last nonzero price change is positive and vice versa). The Lee and Ready algorithm uses the fact that seller-initiated trades tend to execute at a lower price than buyer-initiated trades. We apply the tick test up to the past five price changes. If the past five price changes are zero then we do not use it in the computation of buys or sells. As Lee and Ready (1991) note, the timestamps on quotes are not always correctly synchronized with those for trades and hence they recommend that the quotes be matched to trades with a five-second delay. We follow this five-second delay rule until 1998. Since such recording errors are not observed in the more recent data (see, for example, Madhavan, Porter, and Weaver, 2005 as well as Chordia, Roll, and Subrahmanyam, 2005), we do not impose any delays after 1998.

One concern with the Lee and Ready (1991) algorithm is that it may misclassify the side that initiates a particular trade, even if the trade initiator places a market order. Lee and Radhakrishna (2000) and Odders-White (2000) examine the trade-level accuracy of the Lee and Ready algorithm for NYSE traded stocks and report accuracy rates of 93% and 85%, respectively. Both Lee and Radhakrishna and Odders-White use data from the pre-decimalization era, and it is important to assess the reliability of the Lee and Ready algorithm in the post-decimalization era as well. The most recent study that examines this issue is Chakrabarty, Moulton, and Shkilko (2012). They find that the transaction level accuracy of the Lee and Ready algorithm during the June to December 2005 period is about 68%. The study of Chakrabarty, Moulton, and Shkilko, however, is not directly comparable to Lee and Radhakrishna and Odders-White because it examines Nasdaq stocks, and focuses solely on short sales. Ellis, Michaely, and

[3] The empirical results are largely the same when these trades are included.

O'Hara (2000) is more directly comparable to Chakrabarty, Moulton, and Shkilko because the former also examine the pre-decimalization period accuracy of the Lee and Ready algorithm with Nasdaq stocks. Ellis, Michaely, and O'Hara find an accuracy rate of 81%. Although the lower accuracy rate in Chakrabarty, Moulton, and Shkilko may be partly due to the fact that it focuses only on short sales, it is quite likely that decimalization and increasing prevalence of high frequency trading, with greater limit order activity (O'Hara, 2015) contributed to this phenomenon as well.[4]

What is important from the perspective of our study, however, is not the trade-level accuracy, but the accuracy when trade-level classifications are aggregated. For example, even if a fraction of seller-initiated trades on a particular day is misclassified as buyer-initiated trades and a similar fraction of buyer-initiated trades is also misclassified, then daily-level accuracy would be much greater than trade-level accuracy. In fact, Chakrabarty, Moulton, and Shkilko find that daily-level error rate is close to zero, and statistically insignificant. Therefore, any trade-level misclassification is unlikely to meaningfully impact our tests based on aggregated data. Further, even if trades are signed with error, there is no compelling reason for why the empirical pricing of order flow volatility should be strengthened by this error,[5] so that the signing error does not bias the results in our favor. Nonetheless, we revisit this issue in Section 5.3.

3.1 Measures of Order Imbalance, Order Imbalance Volatility and Shocks to Order Imbalance Volatility

We define order imbalance, order imbalance volatility and shocks to order imbalance volatility as follows:[6]

OIB (order imbalance): We create two measures of order imbalance (OIB), one based on shares

[4]Faster execution of limit orders in the recent era might induce order-initiators to submit limit orders due to a reduction in execution risk, and this would induce errors in the signing rule. Thus, for example, a limit order to buy at a bid, if it executes quickly, might be a superior alternative to a trader than a market buy order which executes at the ask. Note that the order initiator is a buyer, but the trade would execute below the quote mid-point and would be recorded instead as a sell.

[5]For example, in a bivariate cross-sectional regression of returns on order flow volatility, provided returns are not cross-sectionally correlated with the signing error variance, the slope coefficient is in fact attenuated because of cross-sectional variation in the error variance.

[6]The terms "order imbalance" and "order flow" are used interchangeably.

traded and the other based on the number of trades. OIB, in terms of shares traded, is denoted OIB_SHR and is constructed as the number of shares bought less the number of shares sold as a fraction of the sum of shares bought and sold during an observation period. Similarly, OIB, in terms of the number of trades is denoted OIB_NUM and is constructed as the number of buy trades less the number of sell trades as a fraction of the sum of the total number of trades during an observation period. Order imbalance is scaled by the total number of trades or the total number of shares traded so as to eliminate the impact of total trading activity. Actively traded stocks are likely to have higher order imbalances. The scaling standardizes the order imbalance measure. We calculate daily OIB to construct the volatility of order imbalance. For asset pricing tests later, we estimate monthly OIB using all the buy and sell trades in a month.

$VOIB$ (volatility of order imbalance): The OIB volatilities are computed each month as the standard deviation of the daily OIB_SHR and OIB_NUM, and are denoted by $VOIB\_SHR$ and $VOIB\_NUM$, respectively.

$SVOIB$ (shocks to volatility of order imbalance): We compute shocks to $VOIB\_SHR$ and $VOIB\_NUM$ in each month ($SVOIB\_SHR$ and $SVOIB\_NUM$) by subtracting the k-month moving average of $VOIB\_SHR$ and $VOIB\_NUM$ in the previous month. In our main tests, we use the six-month moving average. Our results are not sensitive to the choice of the moving average lag as shown in the robustness tests.[7]

3.2 Summary Statistics

Panel A of Table 1 provides the summary statistics (computed as the time series averages of the monthly cross-sectional statistics) of the above variables. All variables except realized returns are cross-sectionally winsorized at the 0.5% and 99.5% levels.

There are 2948 stocks per month on average in our sample. Both OIB_SHR and OIB_NUM have negative means and medians indicating that, in general, there is more seller initiated trades than buyer initiated trades. The mean (median) of $VOIB\_SHR$ is 0.36 (0.34) and for

[7]We also use an ARMA(1,1) model for $VOIB$ to extract shocks to order flow volatility and find that our central results are unchanged. This analysis is available in the internet appendix (Table A3).

$VOIB\_NUM$ it is 0.29 (0.25) suggesting that OIB_SHR is more volatile. $SVOIB\_SHR$ and $SVOIB\_NUM$ are both close to zero, albeit negative, suggesting that on average, within our sample, there are more, or larger declines (as opposed to increases), in the volatility of order imbalance.[8]

Panel B reports the time-series averages of the cross-sectional correlations between the volatility of OIB and shocks to these volatilities and the well-known liquidity measures, stock returns (RET) and PIN. The liquidity measures include the Amihud (2002) illiquidity measure ($ILLIQ$), bid-ask spread ($SPRD$, obtained using the Holden and Jacobsen (2014) method), and stock share turnover ($TURN$, calculated as the logarithm of the monthly average of the daily ratio of the stock's trading volume to the total number of shares outstanding).[9] PIN is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). The liquidity shocks are computed in a manner similar to shocks to the volatility of order imbalance. For example, the Amihud illiquidity shock ($SILLIQ$) is defined as $ILLIQ$ in the current month minus the average of $ILLIQ$ in the previous six months.

$VOIB\_SHR$ ($VOIB\_NUM$) has a correlation of 0.25 (0.26) with $ILLIQ$; a correlation of -0.51 (-0.50) with the turnover ratio and a correlation of 0.50 (0.51) with the bid-ask spread. This suggests that stocks with a higher volatility of order imbalance have lower share turnover, larger price impact, and wider bid-ask spreads. The fact that the order imbalance volatilities behave in a manner similar to the traditional illiquidity measures supports the notion that the volatility of OIB partially captures liquidity dynamics. $VOIB\_SHR$ ($VOIB\_NUM$) has a correlation of 0.55 (0.58) with PIN suggesting that order flow volatility may be partially proxying for adverse selection.

Turning now to $VOIB$ shocks, in Panel B of Table 1 we find that lagged and concurrent values of $SVOIB\_SHR$ and $SVOIB\_NUM$ are positively correlated with $SILLIQ$ and $SSPRD$, and negatively correlated with $STURN$, indicating that positive shocks to the volatility of order imbalance is associated with liquidity deterioration. However, the correlations are low. For

[8] We also present the statistics for the subperiods before and after 2001 in an internet appendix (Table A2).

[9] Given that Nasdaq trading volume could be overstated due to interdealer trades we follow the methodology of Gao and Ritter (2010) to adjust Nasdaq volume.

instance, the correlation between $SVOIB\_SHR$ and $SILLIQ$ is 0.06 while the correlation between $VOIB\_SHR$ and $ILLIQ$ is 0.25.

Overall, we find that levels of and shocks to order flow volatility are positively correlated with illiquidity proxies and innovations to these proxies, respectively. Based on our discussion in Section 2, these findings accord with the notion that variations in order flow volatility and its shocks are primarily driven by shocks to informed trading parameters, resulting in a positive relation between order flow volatility and illiquidity. However, the low correlations between shocks to volatility of order imbalance and $ILLIQ$ indicate that the time-series variation of the order imbalance volatility can capture some information not contained in the variation of $ILLIQ$. Lagged values of $VOIB\_SHR$ and $VOIB\_NUM$ are positively correlated with RET, suggesting that more illiquid stocks (as measured by the volatility of OIB) have higher expected returns in the cross-section. Both $SVOIB\_SHR$ and $SVOIB\_NUM$ are significantly negatively correlated with stock returns.

Panel C reports the times series averages of the cross-sectional correlations between the order imbalance volatility, the shocks to these volatilities, the return standard deviation (*Ret_Std*), and the shock to the return standard deviation(*SRet_Std*). *Ret_Std* is the standard deviation of daily returns in a month. *SRet_Std* is defined as *Ret_Std* in the current month minus the average of *Ret_Std* in the previous six months. The correlation between the order imbalance volatility and the return standard deviation is positive but low. For instance, the correlation between $VOIB\_SHR$ and *Ret_Std* is 0.06, suggesting that $VOIB$ contains different information from *Ret_Std*. Although $VOIB\_SHR$ and *Ret_Std* are positively correlated, it is interesting that the shocks to these two variables have negative correlations at comparable magnitude, indicating that the dynamics of the two variables are very different.

Figure 1 plots the equally-weighted and value-weighted $VOIB$. There is a clear downward trend for both $VOIB\_SHR$ and $VOIB\_NUM$. In unreported analysis, we regress $VOIB$ on a linear time trend and the estimated time trend is negative and highly significant in each case. This is consistent with Chordia, Roll, and Subrahmanyam (2011) who document a significant increase in liquidity over time. We also plot the equally-weighted and value-weighted $SVOIB$

in Figure 2. There is no clear pattern for $SVOIB$.

3.3 Hypotheses

We describe our testable hypotheses in this subsection. We consider order flow volatility as a proxy for adverse selection costs, because it exhibits comparative statics similar to those for such costs, and because empirically, Table 1 demonstrates that $VOIB$ and its shocks are positively correlated with illiquidity proxies and their shocks, respectively (we provide further evidence on this issue in Section 8). Therefore, we have the following hypotheses (stated in alternative form):

Our first hypothesis relates order flow volatility to asset returns:

Hypothesis 1: Higher order flow volatility ($VOIB$) implies higher expected returns.

The existence of a premium for information asymmetry posits that an asset's price falls upon a positive $VOIB$ shock because the asset becomes less attractive to investors:

Hypothesis 2: Positive shocks to order flow volatility ($SVOIB$) imply lower contemporaneous prices.

Finally, such shocks should increase the expected asset returns in the future when the premium is realized:

Hypothesis 3: Positive shocks to order flow volatility ($SVOIB$) imply higher future returns.

4 Portfolio Sorts

In this section, we provide evidence that both $VOIB$ and $SVOIB$ are priced in the cross section using portfolio sorts.

4.1 Volatility of OIB and Stock Returns

Panel A of Table 2 studies the distribution of stock characteristics across quintile sorted $VOIB$ portfolios. Firm size decreases, while the book-to-market ratio and return volatility increase with the volatility of order imbalance. This is not surprising as smaller firms and more volatile firms are likely to be less liquid and subject to more adverse selection.

To test for the pricing of $VOIB$ (Hypothesis 1), we first present univariate portfolio sort results for the volatility of order imbalance and other liquidity measures. In Panel B of Table 2, we sort stocks into quintile portfolios based on order imbalance, volatility of order imbalance, and other liquidity measures, and report the average raw returns of the quintile portfolios in the next month. Also reported are the raw returns and the alphas (with respect to the Fama and French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor) for the portfolios that are long in stocks in the highest quintile and short in stocks in the lowest quintile. The associated Newey-West (1987) t-statistics are in parentheses.

The results show that $VOIB\_SHR$ and $VOIB\_NUM$ positively predict stock returns in the next month. The raw return differential of the top and bottom quintile portfolios is 0.78% (0.72%) per month for $VOIB\_SHR$ ($VOIB\_NUM$), and both are significant at the 1% level. For the alphas, the differential is even higher at 0.99% (0.97%) per month for $VOIB\_SHR$ ($VOIB\_NUM$). For order imbalances themselves, the return differences between the top and bottom quintiles and the alphas are negative. The high minus low, long-short portfolio alpha amounts to -0.47% (-0.38%) per month when sorting on OIB_SHR (OIB_NUM). This negative relation suggests that the price pressure from order imbalances in the current month reverses in the next month.

We also examine the pricing effects of the traditional liquidity measures. The monthly raw return (alpha) differential for quintile portfolios sorted on $ILLIQ$ is 0.86% (1.15%), indicating that the Amihud illiquidity measure is also significantly and positively related to future returns. For $TURN$ ($SPRD$), the alpha differential is -0.46% (1.20%) per month. The results suggest that illiquid stocks are associated with high future returns, consistent with the consensus in the literature.

To examine whether the order imbalance volatility contains information in addition to the traditional liquidity measures, we double sort stocks based first on the traditional liquidity measures and then on $VOIB\_SHR$ or $VOIB\_NUM$. More specifically, at the end of each month t, we first sort stocks into high and low groups based on median $ILLIQ$, $TURN$, or

$SPRD$, and then sort stocks based on $VOIB\_SHR$ or $VOIB\_NUM$ into quintile portfolios within each group separately. Portfolio returns at month $t+1$ are reported in Panel C for $VOIB\_SHR$ and for $VOIB\_NUM$ in Panel D. Across all the columns, for $VOIB\_SHR$ and $VOIB\_NUM$, the return differences between the top and bottom quintiles and the alphas are larger for stocks with higher $ILLIQ$, higher turnover and higher spreads. This is not surprising given the correlation between the different liquidity measures documented in Table 1. Note that even after controlling for the traditional illiquidity measures, all the return differentials are generally significant at the 5% level or better. This suggests that the volatility of order imbalances provides additional information about the illiquidity of a stock that is not captured by the effect of the traditional liquidity measures, possibly because $VOIB$ is related to adverse selection. In general, the results in Table 2 support our Hypothesis 1 that order flow volatility ($VOIB$) is positively related to the expected returns in the cross section of stocks.

4.2 Shocks to Volatility of OIB and Stock Returns

Having documented the pricing effect of $VOIB$, we now focus on shocks to this measure. We predict that an increase in illiquidity due to an increase in adverse selection (proxied by order flow volatility) will decrease contemporaneous prices (Hypothesis 2) but increase future expected returns (Hypothesis 3). We first consider the impact of a shock to order imbalance volatility on contemporaneous and subsequent months' returns using, in turn, the usual portfolio sort and Fama-MacBeth approaches, and then consider longer horizon returns in Section 7.

Panel A of Table 3 presents the distribution of stock characteristics across quintile sorted $SVOIB$ portfolios. Return volatility has a U-shaped pattern across the quintile portfolios. The highest $SVOIB$ portfolio has lower return volatility than the lowest $SVOIB$ portfolio. The book-to-market ratio and firm size respectively follow a U-shaped and an inverted U-shaped pattern across the quintile portfolios. However, there is no material difference in the size or book-to-market ratio across the highest and lowest $SVOIB$ portfolios.

Panel B of Table 3 presents the average raw returns of the quintile portfolios along with the return differences and the alphas between the top and bottom quintile portfolios in the current month and Panel C reports results in the next month. At the end of each month t, we sort stocks into quintile portfolios based on $SVOIB\_SHR$, $SVOIB\_NUM$, $SILLIQ$, $STURN$, and $SSPRD$, and then examine the returns across the portfolios. The average raw returns are generally positive in all of the quintile portfolios, reflecting an average upward trend in the market during our sample period. The returns change monotonically across all the quintile portfolios regardless of the sorting variable. Consistent with Hypothesis 2, shocks to volatility of order imbalance are negatively correlated with contemporaneous returns. The monthly return difference between the portfolios of stocks with high and low shocks to OIB volatility is quite large at -2.68% (-2.00%) and the alpha is -2.53% (-1.91%) for $SVOIB\_SHR$ ($SVOIB\_NUM$). Consistent with BPST, we find that all of the other liquidity shocks measured using traditional methods also have large and significant impact on the contemporaneous stock prices. The return differences and the alphas across the other liquidity shocks are even higher than those for $SVOIB$. The evidence indicates that shocks to $VOIB$ (our adverse selection proxy) are negatively correlated with contemporaneous returns.

Panel C of Table 3 shows that all of the liquidity shocks predict the next month's returns too. However, at odds with Hypothesis 3, shocks to order imbalance volatility imply lower returns in the following month. The monthly raw return differential between stocks in the highest and the lowest $SVOIB\_SHR$ ($SVOIB\_NUM$) quintiles is -0.79% (-0.95%). The raw return differential between stocks in the highest and the lowest $SILLIQ$ ($STURN$) [$SSPRD$] quintiles is -1.53% (1.31%) [-0.89%] in the following month. The alpha differentials are, in fact, slightly larger. The return and alpha differentials are all statistically significant at the 1% level. BPST argue that because of limited attention (Hirshleifer and Teoh, 2003), such an increase in illiquidity may have a prolonged impact on prices. The notion is that owing to limitations on information processing, it might take time for market participants to realize that illiquidity has changed, which results in a delayed impact of illiquidity shocks on prices.

In order to ascertain whether $SVOIB\_SHR$ and $SVOIB\_NUM$ have any information over and above that contained in $SILLIQ$, $STURN$, $SSPRD$, past one month return, RET, and analyst dispersion in earnings forecasts, $DISP$, in Table 4 we provide results from bivariate sorts. The rationale for the variables is as follows: Since $SVOIB$ measures shocks to liquidity it is important to test whether it contains information over and above the shocks to the standard liquidity variables including $SILLIQ$, $STURN$ and $SSPRD$. Since shocks to order imbalance volatility could be driven by dispersion in investor beliefs we control for analyst forecast dispersion. Finally, low past returns could signal financial distress and thus we control for it in the bivariate sorts. We first sort stocks into quintile portfolios in month t based on $SILLIQ$, $STURN$, $SSPRD$, RET and $DISP$ respectively. Then within each group, we further sort stocks based on $SVOIB\_SHR$ and $SVOIB\_NUM$ into quintile portfolios. Returns in month $t + 1$ are reported for $SVOIB\_SHR$ and $SVOIB\_NUM$ portfolios.

In general, predictive ability of $SVOIB$ is not captured by shocks to the traditional liquidity measures. In the case of $SILLIQ$ and $SSPRD$, all the differential alphas except for those of quintile 3 are significantly negative. In the case of $STURN$, all the alphas except for those of quintiles 3 and 4 are significantly negative for $SVOIB\_SHR$. Thus, $SVOIB$ has return relevant information that is not captured by $SILLIQ$, $STURN$ or $SSPRD$. In the case of RET and $DISP$, all the return differentials are significantly negative suggesting that the impact of $SVOIB$ on returns is not being driven by financial distress or differences in investor beliefs.

While the portfolio sorts do provide support for the idea that shocks to order flow volatility ($SVOIB\_SHR$ and $SVOIB\_NUM$) capture shocks to liquidity over and above those contained in the traditional measures of liquidity (including the Amihud, 2002, measure used by BPST), we now explore this idea in more detail in a regression framework.

5 Asset Pricing Regressions

5.1 Methodology

Our cross-sectional asset pricing tests follow Brennan, Chordia, and Subrahmanyam (1998) and Avramov and Chordia (2006), who test factor models by regressing risk-adjusted returns on firm-level attributes such as size, book-to-market, turnover and past returns. Under the null of exact pricing, such attributes should be statistically and economically insignificant in the cross section. The use of individual stocks as test assets avoids the possibility that tests may be sensitive to the portfolio grouping procedure (Lo and MacKinlay, 1990).

We first regress the excess return of stock j, $(j = 1, \ldots, N)$ on asset pricing factors, F_{kt}, $(k = 1, \ldots, K)$, allowing the factor loadings, β_{jkt}, to vary over time as a function of firm size and book-to-market ratio. The conditional factor loadings of security are modeled as:

$$\beta_{jkt-1} = \beta_{jk1} + \beta_{jk2} Size_{jt-1} + \beta_{jk3} BM_{jt-1}, \tag{1}$$

where $Size_{jt-1}$ and BM_{jt-1} are the market capitalization and the book-to-market ratio at time $t-1$.[10]

The dependence of factor loadings on size and book-to-market is motivated by the general equilibrium model of Gomes, Kogan, and Zhang (2003), who justify separate roles for size and book-to-market as determinants of beta. In particular, firm size captures the component of a firm's systematic risk attributable to growth options, and the book-to-market ratio serves as a proxy for the risk of existing projects.

Subtracting the component of the excess returns associated with the factor realizations generates the risk-adjusted returns, R^*_{jt}:

$$R^*_{jt} = R_{jt} - R_{Ft} - \sum_{k=1}^{K} \beta_{jkt-1} F_{jk}, \tag{2}$$

where R_{Ft} is the risk-free rate, β_{jkt-1} is the conditional beta estimated by a first-pass time-series regression over the entire sample period.[11]

The risk-adjusted returns are then regressed on the equity characteristics:

$$R^*_{jt} = c_{0t} + \sum_{m=1}^{M} c_{mt} Z_{mjt} + e_{jt}, \tag{3}$$

where Z_{mjt} is the lagged one month value of the characteristic m for security j at time t, and M is the total number of characteristics. This procedure ensures unbiased estimates of the

[10] We also check the unconditional specification in which $\beta_{jk}(t) = \beta_{jk}$ (constant betas). The results are unaltered.

[11] Fama and French (1992) and Avramov and Chordia (2006) show that using the entire time series to compute the factor loadings generates qualitatively similar results to those obtained from using rolling regressions. The results are quite similar when we use rolling regressions to estimate the factor betas.

coefficients, c_{mt}, without the need to form portfolios, because the errors in estimation of the factor loadings are included in the dependent variable. The standard Fama-MacBeth (1973) (FM) estimators are the time-series averages of the regression coefficients, $\hat{c}_t$. While we use the risk-adjusted returns to estimate the regression coefficients for the main part of the paper, the results are substantially similar when we use alternative return definitions in Section 5.3.

To examine the pricing effect of $VOIB$ and $SVOIB$, we consider the following control variables.

1. OIB: Monthly order imbalance, defined as in Section 3.1.
2. $POIB$: Positive order imbalance, the logistic transform of the ratio of number of days with positive OIB to the total number of trading days in a month.
3. $SIZE$: Firm size measured as the natural logarithm of the market value of the firm's common equity (Banz, 1981).
4. BM: Book equity for the fiscal year-end in a calendar year divided by market equity at the end of December of that year, as in Fama and French (1992).
5. $R212$: The cumulative return on the stock over the eleven months ending at the beginning of the previous month (Jegadeesh and Titman, 1993).
6. $R1$: The lagged one month return (Jegadeesh, 1990).
7. $ILLIQ$: The Amihud illiquidity measure, defined as in Section 3.2.
8. $TURN$: Turnover ratio, defined as in Section 3.2.
9. $StdTURN$: Standard deviation of the monthly turnover over the past 36 months (Chordia, Subrahmanyam, and Anshuman, 2001).
10. $IVOL$: Idiosyncratic volatility, as in Ang, Hodrick, Xing, and Zhang (2006), computed as the standard deviation of the regression residual of the Fama and French (1993) three-factor model using daily data each month.

11. ACC: Accounting accruals, defined as the change in non-cash current assets, less the change in current liabilities (exclusive of short-term debt and taxes payable), less depreciation expense, all divided by average total assets (Sloan, 1996).

12. AG: Asset growth, as in Cooper, Gulen, and Schill (2008), computed as the year-on-year percentage change in total assets.

13. $ISSUE$: New issues, as in Pontiff and Woodgate (2008), measured as the change in shares outstanding from eleven months ago.

14. $PROFIT$: Profitability, as in Fama and French (2006), calculated as earnings divided by book equity, where earnings is defined as income before extraordinary items.

15. SUE: Standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, divided by its standard deviation estimated over the prior eight quarters. This is used as a proxy for earnings surprises, in order to analyze post-earnings-announcement-drift (PEAD) as in Bernard and Thomas (1989, 1990), and Ball and Brown (1968).

16. MAX: The maximum daily return in the previous month, as in Bali, Cakici, and Whitelaw (2011). This variable is included to capture the notion that large returns may be associated with extreme order imbalance.

17. $DISP$: Analyst earnings forecast dispersion, as in Diether, Malloy, and Scherbina (2002), computed as the standard deviation of annual earnings-per-share forecasts scaled by the absolute value of the average outstanding forecast. This variable is included to address the concern that $VOIB$ could potentially proxy for divergent opinions among investors.

18. $DISPD$: Dummy variable which equals to one if the stock is covered by at least two analysts and zero otherwise.[12]

[12] If there is no or only one analyst forecast in the I/B/E/S database, then DISP is set to zero. Such low coverage stocks account for 25% of the observations in the sample. To preserve a reasonable sample size, we include this dummy variable to indicate minimum analyst coverage instead of excluding all low coverage stocks.

19. $SSTT$: defined as the small-trade buyer-initiated turnover minus the small-trade seller-initiated turnover, measured over the previous six months. Hvidkjaer (2008) suggests that this measure proxies for trading by individual investors.

20. $HiLoSprd$: The bid-ask spread estimate from daily high and low prices as in Corwin and Schultz (2012).

21. PIN: The probability of informed trading as in Easley, Hvidkjaer and O'Hara (2002). The idea is to check whether our measures of adverse selection survive after including the measures existing in the literature.

22. $ILLIQV$: The volatility of idiosyncratic illiquidity as in Akbas, Armstrong and Petkova (2011) since it impacts expected returns.

23. Std_Ret: The standard deviation of daily returns computed each month.

24. $INSTV$: The average of the eight most recent quarterly absolute institutional ownership percentage changes.

To control for shocks to other liquidity and order imbalance variables, we also include innovations in OIB, $POIB$, $ILLIQ$, $TURN$, $IVOL$, Std_Ret and $StdTURN$ calculated in a manner similar to $SVOIB$ using the current value and its lagged six-month moving average, resulting in seven additional control variables termed $SOIB$, $SPOIB$, $SILLIQ$, $STURN$, $IVOL$, $SStd\_Ret$ and $SStdTURN$, respectively.[13] Since $SVOIB$ can potentially capture disagreement among investors, we also include shocks to the analyst dispersion, $SDISP$ in the regression. All control variables are cross-sectionally winsorized at the 0.5% and 99.5% levels.

5.2 Regression Results

Table 5 presents the time-series averages of coefficient estimates in the monthly cross-sectional regressions and the associated Newey-West adjusted t-statistics. The order imbalance is calculated

[13]We have also added shocks to $HiLoSprd$, computed in a manner similar to $SVOIB$; We have also added O/S, which is the ratio of option trading volume and stock trading volume, measured as in Roll, Schwartz and Subrahmanyam (2010). The main results are largely unaltered. The analysis is available in the internet appendix (Table A4).

using the number of shares traded in Columns 1 to 4 and using the number of trades in Columns 5 to 8. We report the single variable regressions results for $VOIB$ ($SVOIB$) in Columns 1 and 5 (Columns 2 and 6). The motivation behind studying the single variable FM regressions is to check whether our earlier portfolio results hold in the linear OLS framework. Multivariate regression results are presented in Columns 3, 4, 7 and 8. Note that the sample size in regressions in Columns 1-3 and 5-7 is an average of 2,944 per month. The sample size decreases to an average of 1,560 firms in full multivariate regressions due to the data requirements for the different variables in the monthly cross-sectional regressions.

We examine the $VOIB$ results first. In Column 1, $VOIB\_SHR$ has a coefficient estimate of 1.09 with a t-statistic of 3.02. The regression using $VOIB\_NUM$ in Column 5 generates similar results with a coefficient estimate of 1.08 (t-statistic = 2.94). $VOIB$ remains significant at the 1% in Columns 3, 4, 7 and 8. In economic terms, a one-standard deviation increase in $VOIB\_SHR$ ($VOIB\_NUM$) this month leads to an increase of 20 (18) basis points in next month's return in the regression of Column 1 (5); an increase of 66 basis points in the regression in Columns 4 and 8.

The FM results for $SVOIB$ in Columns 2 and 6 are consistent with the portfolio sorts presented earlier. The coefficient estimate of $SVOIB\_SHR$ ($SVOIB\_NUM$) is negative and significant at the 1% level. In Columns 4 and 8 the $SVOIB$ coefficient estimates are -4.74 and -5.62 with t-statistics of -5.84 and -7.34, respectively. In economic terms, a one-standard deviation increase in $SVOIB$ this month leads to a decrease of 34 to 48 basis points in next month's return depending on the specification of $SVOIB$ used. The negative coefficients suggest that a shock that increases order flow volatility is accompanied by negative returns next month in the cross section. While $SILLIQ$ is insignificant in Columns 3 and 6, both $ILLIQ$ and $SILLIQ$ are significant in the full multivariate regressions in Columns 4 and 8. Further, shocks to turnover ($STURN$) are also significant in Columns 4 and 8.[14]

[14] In unreported analyses, we consider whether $VOIB$ and $SVOIB$ remain significant for subsamples separated by the cross-sectional median of size and institutional holdings. The logic is to ascertain whether $VOIB$ and $SVOIB$ are priced in the more visible firms, or only in "neglected" firms. We find that both $VOIB$ and $SVOIB$ remain significant in all of these subsamples, but $SILLIQ$ and $STURN$ lose significance for large firms. This points to the robustness of order imbalance volatility and its shocks as cross-sectional predictors of returns.

We now briefly summarize the results for other control variables. In Columns 4 and 8, quite a few firm-level characteristics have significant coefficients. The negative coefficient of the one month lagged return is consistent with the reversal effect documented by Jegadeesh (1990). The negative coefficient of analyst forecast dispersion is consistent with Diether, Malloy, and Scherbina (2002). The positive coefficient of turnover and shocks to turnover is consistent with Gervais, Kaniel, and Mingelgrin (2001). The negative coefficients of $SIZE$, ACC, $ISSUE$, and AG are also consistent with prior research as is the positive coefficient of SUE. Consistent with Corwin and Schultz (2012), we also find that the pricing effect of $HiloSprd$ is positive and significant. Surprisingly, PIN has a negative impact on returns, possibly due to the presence of the other variables proxying for adverse selection. The volatility of institutional holdings also has a negative impact on returns in the cross-section. Return volatility has a negative impact on the cross-section of returns and subsumes the impact of idiosyncratic volatility. The negative coefficient on $StdTURN$ is consistent with Chordia, Subrahmanyam and Anshuman (2001). It is important to include $StdTURN$ in the cross-sectional regressions because it could be related to the volatility in order imbalance. The cross-sectional correlations between $VOIB$ (measured using shares as well as number of trades) and $StdTURN$ average around -0.3, which is lower than the correlation between the two measures of $VOIB$ and turnover and bid-ask spreads (See Panel B of Table 1). Order imbalance volatility and shocks to order imbalance volatility have a robust impact on stock returns in the cross-section and this impact survives the inclusion of a long list of variables that could potentially proxy for illiquidity and /or adverse selection.

5.3 Robustness Checks

In this subsection, we show that the pricing effects of $VOIB$ and $SVOIB$ are robust by experimenting with different return definitions, order imbalance definitions, and subsample periods. We present results using the number of shares in Table 6 while the analysis using number of trades generates largely the same results.[15] The control variables are the same as in Column 4 of Table 5. We include the following robustness checks:

[15] We report only results of $SVOIB\_SHR$ for the rest of the paper. The results of $SVOIB\_NUM$ are similar and are available in the internet appendix (Table A5).

- In the first column of Table 6 we use raw returns instead of the risk adjusted returns as the dependent variable in order to investigate the robustness of our results to risk.
- To address bid-ask bounce, in the second column, we use open-to-close mid-quote returns as the dependent variable. The open-to-close mid-quote returns are computed using the opening bid-ask midpoint price on the first trading day of the month and the closing bid-ask midpoint price on the last trading day of the month, adjusted for dividends and stock splits.
- In Column 3, we use dollar trading volume to compute the order imbalance variables. We define OIB as the buyer-initiated minus seller-initiated dollar volume scaled by the total dollar volume. Then, the volatility of OIB and shocks to the volatility of OIB are computed as before.
- In order to show that the pricing effect of $SVOIB$ is robust to the choice of benchmark $VOIB$, in Column 4 (5), $SVOIB$ is calculated as the difference between the current month $VOIB$ and its three-month (twelve-month) moving average in the previous month and all the other shock variables are calculated using three-month (twelve-month) moving averages accordingly.
- In Column 6, we exclude the financial crisis years of 2008 and 2009 because the accentuated stock market volatility may generate outliers in the $VOIB$ and $SVOIB$ estimates, leading to potentially spurious results in the full sample analysis.
- Column 7 uses data after January 2001 only (post-decimalization period) and Column 8 presents results for the period before January 2001 (pre-decimalization period). We are interested in these two subperiods because Chordia, Roll, and Subrahmanyam (2011) have suggested that in the post-decimalization period, the stock market has become more efficient and institutions are trading smaller quantities (suggesting that small trades are not originating from retail traders alone) and we want to ascertain that our results on $VOIB$ and $SVOIB$ survive. Also, the Lee and Ready (1991) algorithm to sign orders

may be prone to more errors after decimalization. So dividing the sample into pre- and post-decimalization allows us to check the robustness of the results.

- To address potential liquidity biases in our tests, Column 9 uses a Weighted Least Squares (WLS) estimation following Asparouhova, Bessembinder, and Kalcheva (2010). Specifically, in the FM regressions, we use the gross return in the previous month $(1 + Ret_{it-1})$ as the weight of each stock in WLS. We then report the time-series averages of the estimated WLS coefficients, with Newey-West corrections for the standard errors.

- Finally, to account for the potential measurement error in OIB that might affect the interpretation of the coefficients of $VOIB$ and $SVOIB$, we sort all stocks into decile portfolios every day based on OIB and then assign the average portfolio OIB, denoted by $\widehat{OIB}$, to all the stocks in that portfolio. This new order imbalance variable should be less prone to signing error, as long as the errors are not materially correlated in the cross section. We calculate $VOIB$ and $SVOIB$ using $\widehat{OIB}$, and replicate the main regression analysis in Column 10.[16]

Consistent with Hypothesis 1, in Table 6, the coefficient estimates of $VOIB\_SHR$ are all significant at the 1%. The only exception is the pre-2001 sample where the coefficient on $VOIB\_SHR$ is significant at the 5% level. The coefficient estimates of $SVOIB\_SHR$ are all significant at the 1% level. While the coefficient of $ILLIQ$ is significant, those of $SILLIQ$ are not significant in Column 5 (where $SILLIQ$ is computed as the current measure of $ILLIQ$ less its past 12-month moving average) and during the pre-decimalization period as shown in Column 8 of Table 6. The coefficient estimates of $VOIB$ and of $SVOIB$ are also significantly lower in the pre-decimalization period than in the post-decimalization period. Possibly, given the prevalence of smaller trade sizes even by institutional investors (see Chordia, Roll and Subrahmanyam, 2011) in recent years, the smaller trades may also be informed and thus

[16]We also perform an alternative test that accounts for measurement error in OIB. In this test, we use 20 randomly formed portfolios as test assets every month, and use the portfolios' order flows to compute their $VOIB$ and $SVOIB$. We then run Fama-MacBeth regressions for the 20 portfolios, using equally-weighted open-close quote midpoint returns as in the second column of Table 6. We repeat this procedure 100 times, and continue to find overall evidence that shocks to and levels of order flow volatility are priced in the cross section. These results are available in the internet appendix (Table A6).

any order imbalance volatility is likely to be a better measure of adverse selection in the post-decimalization period. In the pre-decimalization period with larger depths, the larger trades were more likely to be informed and thus easier to identify leading to the $VOIB$ being a poorer measure of adverse selection. Note also that the coefficient on PIN is positive and significant only in the pre-decimalization period. This sample period is similar to the one in Easley, Hvidkjaer and O'Hara (2002) suggesting that post-decimalization, PIN may be a poor proxy for adverse selection.

Overall, the impact of $VOIB$ and $SVOIB$ on the cross section of returns is robust to a number of different specifications.

6 Limits to Arbitrage and the Pricing of Order Imbalance Volatility Shocks

Thus far we have presented robust evidence of the positive impact of order flow volatility and the negative impact of shocks to order flow volatility on the cross section of expected returns. The implicit notion is that agents are slow to react to such shocks, so that the impact of such shocks spills over to next month's prices. However, if arbitrageurs are cognizant of such information processing delays, they could arbitrage this delay. Could it be the case that the impact of these shocks on prices is stronger for stocks with high trading costs and high barriers to arbitrage such as small stocks, stocks with low institutional holding ($INST$), or stocks with high idiosyncratic volatility ($IVOL$)? These stocks are those where investors are likely to face limits of arbitrage.

We obtain institutional holdings from Thomson Reuters. In Table 7, we first sort stocks by the arbitrage variables (firm size, institutional holding, and idiosyncratic volatility) into high and low categories by median, each month, and then within each category we sort stocks by $SVOIB\_SHR$ into quintiles. Table 7 presents the long-short quintile portfolio return and the long-short alpha.

Panel A shows that for small firms, the long-short return differential (alpha) between the high and low $SVOIB\_SHR$ portfolio is -1.18% (-1.11%) per month, and both are significant at the

1% level. For large firms, the return differential (alpha) at −0.23% (−0.31%) though smaller is also statistically significant. The return differential is statistically lower (more negative) for small firms as compared to large firms (t-statistic = −6.38), and the same is true of the alpha of the differential (t-statistic = −5.82).[17] Panel A also presents the results for sorts on institutional holdings. The return differential across extreme $SVOIB$ portfolios is far larger for stocks with low institutional holdings than for those with high holdings although the return differentials and alphas across the $SVOIB$ quintiles for low and high institutional holdings are all statistically significant. The return differential is statistically lower (more negative) for firms with low institutional holdings as compared to firms with high institutional holdings (t-statistic = −3.93), and the same is true of the alpha of the differential (t-statistic = −3.90). Finally, Panel A provides the findings for stocks sorted on $IVOL$. Even though the long-short return differential and the alpha are significant for both the high and low $IVOL$ stocks, the return differential and the alpha are more than twice as large for the high $IVOL$ stocks. Further, the return differential and alpha are statistically higher (less negative) for the low $IVOL$ stocks.[18] In Panel B, we repeat the analysis in Panel A using $SVOIB\_NUM$ and find similar results.

In sum, the results are consistent with the notion that shocks to order flow volatility predict returns less strongly in larger stocks, stocks with higher institutional holdings, and stocks with lower $IVOL$, possibly because these stocks have lower arbitrage costs.

7 Dynamics of Shocks to Order Flow Volatility

The central prediction of the cost of capital hypothesis on asset prices is that investors pay lower prices for stocks with higher adverse selection costs. Therefore, a shock that increases such costs (measured here by a positive shock to order flow volatility) should lower the current price of the asset and thus increase the future expected return. Thus far, we have documented that $SVOIB$ negatively impacts the contemporaneous and the following month's return. This is consistent

[17]The cross-sectional standard deviation of $SVOIB$ is smaller in the large firms (0.123 for $SVOIB\_SHR$ and 0.097 for $SVOIB\_NUM$) than in the small firms (0.180 for $SVOIB\_SHR$ and 0.160 for $SVOIB\_NUM$).

[18]We find that shocks to order flow volatility significantly reduce the contemporaneous return for large stocks, stocks with high institutional holdings, and stocks with low idiosyncratic volatility. These results are available in the internet appendix (Table A7).

with the economic notion that shocks which decrease liquidity should result in lower concurrent prices; though the negative impact of the shock persists beyond the current month, possibly due to the limits to arbitrage as discussed in the previous section or due to delayed responses of investors to the shock, as suggested by BPST. Note, however, that a liquidity premium in asset prices should not only result in lower current prices but also eventually lead to higher expected returns, as stated in our Hypothesis 3. So, a relevant question is how long it takes for the impact of $SVOIB$ on returns to turn positive. This section examines the dynamic effects of the shocks to $VOIB$ on returns.

Panel A of Table 8 reports the univariate portfolio results as well as the FM coefficient estimates over time. For the portfolio results, we sort stocks into quintile portfolios based on $SVOIB\_SHR$ and $SILLIQ$ and report the alphas of the long-short portfolios over the next one month, months 2-3, 4-6, 7-9 and 10-12.[19] When sorting on $SVOIB\_SHR$, the long-short alphas are essentially zero over months 2-3 but they are positive and statistically significant over months 4-6, 7-9 and 10-12. This suggests that it takes about three months for the illiquidity shock to be absorbed into prices and for investors to start earning the illiquidity premium.

When sorting on $SILLIQ$ the negative impact of the initial shock is strong in month 1. It is also felt in months 2-3 (alpha=-0.64% per month, t-statistic=-4.47), after which the alphas are negative but insignificant in months 4-6 and 7-9. The alphas turn positive in months 10-12 but they are insignificantly different from zero. This result for $SILLIQ$, while consistent with BPST, suggests that liquidity shocks as measured by $SILLIQ$ do not cause a sufficient drop in prices such that eventually the illiquidity premium obtains.

Panel A of Table 8 also reports the FM coefficients of $SVOIB\_SHR$ and $SILLIQ$ as a measure of liquidity shocks with future returns as the dependent variables. All the control variables are included in the regressions as in Table 5 but are not reported for brevity. The FM coefficients of $SVOIB\_SHR$ are positive and significant for the future returns over months 4-6, 7-9 and 10-12. Therefore, our findings support Hypothesis 3 in that the premium obtains after three months. Interestingly, after including all the control variables in Table 5, the coefficients

[19]The alphas for the next one month are presented in Panel C in Table 3; the FM coefficients for the next one month are obtained from Column 4 in Table 5.

of $SILLIQ$ are significantly negative for months 2-3 and 4-6. In months 7-9, the coefficient estimate of $SILLIQ$ is negative but significant only at the 10% level and in months 10-12 the coefficient estimate while negative is not statistically significant.[20]

Panel B of Table 8 also presents the FM coefficients when the dependent variable is future returns and when all the control variables are included as in Table 5, for size-based sorts (where the sorting is done as in Section 6). Consider first the small stocks. For these stocks there is some evidence that the liquidity shock is not immediately impounded into prices but persists through months 2-3. For instance, the coefficient of the regression of future returns in months 2-3 on $SVOIB\_SHR$ is -0.67 with a t-statistic of -2.62. More importantly, the coefficient estimates in months 4-6, 7-9 and 10-12 are positive and significant, suggesting that, for these stocks, shocks to liquidity result in a positive illiquidity premium after three months. Turning now to large stocks, we see that there is no spillover of the negative impact of the shock to $VOIB$ beyond the first month. The coefficient on $SVOIB$ is -2 (t-statistic $= -1.84$) in month 1 but the impact does not turn significantly positive thereafter.[21] This is consistent with the notion that liquidity shocks are absorbed promptly for large companies. We obtain similar results for sorts by institutional holdings and $IVOL$. These results are presented in the internet appendix (Table A8, Panel B).

Our work raises at least two puzzles. First, it is interesting that it takes three months for the premium to turn positive; BPST have suggested that the slow incorporation of liquidity shocks into prices could be due to investor inattention (Hirshleifer and Teoh, 2003). Second, it also is intriguing that the coefficient of $SILLIQ$ (shocks to Amihud illiquidity) does not turn positive even after 12 months. This finding suggests that shocks to $ILLIQ$ have an initial negative impact on future returns, but the effect does not convert to a standard liquidity premium in the longer term, unlike our $SVOIB$ coefficients. These issues deserve emphasis in future research. Overall, however, the pricing of order flow volatility shocks ($SVOIB$) in the cross section is

[20] For $SVOIB\_NUM$, the alphas (FM coefficients) are positive and significant for the future returns over months 4-6, 7-9 and 10-12 (months 7-9 and 10-12). The results are presented in the internet appendix (Table A8, Panel A).

[21] Paired t-tests show that the estimated $SVOIB$ coefficient of small stocks is significantly more negative than that of large stocks in month 1 and becomes significantly more positive than that of large stocks in months 7-9 and 10-12.

robust, and exhibits reliable economic and statistical significance.

8 Order Flow Volatility around Corporate Events

Thus far, we have postulated that order flow volatility is a proxy for asymmetric information. We now conduct tests to empirically validate this assumption. We focus on the period prior to earnings announcements or merger and acquisition (M&A) announcements when information asymmetry is likely to be high (see, for example, Baruch, Panayides, and Venkataraman, 2016). More specifically we compare $VOIB$ in event periods to $VOIB$ in non-event periods. We collect quarterly earnings announcement data from the I/B/E/S database and the M&A deal information between 1983 and 2012 from the SDC database, for the period from 1983 to 2012. For the earnings announcements, the event period is defined as trading days -18 to $+2$ relative to the announcement day, and the non-event period is defined as all the other days between days -31 and $+31$ relative to the announcement day. We include two days after the announcement in the event period because both theoretical arguments (Kim and Verrecchia, 1994) and empirical evidence (Lee, Mucklow and Ready, 1993; Krinsky and Lee, 1996) indicate that earnings announcements increase information asymmetry before and after the announcement periods. For M&A announcements, the event (non-event) period is defined as days -30 to -1 (-60 to -31) relative to the announcement day.

We first investigate earnings announcements in Panel A of Table 9. After merging the I/B/E/S data with daily order imbalance calculated using ISSM and TAQ data, we have 287,783 announcements in the sample. The mean $VOIB\_SHR$ is 0.385 (0.306) in the event (non-event) periods and the mean $VOIB\_NUM$ is 0.319 (0.256) in the event (non-event) periods. The difference in order imbalance volatility between the event and non-event periods reaches 0.079 (0.063) for $VOIB\_SHR$ ($VOIB\_NUM$), which amounts to about 20% of the mean order imbalance volatility and is statistically significant at the 1% level in each case. The fact that order imbalance volatility is significantly elevated in the period before an earnings announcement is consistent with its theoretical relation to information asymmetry.

Further, we compare the $VOIB$ difference between earnings and non-earnings periods for low and high absolute earnings surprises. The idea here is that the $VOIB$ difference should be larger when the absolute earnings surprise is higher because a higher absolute earnings surprise signals more uncertainty and thus greater information asymmetry. We calculate the earnings surprise as the difference between the actual value and the median analyst forecast in I/B/E/S scaled by the market price at the end of the month preceding the earnings announcement. We lose 598 observations because of missing median analyst forecasts. We sort the remaining 287,185 observations into high (N = 143,595) and low (N = 143,590) groups based on the median absolute earnings surprise. We find that the $VOIB\_SHR$ difference between event and non-event periods is 0.075 (0.083) when the absolute earnings surprise is low (high) while the $VOIB\_NUM$ difference is 0.058 (0.069). The difference-in-difference is significant at 1% level for both $VOIB\_SHR$ and $VOIB\_NUM$. In other words, $VOIB$ is more elevated just prior to an earnings announcement when the subsequent earnings surprise is large than when the surprise is small.

Table 9, Panel B reports $VOIB$ results for M&A events. Out of 279,895 records we extract from the SDC database, 27,767 deals involve a target firm being a public company that has price information available on CRSP and transactions data on either ISSM or TAQ. We focus on meaningful deals above one million dollars with at least 50% of shares sought by the acquirer. The final sample contains 5,209 deals. The mean target $VOIB\_SHR$ is 0.492 (0.480) in M&A event (non-event) period and the mean $VOIB\_NUM$ is 0.432 (0.423). The difference in order imbalance volatility between the event and non-event periods is a statistically significant 0.012 (0.009) for $VOIB\_SHR$ ($VOIB\_NUM$). Conjecturing that high takeover premium deals should attract more informed traders and have greater increase in $VOIB$, in Panel B, we also report the average $VOIB$ difference between M&A event and non-event periods for deals with the one-day takeover premium below the median (2,366 observations) and above the median (2,364 observations) separately. The $VOIB\_SHR$ difference between M&A and non-M&A periods is 0.007 (0.016) when the takeover premium is low (high) and the $VOIB\_NUM$ difference is 0.005 (0.012). The difference-in-difference between the two groups is significant at the 5% (10%) level for $VOIB\_SHR$ ($VOIB\_NUM$).

In Panel C, we run Fama-MacBeth regressions of the risk-adjusted returns on an earnings and M&A announcement dummy, $VOIB$, $SVOIB$, and interaction terms between the dummy and $VOIB$ as well as $SVOIB$, with the full set of control variables in Column 4 of Table 5. The event dummy is equal to one if there is an earnings or M&A announcement in the following month, and zero otherwise. We report the coefficient estimates of the variables of interest only in Panel C for brevity. We find that both $VOIB$ and the interaction term are positive and significant at the 1% level. Similarly, both $SVOIB$ and its interaction term are negative, and significant at the 10% level or better. Therefore, the impact of $VOIB$ and $SVOIB$ on the cross section of expected returns is stronger for periods when information asymmetry is likely to be high.

In summary, we find empirical support for the proposition that order flow volatility is a proxy for information asymmetry in event studies that use two material informational events: earnings and M&A announcements.

9 Limit Order Book Imbalance Volatility

Collin-Dufresne and Fos (CF) (2015) show that measures of adverse selection that are based mostly on market orders may not capture the presence of informed trading. They find that insiders make extensive use of limit orders for exploiting "long-lived" (CF, p. 1563) information signals, as indicated by 13(d) filings, which require reporting within 10 days of acquiring more than 5% of outstanding shares in a firm. However, it is worth noting that with quickly-perishable, short-lived information, agents who are informed may prefer to use market orders, rather than limit orders. So informed agents may use market orders or limit orders, depending on the nature of the information signal they possess. In this section, we study an order imbalance volatility using limit orders ($VOIB\_LOB$) calculated similarly to our main measure, $VOIB$.

We acquire limit order book data from LOBSTER database.[22] The data start from the 27th of June 2007. They contain most stocks listed on NYSE, AMEX and NASDAQ and cover about 95% of stocks during the period July 2007-December 2012. Since the dataset is extremely large, we parsimoniously take snapshots of the limit order book every five minutes during trading

[22]See https://lobsterdata.com for detailed information.

hours. The first three steps of limit buy and sell orders (starting from the best buy or sell price) are aggregated to the total buy and sell volumes for each snapshot. We then calculate the average buy and sell shares available to trade for all snapshots taken on a day and the imbalance between them as the daily limit order imbalance (OIB_LOB). $VOIB\_LOB$ is the standard deviation of OIB_LOB.

We initially compare the order imbalance volatility patterns around earnings announcements and M&A announcements periods in Table 10, Panel A in this subsample period similarly to the analysis in Table 9. For earnings announcements, we find that the difference in $VOIB\_SHR$ ($VOIB\_NUM$) between the event and non-event periods is a statistically significant 0.064 (0.057). However, the difference in $VOIB\_LOB$ is not significant. For M&A announcements, the volatility difference between the event and non-event periods is not significant for $VOIB$ or $VOIB\_LOB$.[23]

We next examine the pricing effect of $VOIB\_LOB$ and $SVOIB\_LOB$. Both $VOIB\_LOB$ and $SVOIB\_LOB$ are computed in a manner analogous to $VOIB$ and $SVOIB$. In Panel B, we sort all stocks in each month t into quintile portfolios based on $VOIB\_SHR$, $SVOIB\_SHR$, $VOIB\_NUM$, $SVOIB\_NUM$, $VOIB\_LOB$ or $SVOIB\_LOB$. The equally-weighted portfolio returns for month $t + 1$ are reported. Panel B also reports the return differences between the high and low quintiles and the alphas with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor. For the $VOIB$ used in earlier sections, the return differences between the top and bottom quintiles and the alphas are positive and significant. The long-short portfolio alpha amounts to 0.63% (0.66%) per month when sorting on $VOIB\_SHR$ ($VOIB\_NUM$). For $SVOIB\_SHR$ ($SVOIB\_NUM$), the monthly return difference is -0.81% (-0.87%) and the alpha is -0.74% (-0.82%). Both the return differences and the alphas are significant at the 1% level. The last two columns of Panel B present results for $VOIB\_LOB$ and $SVOIB\_LOB$. We find that the return differences and alphas for $VOIB\_LOB$ and $SVOIB\_LOB$ are not significant. To further ascertain pricing effects of $VOIB\_LOB$ and $SVOIB\_LOB$, we also run Fama-MacBeth regressions. The results

[23] The insignificance may be due to the size of the M&A sample. During the sample period from July 2007 to December 2012, there are 65,796 earnings event observations, but only 611 M&A observations.

are presented in Panel C. In Columns 1 and 2, the coefficient estimates of $VOIB\_LOB$ and $SVOIB\_LOB$ are not significant in univariate regressions. The regression using $VOIB\_LOB$ and $SVOIB\_LOB$ in Column 3 also generates insignificant results for these variables. In Column 4 (Column 5), we add $VOIB\_SHR$ ($VOIB\_NUM$) and $SVOIB\_SHR$ ($SVOIB\_NUM$), while including the control variables in Column 4 of Table 5. The results show that while $VOIB\_LOB$ and $SVOIB\_LOB$ are insignificant return predictors, $VOIB$ and $SVOIB$ (using number of trades or shares traded to measure order imbalances) remain significant.

Importantly, CF's point that informed agents use limit orders for long-lived information is not invalidated by our analysis. Instead, our measure of order flow volatility based on executed trades accords with a proxy for informed trading on short-lived information around corporate announcements. Further, our proxy is also a robust determinant of equity costs of capital in the cross section.

10 Conclusion

While it is challenging to measure adverse selection costs in financial markets, other endogenous parameters that are related to such costs might be easier to measure and thus might shed additional light on the link between trading costs and asset prices. We consider the notion that both information asymmetry costs and the volatility of order flow are driven by the same exogenous parameters in models of informed trading such as that of Kyle (1984, 1985) and its extension in Subrahmanyam (1991). We document that order flow volatility is higher prior to earnings announcements or M&A announcements when adverse selection is likely to be higher, indicating that order flow volatility is indeed related to adverse selection.

Order flow volatility is positively related to illiquidity proxies in the cross section and is priced in the cross section. Shocks to order flow volatility are strongly and positively related to illiquidity innovations, and are also strongly negatively related to both current and next month's returns even after controlling for various characteristics, including the level of and shocks to the Amihud (2002) illiquidity measure. These new findings are consistent with the notion that positive shocks to order flow volatility proxy for an increase in the true (unobserved) adverse selection costs which translates to a drop in current prices. We also find that in the longer run (between 4 and 12 months), a shock to order flow volatility does have a positive impact on returns, consistent with a premium for information asymmetry.

Our study points to several topics for research. First, why the additional premium induced by an order flow volatility shock takes a relatively long time to manifest itself in equity prices needs further exploration. Second, the link between order flow volatility and trading costs could be explored in other contexts. Since the link primarily emanates via informed trading, markets with less informed trading, such as those for index ETFs, with minimal firm-specific informed trading, might exhibit a more modest link between order flow volatility and liquidity. Third, the dynamics of order flow volatility might serve as a proxy for intertemporal variations in informed trading, which might, in turn, shed light on when prices are more likely to be predictive of future fundamentals. These and other related topics are left for future research.

Appendix

Order Imbalance Volatility: The Theory

In this appendix, we provide a brief theoretical motivation for our study, and provide empirical evidence to support the analytics. The objective of this simple setting is to demonstrate an economic link between order flow volatility and adverse selection costs in financial markets.

Consider an asset that is traded in a one-period Kyle (1984, 1985) and Subrahmanyam (1991) market at date 1, and pays off $F = \bar{F} + \sum_i^N \delta_i$ at date 2, where δ_i's are i.i.d. with mean zero and variance v_δ, and $\bar{F}$ is non-stochastic. Assume that there are N informed traders, and that each receives a signal δ_i.[24] The total noise trade is $z \sim N(0, v_z)$, and z, ϵ_i, and δ are each independent of all other random variables. At date 1, the informed traders and the noise traders submit market orders to a market maker, who sets the price while only observing the combined (net) order flow from all of the agents. This order flow is denoted by ω.

We consider the standard linear equilibrium in this setting where informed traders use symmetric strategies. At date 1 the market maker sets the price according to the linear rule $P = \bar{F} + \lambda\omega$ where ω is the net order flow. Let x_i $(i = 1, \ldots, N)$ denote the order of informed trader i, so that

$$\omega = \sum_{i=1}^{N} x_i + z. \tag{4}$$

Suppose the i'th informed trader conjectures that other informed traders use linear strategies of the form $\bar{\beta}(\delta_i)$. The trader chooses x_i to maximize $E[\{x_i(F - P)\}|\delta_i]$. Taking expectations and differentiating with respect to x_i, we obtain

$$x_i = \frac{\delta_i}{2\lambda}, \tag{5}$$

so that the informed strategy is of the form $\beta\delta_i$. In a symmetric Nash equilibrium $\bar{\beta} = \beta$. From Eq. (5) we then have

$$\beta = \frac{1}{2\lambda}. \tag{6}$$

[24] Alternative specifications of information do not change our basic intuition that order flow volatility is higher when the information asymmetry problem is greater.

Since all informed strategies are linear functions of the normally distributed signals, and the noise trade is normally distributed, the order flow ω also follows a normal distribution. Indeed, from Eq. (4), we have that $\omega = \sum_{i=1}^{N} \beta\delta_i + z$. Thus,

$$\omega = \frac{\sum_{i=1}^{N} \delta_i}{2\lambda} + z. \tag{7}$$

We assume that the market maker is risk averse and has negative exponential utility with coefficient A.[25] As in Subrahmanyam (1991), the market maker earns the autarky utility (due to Bertrand competition), i.e., the utility he would earn from not making the market. Given exponential utility and normally distributed order flow, the market maker's utility can be represented in mean-variance format. We then have that

$$E[\omega(P-F)|\omega] - \frac{A}{2}\text{var}[\omega(P-F)|\omega] = 0.$$

This implies that λ is given by

$$\lambda = \nu + (A/2)\text{var}(F|\omega) \tag{8}$$

where ν is the regression coefficient of F on ω. Substituting for ω from Eq. (7) above, we get a quadratic equation in λ. We choose the positive root, since, from Eq. (5), a positive λ is required to satisfy the second order condition for the informed trader.[26] We thus obtain the following expression for λ:

$$\lambda = \frac{A}{2} + \frac{1}{2}\sqrt{A^2 + \frac{Nv_\delta}{v_z}}. \tag{9}$$

The ex ante expected losses of the noise traders (i.e., the adverse selection costs of trading), denoted by L, are given by $E[z(P-F)] = \lambda v_z$, which implies that

$$L = \frac{Av_z}{2} + \frac{1}{2}\sqrt{A^2 v_z^2 + Nv_\delta v_z}.$$

Note that L is increasing in N and v_δ, the parameters representing private information, as well as A and v_z.

[25]Making the informed traders also risk averse, again, does not change our basic intuition.

[26]Kyle (1984) and Chordia, Huh, and Subrahmanyam (2009) provide further details on the derivation of the closed-form expression for λ.

Now, from Eq. (7), we find that the closed-form expression for variance of the order flow, denoted v_ω, is:

$$v_\omega = \frac{v_\delta}{\theta} + v_z \tag{10}$$

where

$$\theta \equiv \frac{2A^2}{N} + \frac{v_\delta}{v_z} + 2A\sqrt{\left[\frac{A}{N}\right]^2 + \frac{v_\delta}{N v_z}}$$

While A influences L and v_ω in opposite directions, it is evident that both L and v_ω are increasing in N, v_δ, and v_z. Since N and v_δ represent the extent of information asymmetry, both L and v_ω are positively related to the extent of adverse selection in the market. Indeed, holding market maker risk aversion constant, changes in all the other exogenous parameters affect adverse selection costs and the volatility of order flow in the same direction. Thus, controlling for A, if adverse selection costs command a premium in asset returns, so should the variability of order flow.[27] To test these implications, each month, we regress the monthly standard deviation of daily order imbalance on proxies for A, N, v_δ, and v_z in the cross section.

A reasonable proxy for A is market capitalization. The reasoning is as follows. We expect institutions to be active in market making because designated dealers such as specialists participate in only 10% of transactions during the bulk of our sample period (United States Government Accountability Office, 2005). Further, as Gompers and Metrick (2001) show, larger firms are held more heavily by institutions than individuals, and the former are likely to be well-capitalized (and thus less risk averse) in their *de facto* market making activities (viz. Cheng et al., 2015). By controlling for market capitalization in our regressions, we at least partially account for the cross-sectional variation in A.

The number of informed agents, N, is measured as the number of informed institutional investors. Following Abarbanell, Bushee and Raedy (2003), we categorize institutional investors

[27] Inventory costs also affect liquidity and can also be related to order flow. However, increased order flow volatility, in a dynamic setting, does not necessarily mean increased inventory risk. Thus, for example, an equal number of buys and sells imply order flow volatility but zero incremental inventory exposure, but a wave of equal buys in one direction imply zero order flow volatility but considerable inventory exposure. Modeling this issue requires an intertemporal setting, whereas our modeling, in a static setting, simply establishes the notion that informed trading-related parameters that make the market more illiquid (N and v_δ), also make the order flow more volatile. Our finding in Section 8, that order flow volatility rises prior to informational events, supports the link between informed trading and the volatility of order flows.

as informed and uninformed types, where the informed institutions are defined as investment companies and independent investment advisors because such institutions are more likely to be active investors. Other institutions, such as bank trusts, insurance companies, corporate/private pension funds, public pension funds, university and foundation endowments, have longer investment horizons and trade less actively.[28] Following Chordia, Huh and Subrahmanyam (2009), we employ earnings volatility as a proxy for v_δ, where earnings volatility is the standard deviation of earnings per share (EPS) from the most recent eight quarters.[29] Finally, we employ the average of daily dollar volume (in million dollars) as a proxy for v_z. We recognize that $VOIB$ likely depends on inputs beyond the ones we consider, but our proxies aim to capture cross-sectional determinants of $VOIB$ in an intuitive and parsimonious way.

The results are presented in the internet appendix (Table A1). We report the time-series averages of coefficient estimates from monthly cross-sectional regressions, together with the associated Newey-West (1987) t-statistics. We find that the average slope coefficients of the proxies for N, v_δ and v_z are positive. All of the average slope coefficients are statistically significant at the 1% level. Also, firm size is negative and significant, suggesting that small firms have higher levels of order flow variability. Thus, overall, the regression results accord with our analytical comparative statics.

[28]The data are obtained from Brian Bushee's website: http://acct3.wharton.upenn.edu/faculty/bushee/IIclass.html.

[29]Order flow variability is measured in standard deviation units, hence our proxy for v_δ is also measured as a standard deviation.

References

Abarbanell, J., B. Bushee, and J. Raedy, 2003, Institutional investor preferences and price pressure: The case of corporate spin-offs, *Journal of Business* 76, 233-261.

Acharya, V., and L. Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 385-410.

Akbas, Ferhat, R. Petkova, and and W. Armstrong, 2011, The Volatility of Liquidity and Expected Stock Returns. Working Paper, Case Western Reserve University.

Amihud, Y., 2002, Illiquidity and stock returns: cross-section and time series effects, *Journal of Financial Markets* 5, 31-56.

Amihud, Y., H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Arbel, A., S. Carvell, and P. Strebel, 1983, Giraffes, institutions and neglected firms, *Financial Analysts Journal* 39, 57-63.

Asparouhova, E., H. Bessembinder and I. Kalcheva, 2010, Liquidity biases in asset pricing tests, *Journal of Financial Economics* 96, 215-237.

Avramov, D., T. Chordia, 2006, Asset pricing models and financial market anomalies, *Review of Financial Studies* 19, 1001-1040.

Back, Kerry and Crotty, Kevin and Li, Tao, 2016, Estimating Information Asymmetry in Securities Markets. Working paper, Rice University.

Bali, T., L. Peng, Y. Shen, and Y. Tang, 2014, Liquidity shocks and stock market reactions, *Review of Financial Studies* 27, 1434-1485.

Bali, T., N. Cakici, and R. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Ball, R., and P. Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of*

Accounting Research 6, 159-178.

Baker, M., and J. Stein, 2004, Market liquidity as a sentiment indicator, *Journal of Financial Markets* 7, 271-299.

Banz, R., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Barber, B., T. Odean, and N. Zhu, 2009, Do retail traders move markets?, *Review of Financial Studies* 22, 151-186.

Baruch, S., M. Panayides, and K. Venkataraman, 2016, Informed trading and price discovery before corporate events, forthcoming, *Journal of Financial Economics.*

Bernard, V., and J. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or characteristic premium?, *Journal of Accounting Research* 27, 1-36.

Bernard, V., and J. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305-340.

Brennan, M., T. Chordia, A. Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345-373.

Brennan, M., and A. Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441-464.

Chakrabarty, B., P. Moulton, and A. Shkilko, 2012, Short sales, long sales, and the Lee-Ready trade classification algorithm revisited, *Journal of Financial Markets* 15, 467-491.

Cheng, S., A. Hameed, A. Subrahmanyam, and S. Titman, 2015, Short-term reversals: The effects of past returns and institutional exits, forthcoming, *Journal of Financial and Quantitative Analysis.*

Choi, J., L., Jin, and H. Yan, 2016, Informed Trading and Expected Returns. Working paper, Yale University.

Chordia, T., S. Huh, and A. Subrahmanyam, 2009, Theory-based illiquidity and asset pricing, *Review of Financial Studies* 22, 3629-3668.

Chordia, T., R. Roll, and A. Subrahmanyam, 2001, Market liquidity and trading activity, *Journal of Finance* 56, 501-530.

Chordia, T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111-130.

Chordia, T., R. Roll, and A. Subrahmanyam, 2005, Evidence on the speed of convergence to market efficiency, *Journal of Financial Economics* 76, 271-292.

Chordia, T., R. Roll, and A. Subrahmanyam, 2011, Recent trends in trading activity and market quality, *Journal of Financial Economics* 101, 243-263.

Chordia, T., and A. Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485-518.

Chordia, T., A. Subrahmanyam, and R. Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3-32.

Chordia, T., A. Subrahmanyam, and Q. Tong, 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* 58, 41-58.

Collin-Dufresne P., and V. Fos, 2015, Do Prices Reveal the Presence of Informed Trading?, *Journal of Finance* 70, 1555-1582.

Corwin, S., A., and P. Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* 67, 719-759.

Cooper, M., H. Gulen, and M. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Datar, V., N. Naik, and R. Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203-219.

Diether, K. B., C. J. Malloy, and A. Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

Easley, D., S. Hvidkjaer, and M. O'Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185-2221.

Easley, D., N. Kiefer, M. O'Hara, and J. Paperman, 1996, Liquidity, Information and Infrequently Traded Stocks, *Journal of Finance* 51, 1405-1436.

Easley, D., M. Lopez de Prado, and M. O'Hara, 2012, Flow Toxicity and Liquidity in a High-frequency World, *Review of Financial Studies* 25, 1457-1493.

Easley, D., and M. O'Hara, 2004, Information and the cost of capital, *Journal of Finance* 59, 1553-1583.

Ellis, K., R. Michaely, and M. O'Hara, 2000, The accuracy of trade classification rules: Evidence from Nasdaq, *Journal of Financial and Quantitative Analysis* 35, 529-551.

Fama, E., and K. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, E., and K. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491-518.

Fama, E., and J. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Gao, Xiaohui, and Jay Ritter, 2010, The marketing of seasoned equity offerings, *Journal of Financial Economics* 97, 33-52.

Gervais, S., and R. Kaniel, Mingelgrin, D., 2001, The high volume return premium, *Journal of Finance* 56, 877-919.

Gomes, J., L. Kogan, and L. Zhang, 2003, Equilibrium cross-section of returns, *Journal of Political Economy* 111, 693-732.

Gompers, P., and A. Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229-259.

Hasbrouck, J., 2009, Trading costs and returns for U.S. equities: Estimating effective costs from daily data, *Journal of Finance* 64, 1445-1477.

Hirshleifer, D., and S. H. Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337-386.

Holden, C., and S. Jacobson, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *Journal of Finance* 69, 1747-1785.

Hvidkjaer, S., 2008, Small trades and the cross-section of stock returns, *Review of Financial Studies* 21, 1123-1151.

Hwang, C., and X. Qian, 2011, Is Information Risk Priced? Evidence from the Price Discovery of Large Trades. Working paper, Nanyang Technological University.

Jacoby, G., D. Fowler, and A. Gottesman, 2000, The capital asset pricing model and the liquidity effect: A theoretical approach, *Journal of Financial Markets* 3, 69-81.

Jegadeesh, N., 1990, Evidence of predictable behavior in security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-92.

Johnson, T. and Eric So, 2015, A Simple Multimarket Measure of Information Asymmetry. Working paper, The University of Texas at Austin.

Jones, C., 2002, A century of stock market liquidity and trading costs, unpublished manuscript, Columbia University.

Kelly, B. and A. Ljungqvist, 2012, Testing Asymmetric-Information Asset Pricing Models, *Review of Financial Studies* 25, 1366-1413.

Kim, O., and R. Verrecchia, 1991, Trading Volume and Price Reactions to Public Announcements, *Journal of Accounting Research* 29, 302-221.

Kim, O., and R. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 41-67.

Krinsky, I., and J. Lee, 1996, Earnings announcements and the components of the bid-ask spread, *Journal of Finance* 51, 1523-1535.

Kyle, A., 1984, Market structure, information, futures markets, and price Formation, in Gary G. Storey, Andrew Schmitz and Alexander H. Sarris (ed.), *International Agricultural Trade: Advanced Readings in Price Formation, Market Structure, and Price Instability*, Westview Press, Boulder and London, 45-64.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Lee, C., B. Mucklow, and M. Ready, 1993, Spreads, depths, and the impact of earnings information: An intraday analysis, *Review of Financial Studies* 6, 345-374.

Lee, C., and B. Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, *Journal of Financial Markets* 3, 88-111.

Lee, C., and M. Ready, 1991, Inferring trade direction from intradaily data, *Journal of Finance* 46, 733-746.

Lo, A., and C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

Madhavan, A., D. Porter, and D. Weaver, 2005, Should securities markets be transparent? *Journal of Financial Markets* 8, 266-288.

Miller, E., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Newey, W., and K. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Odders-White, E., 2000, On the occurrence and consequences of inaccurate trade classification, *Journal of Financial Markets* 3, 259-286.

O'Hara, M., 2015, High frequency market microstructure, *Journal of Financial Economics* 116, 257-270.

Pástor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 113, 642-685.

Pontiff, J., and A. Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921-945.

Roll, R., E. Schwartz, and A. Subrahmanyam, 2010, O/S: The Relative Trading Activity in Options and Stock, *Journal of Financial Economics* 96, 1-17.

Sloan, R., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* 71, 289-315.

Subrahmanyam, A., 1991, Risk aversion, market liquidity, and price efficiency, *Review of Financial Studies* 4, 417-441.

United States Government Accountability Office, 2005, Decimal pricing has contributed to lower trading costs and a more challenging trading environment, GAO-05-535, submitted to Congressional Requesters, Washington, DC.

Yang, C., B. Zhang, and C. Zhang, 2015, Is Information Risk Priced? Evidence from Abnormal Idiosyncratic Volatility. Working paper, Hong Kong University of Science and Technology.

Table 1: Summary Statistics of Order Imbalance Variables

Panel A presents the time-series averages of the cross-sectional statistics for common stocks listed on NYSE, AMEX and NASDAQ from January 1983 to December 2012. The stock-month observation must have valid information to calculate the return, market capitalization, book-to-market ratio, and order imbalance, and must have the month-end price above one dollar. *OIB_SHR* is the monthly order imbalance defined as *(B-S)/(B+S)*, where *B (S)* is the number of shares traded initiated by buyers (sellers). *VOIB_SHR* is the standard deviation of daily order imbalance in a month. *SVOIB_SHR* is the difference between *VOIB_SHR* in the current month and the six-month moving average of *VOIB_SHR* in the previous month. The variables calculated using the number of trades are termed as *OIB_NUM*, *VOIB_NUM*, and *SVOIB_NUM*. Panels B and C present the time-series averages of the monthly cross-sectional correlations. The Amihud illiquidity (*ILLIQ*) is calculated as the monthly average of the daily ratio of the absolute return to the dollar volume. *TURN* is the logarithm of the monthly average of the daily turnover ratio calculated as the number of shares traded divided by shares outstanding. *SPRD* is the spread measure using the cheap alternative solution by Holden and Jacobsen (2014). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). *RET* is the monthly stock return. *Ret_Std* is the standard deviation of daily returns in a month. The shocks to the Amihud illiquidity (*SILLIQ*), turnover (*STURN*), spread (*SSPRD*), and return standard deviation (*SRet_Std*) are computed similarly to *SVOIB*. The corresponding z-statistics are reported in parentheses.

Panel A: Descriptive statistics

Statistics	N	Mean	St. dev.	Median	Minimum	Maximum
OIB_SHR	2,948	-0.051	0.325	-0.068	-0.782	0.642
VOIB_SHR	2,948	0.361	0.185	0.337	0.061	0.894
SVOIB_SHR	2,948	-0.001	0.102	-0.001	-0.463	0.461
OIB_NUM	2,948	-0.041	0.279	-0.051	-0.678	0.575
VOIB_NUM	2,948	0.287	0.164	0.253	0.047	0.985
SVOIB_NUM	2,948	-0.001	0.085	-0.003	-0.400	0.441

Panel B: Correlations with other liquidity measures

	ILLIQ	*TURN*	*SPRD*	*SILLIQ*	*STURN*	*SSPRD*	*PIN*	*RET*
VOIB_SHR	0.248	-0.514	0.496	0.005	-0.074	-0.005	0.549	-0.004
	(31.96)	(-30.27)	(30.72)	(0.90)	(-13.13)	(-0.28)	(25.77)	(-0.92)
lag*(VOIB_SHR)*	0.236	-0.494	0.487	-0.006	-0.019	-0.012	0.545	0.015
	(32.31)	(-30.4)	(30.18)	(-1.05)	(-3.77)	(-1.05)	(25.62)	(3.58)
SVOIB_SHR	0.060	-0.103	0.061	0.062	-0.235	0.047	0.034	-0.064
	(16.88)	(-17.81)	(10.09)	(14.68)	(-24.48)	(6.47)	(5.93)	(-16.80)
lag*(SVOIB_SHR)*	0.041	-0.068	0.059	0.033	-0.110	0.037	0.035	-0.012
	(14.23)	(-15.18)	(10.62)	(11.9)	(-24.33)	(5.89)	(6.73)	(-6.30)
VOIB_NUM	0.257	-0.497	0.512	0.003	-0.079	-0.014	0.580	0.005
	(30.12)	(-28.74)	(31.79)	(0.43)	(-13.78)	(-1.11)	(27.86)	(1.41)
lag*(VOIB_NUM)*	0.243	-0.475	0.504	-0.008	-0.019	-0.017	0.576	0.012
	(30.9)	(-28.85)	(31.46)	(-1.39)	(-3.77)	(-1.54)	(27.75)	(2.92)
SVOIB_NUM	0.068	-0.110	0.065	0.060	-0.253	0.027	0.038	-0.047
	(15.34)	(-17.1)	(9.39)	(11.33)	(-24.48)	(3.05)	(5.67)	(-12.03)
lag*(SVOIB_NUM)*	0.044	-0.066	0.062	0.026	-0.107	0.019	0.039	-0.015
	(12.38)	(-13.1)	(9.77)	(7.69)	(-22.94)	(2.52)	(6.53)	(-7.79)

Table 1 (continued)

Panel C: Correlations with return standard deviation						
	VOIB_SHR	*SVOIB_SHR*	*VOIB_NUM*	*SVOIB_NUM*	*Ret_Std*	*SRet_Std*
VOIB_SHR	1					
SVOIB_SHR	0.443	1				
	(37.95)					
VOIB_NUM	0.911	0.342	1			
	(67.37)	(39.45)				
SVOIB_NUM	0.344	0.789	0.446	1		
	(37.06)	(160.75)	(34.39)			
Ret_Std	0.061	-0.056	0.106	-0.034	1	
	(13.34)	(-10.13)	(27.38)	(-6.21)		
SRet_Std	-0.015	-0.084	0.000	-0.053	0.603	1
	(-4.17)	(-20.30)	(0.06)	(-12.11)	(56.95)	

Table 2: Portfolio Sorts for *VOIB*

In Panel A, for each month t from June 1983 to December 2012, we sort all stocks in the sample into quintile portfolios based on order imbalance volatility, and report the characteristics of the stocks in each portfolio. *SIZE* represents market capitalization in millions. *BM* is the book-to-market ratio. *RET* is the monthly stock return. *Ret_Std* is the standard deviation of daily returns in a month. In Panel B, we sort all stocks into quintile portfolios based on order imbalance, order imbalance volatility and other liquidity measures, and report the equally-weighted portfolio returns in month t + 1. O*IB_SHR* is the monthly order imbalance defined as *(B-S)/(B+S)*, where *B (S)* is the number of shares traded initiated by buyers (sellers). *VOIB_SHR* is the standard deviation of daily order imbalance in a month. Variables calculated using the number of trades are termed *OIB_NUM* and *VOIB_NUM*. *ILLIQ* represents the Amihud measure of illiquidity. *TURN* is the logarithm of the monthly average of the daily turnover ratio calculated as the number of shares traded divided by shares outstanding. *SPRD* is the spread measure using the cheap alternative solution by Holden and Jacobsen (2014). Also reported are the return differences between the high and low quintiles and the alphas with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor. Newey-West *t*-statistics are reported in parentheses. In Panel C, we first sort stocks into high and low groups based on *ILLIQ*, *TURN*, and *SPRD* separately, and then sort on *VOIB_SHR* into quintile portfolios in each group at month t. Portfolio returns and return differences in month t + 1 are reported. In Panel D, we perform the double sorting analysis using *VOIB_NUM*. All returns are reported in percent. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Quintile portfolio characteristics

	VOIB_SHR				*VOIB_NUM*			
Quintile	*SIZE*	*BM*	lag(*RET*)	lag(*Ret_Std*)	*SIZE*	*BM*	lag(*RET*)	lag(*Ret_Std*)
Low-1	6907.394	0.553	1.863	0.021	7234.333	0.562	1.509	0.019
2	1846.671	0.622	1.529	0.023	1639.776	0.624	1.637	0.023
3	842.499	0.685	1.403	0.025	747.849	0.678	1.463	0.025
4	407.171	0.781	1.149	0.026	367.079	0.783	1.220	0.028
High-5	185.406	0.978	0.849	0.028	200.452	0.972	0.966	0.028
High-Low	-6738.390***	0.426***	-1.014***	0.006***	-7050.95***	0.411***	-0.543***	0.009***
	(-30.72)	(23.37)	(-4.86)	(20.65)	(-31.65)	(22.17)	(-3.28)	(35.14)

Panel B: Univarite sorts

Quintile	*OIB_SHR*	*OIB_NUM*	*VOIB_SHR*	*VOIB_NUM*	*ILLIQ*	*TURN*	*SPRD*
Low-1	1.496	1.506	0.831	0.873	0.96	1.203	0.98
2	1.318	1.257	1.137	1.12	1.134	1.436	1.134
3	1.181	1.115	1.204	1.256	1.081	1.348	1.151
4	1.146	1.155	1.514	1.457	1.308	1.328	1.264
High-5	1.159	1.267	1.613	1.594	1.817	1.007	1.885
High-Low	-0.336***	-0.238	0.782***	0.722***	0.857***	-0.196	0.905***
	(-2.81)	(-1.51)	(4.27)	(4.56)	(3.60)	(-0.79)	(3.35)
Alpha	-0.465***	-0.375*	0.992***	0.972***	1.146***	-0.458**	1.202***
	(-3.12)	(-1.93)	(5.18)	(5.83)	(5.57)	(-2.24)	(5.17)

Table 2 (continued)

Panel C: Bivariate sorts on *VOIB_SHR*						
	ILLIQ		*TURN*		*SPRD*	
Quintile	Low	High	Low	High	Low	High
Low-1	0.788	1.01	0.955	0.723	0.838	0.976
2	1.125	1.467	1.23	1.089	1.151	1.482
3	1.075	1.565	1.371	1.066	1.095	1.624
4	1.155	1.513	1.474	1.23	1.122	1.577
High-5	1.186	1.713	1.66	1.847	1.183	1.777
High-Low	0.398**	0.703**	0.705***	1.124***	0.345**	0.801***
	(2.17)	(2.43)	(4.69)	(5.76)	(2.18)	(2.78)
Alpha	0.405***	0.823***	0.924***	1.158***	0.354**	0.927***
	(2.65)	(2.67)	(5.27)	(6.54)	(2.24)	(2.89)

Panel D: Bivariate sorts on *VOIB_NUM*						
	ILLIQ		*TURN*		*SPRD*	
Quintile	Low	High	Low	High	Low	High
Low-1	0.869	1.066	0.978	0.834	0.902	1.054
2	1.062	1.459	1.22	1.076	1.076	1.487
3	1.109	1.564	1.404	1.107	1.15	1.605
4	1.168	1.496	1.406	1.221	1.14	1.542
High-5	1.123	1.684	1.682	1.717	1.122	1.749
High-Low	0.254**	0.619**	0.704***	0.883***	0.220**	0.695***
	(2.21)	(2.26)	(4.45)	(3.83)	(1.97)	(2.61)
Alpha	0.310***	0.729**	0.947***	0.959***	0.266**	0.823***
	(3.17)	(2.41)	(5.17)	(5.09)	(2.57)	(2.66)

Table 3: Portfolio Sorts for Liquidity Shocks

In Panel A, for each month t from June 1983 to December 2012, we sort all stocks in the sample into quintile portfolios based on *SVOIB_SHR* and *SVOIB_NUM*, and report the characteristics of the stocks in each portfolio. *SIZE* represents market capitalization in millions. *BM* is the book-to-market ratio. *RET* is the monthly stock return. *Ret_Std* is the standard deviation of daily returns in a month. In Panel B, we sort all stocks in the sample into quintile portfolios based on *SVOIB_SHR*, *SVOIB_NUM*, *SILLIQ*, *STURN*, and *SSPRD*. The equally-weighted portfolio returns for month t are reported. Also reported are the return differences between the high and low quintiles and the alphas with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor. Newey-West *t*-statistics are reported in parentheses. Panel C reports the equally-weighted portfolio returns and alphas for month t+1. *VOIB_SHR* is the standard deviation of daily order imbalance in a month, where the order imbalance is defined as *(B-S)/(B+S)* with *B (S)* being the number of shares traded initiated by buyers (sellers). *SVOIB_SHR* is the difference between *VOIB_SHR* in the current month and the six-month moving average of *VOIB_SHR* in the previous month. The shock to order flow volatility calculated using the number of trades is termed *SVOIB_NUM*. *ILLIQ* represents the Amihud measure of illiquidity. *TURN* is the logarithm of the monthly average of the daily turnover ratio calculated as the number of shares traded divided by shares outstanding. *SPRD* is the spread measure using the cheap alternative solution by Holden and Jacobsen (2014). *SILLIQ*, *STURN*, and *SSPRD* are the shocks to *ILLIQ*, *TURN*, and *SPRD* calculated similarly to *SVOIB*. All returns are reported in percent. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Quintile portfolio characteristics

	SVOIB_SHR				*SVOIB_NUM*			
Quintile	*SIZE*	*BM*	lag(*RET*)	lag(*Ret_Std*)	*SIZE*	*BM*	lag(*RET*)	lag(*Ret_Std*)
Low-1	888.141	0.797	3.617	0.027	773.247	0.805	3.359	0.027
2	2514.686	0.678	1.553	0.023	2,404.570	0.674	1.558	0.023
3	3526.090	0.657	1.031	0.023	3,797.441	0.651	1.054	0.022
4	2356.707	0.690	0.490	0.024	2,409.267	0.680	0.585	0.023
High-5	898.332	0.796	0.113	0.026	799.227	0.809	0.249	0.027
High-Low	10.202	-0.000	-3.504***	-0.001***	26.023	0.005	-3.109***	-0.001***
	(0.29)	(-0.017)	(-17.42)	(-2.72)	(0.72)	(0.59)	(-15.86)	(-3.19)

Panel B: Contemporaneous returns

Quintile	*SVOIB_SHR*	*SVOIB_NUM*	*SILLIQ*	*STURN*	*SSPRD*
Low-1	3.233	2.891	4.734	-1.13	5.238
2	1.415	1.38	2.685	0.192	2.578
3	1.037	1.022	1.259	0.907	1.154
4	0.816	0.862	-0.215	1.778	-0.233
High-5	0.550	0.896	-1.412	5.375	-1.636
High-Low	-2.684***	-1.995***	-6.146***	6.505***	-6.874***
	(-8.28)	(-6.26)	(-25.36)	(14.51)	(-17.37)
Alpha	-2.530***	-1.911***	-5.987***	6.094***	-6.589***
	(-8.70)	(-6.65)	(-24.01)	(15.35)	(-16.88)

Panel C: Next month' returns

Quintile	*SVOIB_SHR*	*SVOIB_NUM*	*SILLIQ*	*STURN*	*SSPRD*
Low-1	1.829	1.934	2.326	0.724	1.979
2	1.180	1.159	1.264	1.028	1.183
3	1.117	1.104	1.097	1.192	1.037
4	1.139	1.115	0.819	1.374	1.119
High-5	1.035	0.989	0.795	2.033	1.091
High-Low	-0.793***	-0.945***	-1.530***	1.308***	-0.888***
	(-6.50)	(-7.49)	(-8.94)	(8.13)	(-4.87)
Alpha	-0.816***	-0.968***	-1.534***	1.325***	-1.015***
	(-6.45)	(-7.84)	(-9.08)	(8.16)	(-5.01)

Table 4: Bivariate Portfolio Sorts Based on SVOIB and Other Liquidity Shocks

In Panel A (B), for each month t from July 1983 to December 2012, we first sort stocks into quintile portfolios based on *SILLIQ, STURN, SSPRD, RET, DISP* separately, and then sort on *SVOIB_SHR* (*SVOIB_NUM*) into quintile portfolios in each group. The equally-weighted portfolio returns in month t+1 are reported. The return difference between the high and low quintiles and the alpha with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor are also reported with Newey-West *t*-statistics in parentheses. *VOIB_SHR* is the standard deviation of daily order imbalance in a month, where the order imbalance is defined as *(B-S)/(B+S)* with *B (S)* being the number of shares traded initiated by buyers (sellers). *SVOIB_SHR* is the difference between *VOIB_SHR* in the current month and the six-month moving average of *VOIB_SHR* in the previous month. The shock to order flow volatility calculated using the number of trades is termed *SVOIB_NUM*. *ILLIQ* represents the Amihud measure of illiquidity. *TURN* is the logarithm of the monthly average of the daily turnover ratio calculated as the number of shares traded divided by shares outstanding. *SPRD* is the spread measure using the cheap alternative solution by Holden and Jacobsen (2014). *DISP* is the analyst dispersion in earnings forecasts. *SILLIQ*, *STURN*, and *SSPRD* are the shocks to *ILLIQ*, *TURN*, and *SPRD* calculated similarly to *SVOIB*. All returns are reported in percent. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Bivariate sorts on *SVOIB_SHR*										
			SILLIQ					*STURN*		
Quintile	Low	2	3	4	High	Low	2	3	4	High
Low-1	2.893	1.537	1.135	0.978	1.134	1.020	1.283	1.427	1.699	2.984
2	2.157	1.329	0.970	0.862	0.773	0.688	0.952	1.208	1.291	1.863
3	2.344	1.161	1.006	0.880	0.734	0.832	1.020	1.023	1.207	1.682
4	2.115	1.241	1.208	0.798	0.857	0.650	0.981	1.079	1.287	1.701
High-5	2.122	1.054	1.165	0.575	0.479	0.432	0.903	1.226	1.388	1.937
High-Low	-0.771***	-0.483***	0.030	-0.404***	-0.655***	-0.589***	-0.380***	-0.200*	-0.311**	-1.047***
	(-4.23)	(-3.28)	(0.31)	(-3.32)	(-4.06)	(-4.11)	(-3.33)	(-1.66)	(-2.49)	(-5.64)
Alpha	-0.716***	-0.392***	0.137	-0.375***	-0.660***	-0.449***	-0.295**	-0.078	-0.198	-0.973***
	(-3.87)	(-2.68)	(1.40)	(-2.75)	(-3.79)	(-3.20)	(-2.49)	(-0.70)	(-1.57)	(-5.54)
			SSPRD					*RET*		
Quintile	Low	2	3	4	High	Low	2	3	4	High
Low-1	2.599	1.462	1.114	1.379	1.800	2.340	1.783	1.601	1.389	1.879
2	1.719	1.187	0.979	1.032	1.144	1.424	1.296	1.230	1.037	0.975
3	2.002	1.022	1.067	1.117	1.032	1.419	1.195	1.166	1.035	0.903
4	1.745	1.199	1.017	1.119	0.921	1.657	1.253	1.125	1.050	0.672
High-5	1.832	1.047	1.008	0.950	0.563	1.645	0.793	0.997	1.024	0.621
High-Low	-0.768***	-0.415***	-0.106	-0.429***	-1.237***	-0.696***	-0.990***	-0.604***	-0.366***	-1.257***
	(-3.73)	(-3.16)	(-1.03)	(-4.23)	(-7.75)	(-4.76)	(-7.30)	(-4.80)	(-2.91)	(-6.84)
Alpha	-0.707***	-0.309**	-0.058	-0.437***	-1.244***	-0.652***	-1.032***	-0.627***	-0.350**	-1.356***
	(-3.78)	(-2.21)	(-0.53)	(-4.04)	(-7.08)	(-4.59)	(-7.88)	(-5.26)	(-2.48)	(-7.10)
			DISP							
Quintile	Low	2	3	4	High					
Low-1	2.076	1.564	1.317	1.569	1.264					
2	1.291	1.358	1.080	1.032	0.798					
3	1.172	1.203	0.976	1.043	0.881					
4	1.230	1.221	1.034	1.114	0.752					
High-5	1.049	1.161	1.044	1.008	0.609					
High-Low	-1.028***	-0.395**	-0.273***	-0.561***	-0.655***					
	(-6.75)	(-2.26)	(-2.90)	(-4.78)	(-4.48)					
Alpha	-0.944***	-0.371**	-0.192*	-0.482***	-0.517***					
	(-6.19)	(-2.02)	(-1.83)	(-3.97)	(-3.83)					

Table 4 (continued)

Panel B: Bivariate sorts on *SVOIB_NUM*

	SILLIQ					*STURN*				
Quintile	Low	2	3	4	High	Low	2	3	4	High
Low-1	2.912	1.577	1.227	0.992	1.282	1.089	1.311	1.617	1.826	3.061
2	2.241	1.348	1.006	0.920	0.595	0.715	0.94	1.082	1.277	1.992
3	2.143	1.206	1.081	0.789	0.704	0.634	1.066	1.018	1.228	1.543
4	2.214	1.248	1.161	0.833	0.909	0.795	0.941	1.104	1.230	1.651
High-5	2.120	0.946	1.010	0.560	0.490	0.390	0.883	1.145	1.312	1.922
High-Low	-0.792***	-0.630***	-0.217**	-0.432***	-0.792***	-0.699***	-0.428***	-0.472***	-0.514***	-1.140***
	(-4.32)	(-5.89)	(-2.29)	(-3.24)	(-5.20)	(-5.56)	(-3.82)	(-4.44)	(-3.99)	(-6.27)
Alpha	-0.788***	-0.595***	-0.164	-0.382**	-0.839***	-0.572***	-0.368***	-0.444***	-0.453***	-1.106***
	(-4.27)	(-4.74)	(-1.54)	(-2.31)	(-5.37)	(-4.32)	(-2.71)	(-3.78)	(-3.11)	(-5.86)

	SSPRD					*RET*				
Quintile	Low	2	3	4	High	Low	2	3	4	High
Low-1	2.677	1.605	1.157	1.354	1.903	2.463	1.875	1.693	1.471	1.997
2	1.826	1.148	0.993	1.158	1.017	1.343	1.231	1.256	1.065	0.900
3	1.703	1.047	1.061	1.077	0.992	1.409	1.214	1.115	1.024	0.821
4	1.922	1.085	1.049	1.113	1.078	1.740	1.122	1.113	1.017	0.765
High-5	1.771	1.032	0.925	0.894	0.469	1.530	0.878	0.943	0.958	0.568
High-Low	-0.906***	-0.573***	-0.232**	-0.459***	-1.434***	-0.932***	-0.997***	-0.750***	-0.513***	-1.429***
	(-4.88)	(-4.46)	(-2.08)	(-3.92)	(-9.11)	(-6.40)	(-6.56)	(-6.02)	(-3.61)	(-7.44)
Alpha	-0.836***	-0.497***	-0.172	-0.420***	-1.463***	-0.848***	-0.913***	-0.705***	-0.429***	-1.310***
	(-4.75)	(-3.50)	(-1.37)	(-3.59)	(-9.34)	(-5.40)	(-6.02)	(-5.94)	(-2.98)	(-7.63)

	DISP				
Quintile	Low	2	3	4	High
Low-1	2.162	1.551	1.398	1.606	1.371
2	1.279	1.385	1.015	1.230	0.737
3	1.162	1.212	0.982	1.058	0.798
4	1.265	1.359	0.996	1.014	0.944
High-5	0.951	1.001	1.061	0.857	0.456
High-Low	-1.211***	-0.539***	-0.337***	-0.750***	-0.915***
	(-7.49)	(-2.64)	(-3.26)	(-5.53)	(-6.30)
Alpha	-1.211***	-0.514***	-0.269**	-0.653***	-0.788***
	(-7.48)	(-2.63)	(-2.40)	(-4.78)	(-5.25)

Table 5: Fama-MacBeth Regression Estimates

This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012. The dependent variable is the risk-adjusted return calculated using the Fama-French (1993) factors as well as the momentum factor and the liquidity factor of Pastor and Stambaugh (2003) with loadings conditional on the size and book-to market ratio. All independent variables (except *R1* and *R212*) are lagged one month. *OIB* is the monthly order imbalance defined as *(B-S)/(B+S)*, where *B (S)* is the trades initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *POIB* is the logistic transform of the ratio of number of days with positive order imbalance and total number of trading days in the month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. The order imbalance is calculated using the number of shares traded in Columns 1 to 4 and using the number of trades in Columns 5 to 8. *SIZE* represents the logarithm of market capitalization. *BM* is the logarithm of the book-to-market ratio. *R1* is the lagged one month return. *R212* is the cumulative returns over the second through the twelfth months prior to the current month. *TURN* is the logarithm of the monthly average turnover ratio calculated as the trading volume divided by shares outstanding. *StdTURN* is the standard deviation of *TURN* in the past 36 months. *ILLIQ* represents the Amihud measure of illiquidity. *ACC* represents accruals, measured as in Sloan (1996). *AG* is the asset growth computed in Cooper, Gulen and Shill (2008). *ISSUE* represents new issues as in Pontiff and Woodgate (2008). *IVOL* is the idiosyncratic volatility computed as in Ang, Hodrick, Xing, and Zhang (2006). *PROFIT* is the profitability variable as in Fama and French (2006). *SUE* is the standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. *MAX* is the maximum daily return in the last month. *DISP* is the analyst dispersion in earnings forecasts and *DISPD* is a dummy that equals to one if the stock is covered by at least two analysts and zero otherwise. *SSTT* is the small size trade imbalance as in Hvidkjaer (2008). *HiloSprd* is the high-low spread estimate of Corwin and Schultz (2012). *INSTV* is the average of the eight most recent quarterly absolute institutional ownership percentage changes. *ILLIQV* is the idiosyncratic volatility of liquidity in Akbas, Armstrong and Petkova (2011). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). *Ret_Std* is the standard deviation of daily returns in a month. *SOIB*, *SPOIB*, *STURN*, *SStdTURN*, *SIVOL*, *SILLIQ*, *SDISP*, and *SRet_Std* are defined similarly as *SVOIB*. All variables are winsorized at the 0.5% and 99.5% levels. *N* is the average number of stocks per month. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	SHR				*NUM*			
Model	1	2	3	4	5	6	7	8
Intercept	-0.214	0.179***	-0.390**	2.535***	-0.133	0.177***	-0.362***	2.263***
	(-1.62)	(2.80)	(-2.56)	(4.05)	(-1.21)	(2.79)	(-2.90)	(3.72)
VOIB	1.087***		1.193***	3.548***	1.076***		1.472***	4.033***
	(3.02)		(2.98)	(3.93)	(2.94)		(3.59)	(4.89)
SVOIB		-3.372***	-4.027***	-4.739***		-4.171***	-5.118***	-5.619***
		(-8.21)	(-10.13)	(-5.84)		(-9.21)	(-11.94)	(-7.34)
ILLIQ			1.825***	2.258***			1.826***	2.235***
			(3.81)	(4.96)			(3.77)	(4.97)
SILLIQ			-0.531	-0.979**			-0.512	-0.979**
			(-1.02)	(-2.30)			(-0.97)	(-2.34)
OIB				-0.448				-0.482
				(-1.21)				(-1.00)
SOIB				0.591				1.055**
				(1.63)				(2.34)
POIB				0.075				-0.033
				(1.13)				(-0.51)
SPOIB				-0.045				0.031
				(-0.64)				(0.49)
SIZE				-0.183***				-0.162***
				(-5.15)				(-4.66)
BM				-0.013				-0.009
				(-0.27)				(-0.18)
R212				-0.032				-0.029
				(-0.17)				(-0.15)
R1				-0.049***				-0.049***
				(-10.54)				(-10.51)

Table 5 (continued)

	SHR				*NUM*			
Model	1	2	3	4	5	6	7	8
TURN				0.141*				0.188**
				(1.70)				(2.22)
STURN				0.771***				0.695***
				(8.65)				(8.31)
STDTURN				-0.175***				-0.186***
				(-3.81)				(-4.05)
SSTDTURN				-0.182				-0.159
				(-1.30)				(-1.15)
IVOL				24.428				29.207
				(1.30)				(1.55)
SIVOL				-22.879				-26.216*
				(-1.54)				(-1.76)
ACC				-1.125***				-1.099***
				(-3.41)				(-3.32)
AG				-0.212***				-0.213***
				(-3.18)				(-3.27)
ISSUE				-1.203***				-1.233***
				(-4.06)				(-4.15)
PROFIT				0.068				0.051
				(1.06)				(0.80)
SUE				0.030**				0.029**
				(2.44)				(2.38)
Max				4.775***				4.884***
				(4.20)				(4.30)
DISP				-0.352***				-0.363***
				(-3.37)				(-3.49)
SDISP				0.087				0.090
				(0.74)				(0.78)
DISPD				-0.098				-0.120
				(-1.33)				(-1.63)
SSTT				9.319				0.055
				(0.81)				(0.00)
HiLoSprd				25.488***				27.167***
				(3.31)				(3.48)
INSTV				-0.095*				-0.109**
				(-1.77)				(-2.04)
ILLIQV				-0.037				-0.023
				(-0.65)				(-0.39)
PIN				-3.870***				-3.886***
				(-2.98)				(-3.15)
Ret_Std				-51.36***				-56.39***
				(-3.13)				(-3.44)
SRet_Std				15.715				19.377
				(1.26)				(1.54)
Adj. R-sq	0.0023	0.0009	0.0079	0.0552	0.0023	0.0010	0.0079	0.0553
N	2944	2944	2944	1560	2944	2944	2944	1560

Table 6: Fama-MacBeth Regressions for Robustness Checks Using *VOIB_SHR*

This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012. Model 1 (Model 2) uses raw return (mid quote return from open to close) as the dependent variable. Except for Models 1 and 2, the dependent variable is the risk-adjusted return calculated using the Fama-French (1993) factors as well as the momentum factor and the liquidity factor of Pastor and Stambaugh (2003) with loadings conditional on the size and book-to market ratio. In Model 3, the order imbalance calculation is based on the dollar volume. In Models 4 (Model 5), all shock variables are calculated using the three-month (twelve-month) moving averages accordingly. Model 6 excludes the great financial crisis period of 2008 and 2009. Models 7 (8) uses data before (after) January 2001 only. Model 9 uses the Weighted Least Squares regressions in cross-sectional estimation following Asparouhova, Bessembinder and Kalcheva (ABK, 2010). In Model 10, we form decile portfolios sorted by *OIB* every day, replace each individual firm's *OIB* with the average *OIB* of the decile portfolio to which the firm belongs ($\widehat{OIB}$), and construct other order flow variables using $\widehat{OIB}$. All independent variables (except *R1* and *R212*) are lagged one month. *OIB* is the monthly order imbalance defined as *(B-S)/(B+S)*, where *B (S)* is the number of shares traded initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *POIB* is the logistic transform of the ratio of number of days with positive order imbalance and total number of trading days in the month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SIZE* represents the logarithm of market capitalization. *BM* is the logarithm of the book-to-market ratio. *R1* is the lagged one month return. *R212* is the cumulative returns over the second through the twelfth months prior to the current month. *TURN* is the logarithm of the monthly average turnover ratio calculated as the trading volume divided by shares outstanding. *StdTURN* is the standard deviation of *TURN* in the past 36 months. *ILLIQ* represents the Amihud measure of illiquidity. *ACC* represents accruals, measured as in Sloan (1996). *AG* is the asset growth computed in Cooper, Gulen and Shill (2008). *ISSUE* represents new issues as in Pontiff and Woodgate (2008). *IVOL* is the idiosyncratic volatility computed as in Ang, Hodrick, Xing, and Zhang (2006). *PROFIT* is the profitability variable as in Fama and French (2006). *SUE* is the standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. *MAX* is the maximum daily return in the last month. DISP is the analyst dispersion in earnings forecasts and *DISPD* is a dummy that equals to one if the stock is covered by at least two analysts and zero otherwise. *SSTT* is the small size trade imbalance as in Hvidkjaer (2008). *HiloSprd* is the high-low spread estimate of Corwin and Schultz (2012). *INSTV* is the average of the eight most recent quarterly absolute institutional ownership percentage changes. *ILLIQV* is the idiosyncratic volatility of liquidity in Akbas, Armstrong and Petkova (2011). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). *Ret_Std* is the standard deviation of daily returns in a month. *SOIB*, *SPOIB*, *STURN*, *SStdTURN*, *SIVOL*, *SILLIQ*, *SDISP*, and *SRet_Std* are defined similarly as *SVOIB*. All variables are winsorized at the 0.5% and 99.5% levels. *N* is the average number of stocks per month. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Model	1	2	3	4	5	6	7	8	9	10
	raw ret	o-c ret	*OIB$*	MA=3	MA=12	ex-crisis	post2001	pre2001	ABK	$\widehat{OIB}$
Intercept	3.905***	4.380***	2.507***	2.969***	2.680***	2.563***	3.952***	1.563*	2.540***	2.739***
	(4.33)	(3.42)	(3.98)	(4.96)	(3.98)	(3.89)	(4.89)	(1.85)	(4.06)	(4.36)
VOIB	3.809***	5.426***	3.481***	2.957***	2.984***	3.113***	7.306***	0.969**	3.538***	2.115***
	(4.06)	(3.97)	(3.85)	(3.74)	(3.53)	(3.45)	(3.80)	(2.20)	(3.93)	(3.96)
SVOIB	-5.158***	-6.326***	-4.642***	-4.225***	-3.439***	-4.118***	-8.850***	-1.918***	-4.734***	-3.130***
	(-6.19)	(-5.23)	(-5.70)	(-5.98)	(-4.60)	(-5.32)	(-5.51)	(-4.25)	(-5.83)	(-7.81)
ILLIQ	2.632***	3.621***	2.323***	2.299***	1.968***	2.250***	2.564***	2.048***	2.256***	2.288***
	(5.63)	(6.13)	(5.14)	(5.24)	(3.83)	(4.66)	(4.20)	(3.19)	(4.97)	(4.96)
SILLIQ	-1.256***	-2.487***	-1.044**	-1.255***	-0.757	-0.854*	-2.510***	0.072	-0.979**	-1.067**
	(-2.99)	(-5.97)	(-2.50)	(-3.34)	(-1.50)	(-1.91)	(-4.03)	(0.14)	(-2.30)	(-2.46)
OIB	-0.442	-0.684	-0.122	-0.258	-1.093**	-0.701**	-1.091	-0.006	-0.450	-0.152
	(-1.20)	(-1.18)	(-0.30)	(-1.03)	(-2.29)	(-2.09)	(-1.44)	(-0.02)	(-1.22)	(-1.09)
SOIB	0.624*	0.791	0.271	0.464**	1.268***	0.865***	1.064	0.267	0.592	0.904***
	(1.73)	(1.35)	(0.70)	(1.98)	(2.76)	(2.70)	(1.46)	(0.78)	(1.64)	(3.26)
POIB	0.076	0.270**	0.001	0.066	0.156	0.125*	0.077	0.074	0.075	0.032
	(1.11)	(2.30)	(0.01)	(1.26)	(1.61)	(1.95)	(0.63)	(0.98)	(1.13)	(0.83)

Table 6 (continued)

	raw ret	o-c ret	OIB$	MA=3	MA=12	ex-crisis	post2001	pre2001	ABK	$\widehat{\text{OIB}}$
SPOIB	-0.053	-0.128	0.031	-0.033	-0.137	-0.099	-0.058	-0.036	-0.044	-0.329
	(-0.75)	(-1.13)	(0.40)	(-0.61)	(-1.35)	(-1.51)	(-0.45)	(-0.45)	(-0.63)	(-1.21)
SIZE	-0.205***	-0.238***	-0.179***	-0.214***	-0.195***	-0.186***	-0.240***	-0.144***	-0.184***	-0.192***
	(-3.99)	(-3.44)	(-4.99)	(-6.17)	(-5.18)	(-5.04)	(-4.64)	(-3.13)	(-5.16)	(-5.47)
BM	0.120*	0.074	-0.011	-0.016	0.005	-0.018	-0.011	-0.013	-0.013	-0.008
	(1.78)	(0.84)	(-0.24)	(-0.35)	(0.10)	(-0.37)	(-0.17)	(-0.20)	(-0.27)	(-0.17)
R212	0.131	0.107	-0.033	0.070	-0.198	0.211*	-0.654*	0.394***	-0.033	-0.027
	(0.62)	(0.38)	(-0.17)	(0.37)	(-0.96)	(1.76)	(-1.67)	(2.59)	(-0.17)	(-0.14)
R1	-0.042***	-0.021***	-0.049***	-0.048***	-0.048***	-0.048***	-0.040***	-0.057***	-0.049***	-0.048***
	(-9.24)	(-4.01)	(-10.52)	(-10.32)	(-10.07)	(-9.97)	(-5.91)	(-9.11)	(-10.54)	(-10.34)
TURN	0.092	0.106	0.139*	0.245***	0.052	0.142*	-0.046	0.270**	0.142*	0.102
	(1.08)	(0.93)	(1.65)	(2.90)	(0.58)	(1.66)	(-0.49)	(2.24)	(1.70)	(1.19)
STURN	0.929***	0.829***	0.763***	0.673***	0.763***	0.750***	0.815***	0.741***	0.771***	0.813***
	(8.68)	(6.57)	(8.52)	(8.20)	(7.22)	(8.11)	(7.08)	(5.77)	(8.65)	(8.98)
STDTURN	-0.152**	-0.150	-0.172***	-0.255***	-0.159***	-0.186***	-0.124*	-0.211***	-0.176***	-0.177***
	(-2.35)	(-1.55)	(-3.74)	(-5.60)	(-3.17)	(-4.01)	(-1.65)	(-3.72)	(-3.83)	(-3.84)
SSTDTURN	-0.381**	-0.448**	-0.181	0.081	-0.229**	-0.134	-0.230	-0.148	-0.181	-0.161
	(-2.38)	(-2.16)	(-1.31)	(0.44)	(-2.02)	(-0.97)	(-1.22)	(-0.75)	(-1.29)	(-1.16)
IVOL	13.299	32.432	25.424	13.607	-3.680	21.671	49.429*	7.277	24.592	26.061
	(0.62)	(1.24)	(1.35)	(0.90)	(-0.16)	(1.09)	(1.81)	(0.29)	(1.31)	(1.38)
SIVOL	-16.024	-35.698*	-23.623	-10.703	2.626	-20.903	-49.320**	-4.744	-22.991	-24.437
	(-0.96)	(-1.83)	(-1.59)	(-1.02)	(0.13)	(-1.33)	(-2.13)	(-0.25)	(-1.54)	(-1.63)
ACC	-1.172***	-0.949**	-1.131***	-1.178***	-1.189***	-1.125***	-0.756*	-1.379***	-1.125***	-1.115***
	(-3.22)	(-2.01)	(-3.41)	(-3.45)	(-3.63)	(-3.24)	(-1.85)	(-2.90)	(-3.40)	(-3.38)
AG	-0.214***	-0.204***	-0.212***	-0.211***	-0.187***	-0.213***	-0.118*	-0.277***	-0.212***	-0.217***
	(-3.52)	(-2.80)	(-3.20)	(-3.16)	(-3.14)	(-3.02)	(-1.69)	(-2.76)	(-3.18)	(-3.28)
ISSUE	-1.227***	-1.725***	-1.196***	-1.198***	-1.212***	-1.263***	-1.35***	-1.100**	-1.203***	-1.225***
	(-3.98)	(-3.70)	(-4.04)	(-3.99)	(-3.89)	(-4.11)	(-3.67)	(-2.56)	(-4.06)	(-4.14)
PROFIT	0.018	0.093	0.062	0.068	0.112	0.094	-0.107	0.188**	0.069	0.069
	(0.28)	(1.06)	(0.97)	(1.01)	(1.55)	(1.42)	(-1.25)	(2.17)	(1.06)	(1.08)
SUE	0.030**	0.021	0.030**	0.032***	0.026**	0.038***	0.032*	0.029*	0.030**	0.031**
	(2.20)	(1.13)	(2.42)	(2.59)	(2.06)	(3.24)	(1.68)	(1.79)	(2.47)	(2.49)
Max	5.343***	2.920*	4.733***	4.815***	5.477***	5.035***	3.903**	5.373***	4.772***	4.843***
	(4.51)	(1.90)	(4.13)	(4.30)	(4.49)	(4.38)	(2.39)	(3.45)	(4.18)	(4.21)
DISP	-0.334***	-0.143	-0.361***	-0.351***	-0.385***	-0.366***	-0.350***	-0.352**	-0.352***	-0.356***
	(-2.80)	(-0.80)	(-3.46)	(-3.75)	(-3.17)	(-3.31)	(-2.90)	(-2.26)	(-3.38)	(-3.43)
SDISP	0.025	-0.030	0.087	0.125	0.077	0.084	0.303*	-0.062	0.086	0.091
	(0.20)	(-0.12)	(0.75)	(1.10)	(0.60)	(0.68)	(1.89)	(-0.39)	(0.74)	(0.78)
DISPD	-0.224**	-0.157	-0.111	-0.060	-0.105	-0.074	-0.133	-0.074	-0.096	-0.063
	(-2.44)	(-1.20)	(-1.52)	(-0.82)	(-1.38)	(-1.00)	(-1.11)	(-0.80)	(-1.31)	(-0.88)
SSTT	6.985	-24.900*	9.439	7.435	7.470	10.873	-14.22**	25.467	9.431	8.949
	(0.63)	(-1.77)	(0.83)	(0.65)	(0.65)	(0.89)	(-2.12)	(1.39)	(0.82)	(0.78)
HiLoSprd	28.737***	35.596***	25.509***	27.933***	29.254***	27.871***	13.483	33.724***	25.364***	24.999***
	(4.01)	(3.68)	(3.28)	(3.38)	(4.46)	(3.44)	(1.56)	(2.97)	(3.29)	(3.20)
INSTV	-0.083	-0.153	-0.100*	-0.117**	-0.042	-0.094	-0.170***	-0.041	-0.095*	-0.094*
	(-1.36)	(-0.96)	(-1.84)	(-2.14)	(-0.55)	(-1.64)	(-3.16)	(-0.50)	(-1.78)	(-1.72)
ILLIQV	-0.064	-0.011	-0.030	-0.025	-0.025	-0.033	0.108	-0.136**	-0.037	-0.014
	(-1.00)	(-0.10)	(-0.52)	(-0.43)	(-0.44)	(-0.55)	(1.12)	(-2.05)	(-0.65)	(-0.24)
PIN	-4.277***	-7.474***	-3.890***	-3.166***	-3.565***	-2.948**	-12.100***	1.777**	-3.864***	-2.930***
	(-2.98)	(-3.75)	(-3.00)	(-2.70)	(-2.91)	(-2.31)	(-5.45)	(2.28)	(-2.98)	(-2.79)
Ret_Std	-42.607**	-50.486*	-51.440***	-47.130***	-26.527	-52.940***	-58.370**	-46.560**	-51.380***	-52.170***
	(-2.12)	(-1.95)	(-3.12)	(-3.48)	(-1.33)	(-3.04)	(-2.51)	(-2.06)	(-3.13)	(-3.18)
SRet_Std	6.476	18.772	15.696	11.754	-10.454	17.039	32.303*	4.337	15.708	16.242
	(0.42)	(1.02)	(1.26)	(1.29)	(-0.62)	(1.29)	(1.72)	(0.26)	(1.26)	(1.29)
Adj. R-sq	0.080	0.084	0.055	0.054	0.056	0.055	0.048	0.060	0.055	0.054
N	1561	1497	1560	1576	1514	1528	2082	1198	1560	1560

Table 7: Bivariate Portfolio Sorts Controlling for Limits to Arbitrage Proxies

In Panel A (B), each month between July 1983 and December 2012, stocks are first sorted into high and low groups based on one of the lagged control variables, and then into lagged *SVOIB_SHR* (*SVOIB_NUM*) quintile portfolios within each control variable group. The quintile portfolio returns, the return differences between high and low quintile *SVOIB* portfolios, and the alphas with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor, are reported with the Newey-West *t*-statistics in parentheses. *VOIB* is the standard deviation of daily order imbalance in a month, where order imbalance is defined as *(B-S)/(B+S)* with *B (S)* being the trades initiated by buyers (sellers). *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. The order imbalance is calculated using the number of shares traded in Panel A and using the number of trades in Panel B. *SIZE* represents the logarithm of market capitalization. *INST* is the percentage of shares held by institutional investors. *IVOL* is the idiosyncratic stock return volatility. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: *SVOIB_SHR*

	Controlling for *SIZE*			Controlling for *INST*			Controlling for *IVOL*		
SVOIB	Small	Large	Small-Large	Low	High	Low-High	Low	High	Low-High
Low	2.239	1.192		2.067	1.406		1.423	2.168	
2	1.450	1.004		1.442	1.121		1.204	1.197	
3	1.354	1.006		1.311	1.116		1.149	1.085	
4	1.341	1.000		1.325	1.112		1.173	1.116	
High	1.057	0.959		1.127	1.103		1.076	1.011	
High-Low	-1.182***	-0.233***	-0.949***	-0.940***	-0.303***	-0.637***	-0.347***	-1.157***	0.810***
	(-7.34)	(-2.59)	(-6.38)	(-6.15)	(-2.93)	(-3.93)	(-5.02)	(-7.42)	(5.67)
Alpha	-1.107***	-0.308***	-0.799***	-0.877***	-0.297***	-0.580***	-0.302***	-1.064***	0.763***
	(-7.31)	(-2.94)	(-5.82)	(-5.58)	(-2.62)	(-3.90)	(-4.17)	(-7.66)	(5.95)

Panel B: *SVOIB_NUM*

	Controlling for *SIZE*			Controlling for *INST*			Controlling for *IVOL*		
SVOIB	Small	Large	Small-Large	Low	High	Low-High	Low	High	Low-High
Low	2.338	1.253		2.234	1.452		1.517	2.318	
2	1.423	1.038		1.399	1.134		1.219	1.111	
3	1.234	1.014		1.198	1.08		1.119	1.043	
4	1.453	1.011		1.289	1.142		1.121	1.204	
High	0.993	0.845		1.153	1.05		1.048	0.902	
High-Low	-1.346***	-0.408***	-0.938***	-1.081***	-0.402***	-0.679***	-0.469***	-1.416***	0.947***
	(-8.96)	(-4.14)	(-6.56)	(-6.93)	(-3.59)	(-4.06)	(-6.23)	(-9.46)	(8.02)
Alpha	-1.308***	-0.308***	-1.000***	-1.057***	-0.297***	-0.760***	-0.453***	-1.356***	0.903***
	(-9.77)	(-2.94)	(-6.82)	(-7.01)	(-2.62)	(-4.61)	(-5.64)	(-9.91)	(8.79)

Table 8: Dynamic Effects of Liquidity Shocks

This table examines the effects of *SVOIB* and *SILLIQ* on returns in the next 1, 2-3, 4-6, 7-9, and 10-12 months. We calculate *OIB* as the monthly order imbalance defined as *(B-S)/(B+S)*, where *B (S)* is the number of shares traded initiated by buyers (sellers). VOIB is the standard deviation of daily order imbalance in a month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SILLIQ* is the shock to the Amihud illiquidity measure defined similarly as *SVOIB*. Panel A reports the average alpha with respect to the Fama-French three-factor along with the momentum and liquidity factors of portfolios that buy stocks in the highest quintile of illiquidity shocks and sell stocks in the lowest quintile of illiquidity shocks, and the time-series averages of the coefficient estimates for the illiquidity shocks from cross-sectional regressions as in Table 5. In Panel B, we divide the sample into two groups based on the firm size and report for each group the time-series averages of the coefficient estimates of illiquidity shocks from cross-sectional regressions. All cross-sectional regressions control for the same variables as those in Column 4 of Table 5. For brevity, coefficient estimates are reported only for the illiquidity shock variables. All variables are winsorized at the 0.5% and 99.5% levels. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Price impact in the long run

	Alpha		FM coefficient	
	SVOIB_SHR	*SILLIQ*	*SVOIB_SHR*	*SILLIQ*
Month1	-0.816***	-1.534***	-4.739***	-0.979**
	(-6.45)	(-9.08)	(-5.84)	(-2.30)
Month2-3	0.091	-0.639***	0.171	-1.232***
	(0.92)	(-4.47)	(0.90)	(-3.45)
Month4-6	0.217**	-0.178	0.403**	-1.655***
	(2.07)	(-1.29)	(2.36)	(-4.84)
Month7-9	0.251***	-0.094	0.360*	-1.496*
	(2.85)	(-0.62)	(1.83)	(-1.86)
Month10-12	0.345***	0.148	0.413**	-0.816
	(3.83)	(0.96)	(2.31)	(-0.94)

Panel B: Fama-MacBeth regression coefficient estimates conditioning on SIZE

	Small		Large	
	SVOIB_SHR	*SILLIQ*	*SVOIB_SHR*	*SILLIQ*
Month1	-4.441***	-1.383***	-2.002**	-51.577
	(-5.02)	(-4.33)	(-2.26)	(-1.13)
Month2-3	-0.671***	-0.643***	0.204	-27.094
	(-2.62)	(-3.36)	(0.52)	(-0.59)
Month4-6	0.731***	-0.577***	0.484	-12.268
	(2.99)	(-2.62)	(1.46)	(-0.39)
Month7-9	1.117***	-0.170	0.337	-24.137
	(4.71)	(-0.83)	(1.23)	(-0.86)
Month10-12	1.267***	-0.080	0.483	-11.864
	(5.23)	(-0.46)	(1.60)	(-0.35)

Table 9: *VOIB* around Corporate Events

Panel A presents the average standard deviation of daily order imbalance (*VOIB*) in the earnings and non-earnings announcement periods, as well as the difference of *VOIB* between the earnings and non-earnings periods. We also report the average *VOIB* differences between earnings and non-earnings periods for the low (high) absolute earnings announcements surprise as well as the difference-in-difference results between the high and low surprise groups. The earnings period is defined as trading days -18 to 2 relative to the announcement day, and the non-earnings period is defined as all the other days between days -31 and 31 relative to the announcement day. Earnings surprise is calculated as the difference between the actual value and the median forecast scaled by the market price at the end of the month preceding the earnings announcement. Panel B replicates the analysis using *VOIB* calculated for the event and non-event periods before M&A announcements. The M&A (non-M&A) period is defined as days -30 to -1 (-60 to -31) relative to the announcement day. We also report the average *VOIB* difference for deals with the takeover premium below and above the median separately as well as the difference-in-difference between the two groups. All standard errors are clustered on both time and firm identities and the associated *t*-statistics are reported in parentheses. Panel C reports the Fama-MacBeth regression coefficient estimates similar to Table 5**.** The dependent variable is the risk-adjusted return calculated using the Fama-French (1993) factors as well as the momentum factor and the liquidity factor of Pastor and Stambaugh (2003) with loadings conditional on the size and book-to market ratio. Event dummy is one if there is an earnings or M&A announcement in the following month, and zero otherwise. The regressions have the full set of control variables as in Column 4 of Table 5. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: *VOIB* around earning and non-earnings periods

	Earnings and non-earnings periods			Earnings and non-earnings periods *VOIB* differences for low and high absolute earnings surprise		
	Earnings periods	Non-earnings periods	Earnings minus Non-earnings	Low absolute earnings surprise	High absolute earnings surprise	High-Low
VOIB_SHR	0.385	0.306	0.079***	0.075***	0.083***	0.008***
			(45.70)	(40.87)	(44.19)	(2.98)
VOIB_NUM	0.319	0.256	0.063***	0.058***	0.069***	0.010***
			(42.85)	(39.33)	(41.54)	(4.60)

Panel B: *VOIB* around M&A and non-M&A periods

	M&A and non-M&A periods			*VOIB* differences for low and high premium		
	M&A periods	Non-M&A periods	M&A minus Non-M&A	Low premium	High premium	High-Low
VOIB_SHR	0.492	0.480	0.012***	0.007**	0.016***	0.009**
			(6.13)	(2.36)	(5.59)	(2.19)
VOIB_NUM	0.432	0.423	0.009***	0.005*	0.012***	0.007*
			(4.93)	(1.75)	(4.41)	(1.79)

Panel C: Fama-MacBeth regression coefficient estimates

	VOIB_SHR	*VOIB_NUM*
Event dummy	0.999	-0.632
	(0.70)	(-0.40)
VOIB	2.842***	3.218***
	(3.08)	(3.64)
*VOIB**Event dummy	3.600***	5.029***
	(2.66)	(3.81)
SVOIB	-4.094***	-4.904***
	(-4.81)	(-5.94)
*SVOIB**Event dummy	-1.833*	-1.956**
	(-1.92)	(-2.01)
Control	Yes	Yes

Table 10: Limit Order Book Imbalance Volatility Results

This table examines the effect of limit order imbalance volatility using LOBSTER data from July 2007 to December 2012. For each stock, we take snapshots of the limit order book every five minutes during trading hours. The top three steps of limit buy and sell orders are aggregated to the total buy and sell volumes for each snapshot. We then calculate the average buy and sell volumes of all snapshots taken on a day and the imbalance between them as the daily limit order imbalance. Panel A presents the average standard deviation of daily limit order book imbalance (*VOIB_LOB*) around earnings and M&A periods. The construction of those event and non-event periods is the same as in Table 9. Panel A also reports *VOIB* results in the same sample period. All standard errors are clustered on both time and firm identities and the associated *t*-statistics are reported in parentheses. In Panel B, we sort all stocks in the restricted sample into quintile portfolios based on *VOIB_SHR*, *SVOIB_SHR*, *VOIB_NUM*, *SVOIB_NUM*, *VOIB_LOB*, *SVOIB_LOB*,. The equally-weighted portfolio returns for month t+1 are reported. Also reported are the return differences between the high and low quintiles and the alphas with respect to the Fama-French (1993) factors along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor. Panel C reports Fama-MacBeth regression coefficient estimates similar to Table 5. Models 4 and 5 include the same control variables in Column 4 of Table 5. For brevity, coefficient estimates are reported only for *VOIB_LOB*, *SVOIB_LOB*, *VOIB* and *SVOIB*. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Corporate events

	Earnings and non-earnings periods			M&A and non-M&A periods		
	Earnings periods	Non-earnings periods	Earnings minus Non-earnings	M&A periods	Non-M&A periods	M&A minus Non-M&A
VOIB_SHR	0.246	0.181	0.064***	0.374	0.373	0.000
			(50.34)			(0.10)
VOIB_NUM	0.222	0.165	0.057***	0.337	0.334	0.003
			(54.57)			(0.66)
VOIB_LOB	0.205	0.202	0.002	0.312	0.311	0.001
			(1.47)			(0.39)

Panel B: Univarite sorts

	VOIB_SHR	*SVOIB_SHR*	*VOIB_NUM*	*SVOIB_NUM*	*VOIB_LOB*	*SVOIB_LOB*
Quintile						
Low-1	0.320	1.419	0.346	1.464	0.675	0.903
2	0.477	0.764	0.483	0.800	0.626	0.924
3	0.684	0.87	0.625	0.768	0.374	0.831
4	0.652	0.833	0.599	0.875	0.594	0.767
High-5	1.041	0.613	1.120	0.592	0.905	1.077
High-Low	0.721**	-0.807***	0.775***	-0.872***	0.230	0.174
	(2.43)	(-3.06)	(2.68)	(-3.35)	(0.51)	(0.59)
Alpha	0.626**	-0.742***	0.661**	-0.820***	0.269	0.206
	(2.15)	(-3.52)	(2.29)	(-3.38)	(0.61)	(0.72)

Panel C: Fama-MacBeth regression coefficient estimates

	Model 1	Model 2	Model 3	Model 4	Model 5
VOIB_LOB	-0.160		0.014	1.994*	1.660
	(-0.11)		(0.01)	(1.66)	(1.51)
SVOIB_LOB		0.593	1.030	-0.965	-1.518
		(0.36)	(0.43)	(-0.74)	(-1.14)
VOIB_SHR				15.313***	
				(3.99)	
SVOIB_SHR				-15.942***	
				(-5.68)	
VOIB_NUM					14.964***
					(4.27)
SVOIB_NUM					-16.253***
					(-6.35)
Control	No	No	No	Yes	Yes

Figure 1: Order Imbalance Volatility Over Time

This figure plots the equally-weighted (EW) and value-weighted (VW) order imbalance volatility calculated using the number of shares shared (*VOIB_SHR*) and the number of transactions (*VOIB_NUM*). The time trend line is also plotted. The order imbalance is defined *as (B-S)/(B+S),* where *B (S)* is the trades initiated by buyers (sellers). Order imbalance volatility is measured as the monthly standard deviation of daily order imbalance, and the first observation is in January, 1983.

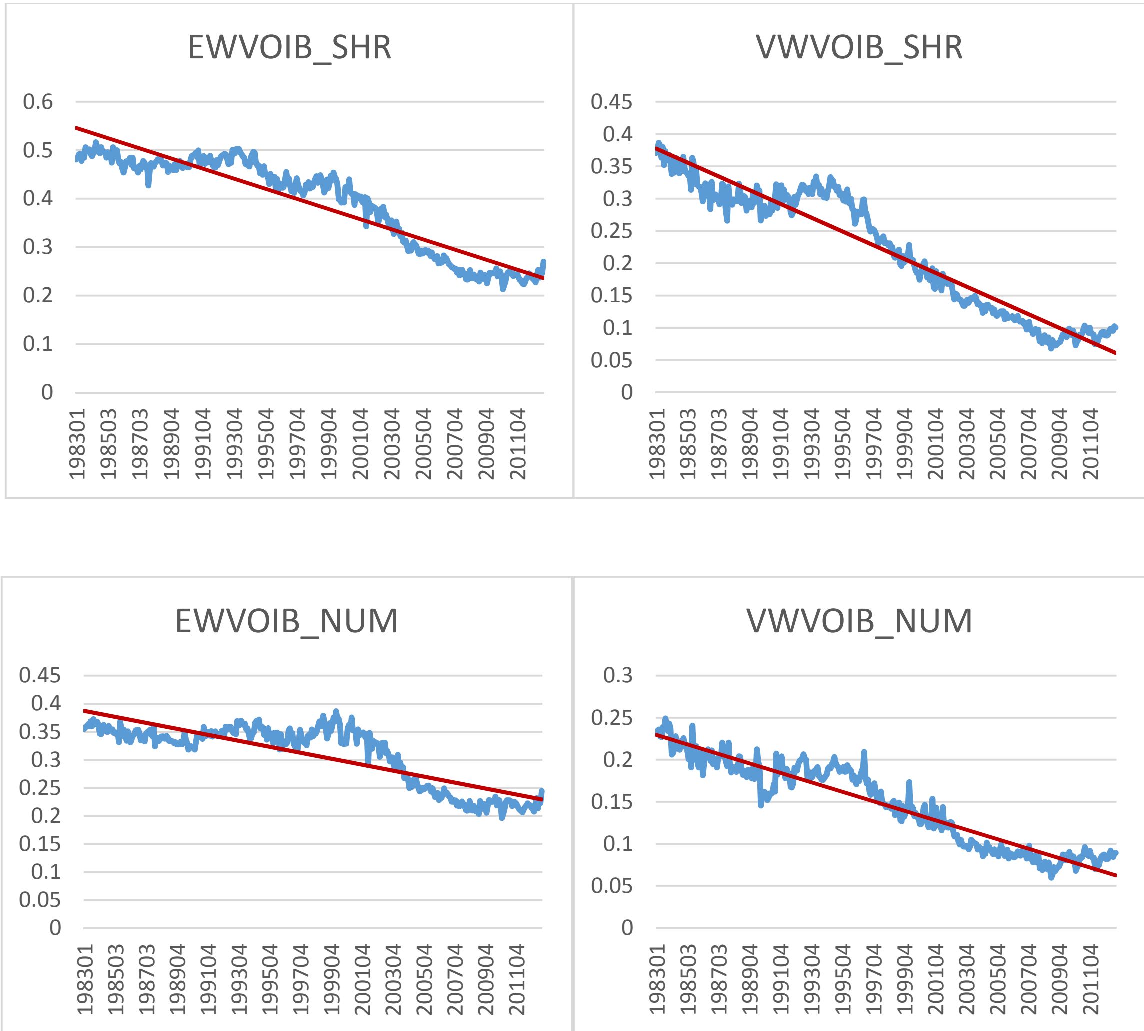
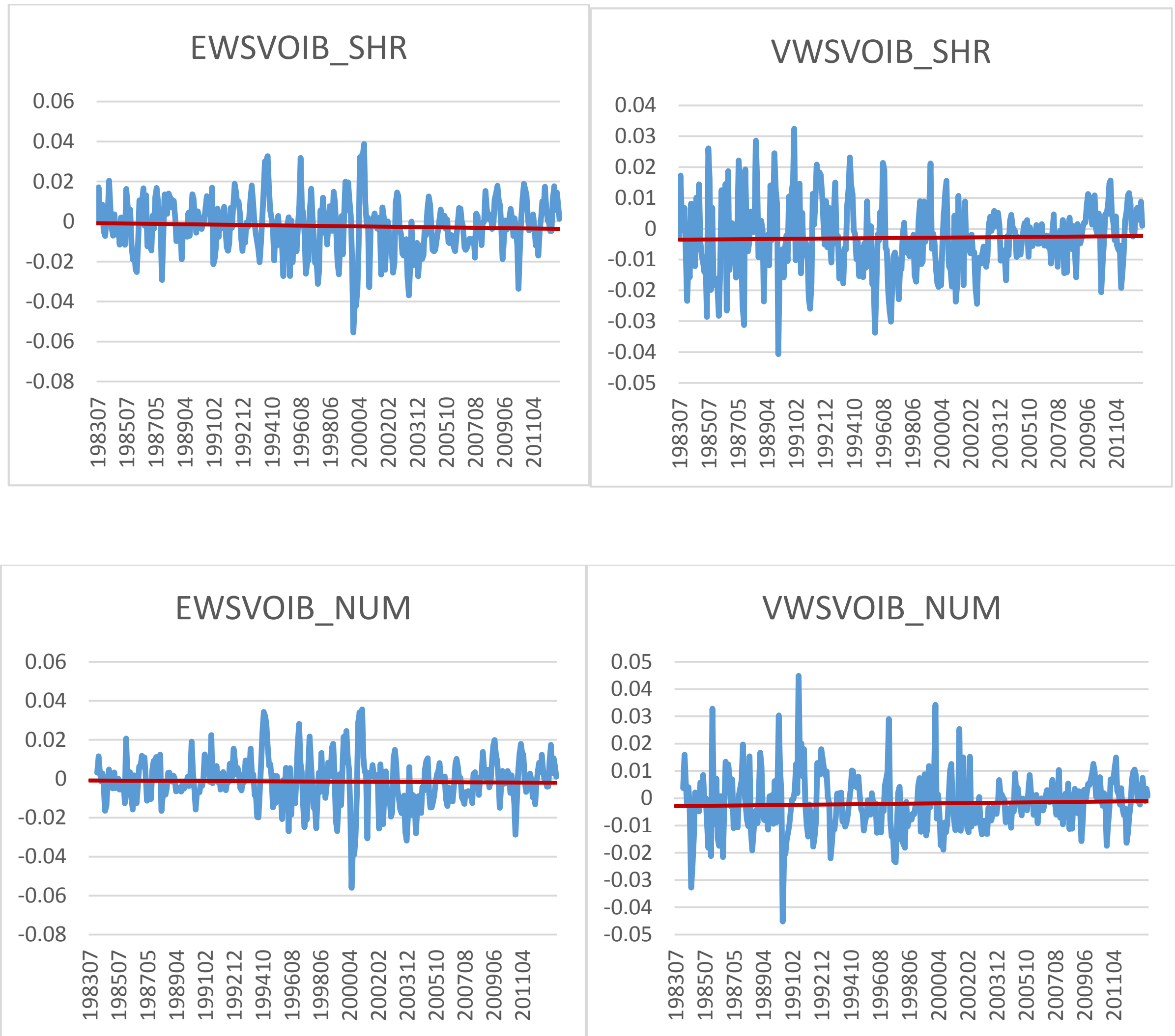


Figure 2: Shocks to Order Imbalance Volatility Over Time

This figure plots the equally-weighted (EW) and value-weighted (VW) averages of shocks to order imbalance volatility calculated using the number of shares traded (*SVOIB_SHR*) and the number of transactions (*SVOIB_NUM*). The order imbalance is defined as *(B-S)/(B+S)*, where *B (S)* is the number of trades initiated by buyers (sellers). Shocks are calculated as the current month's order imbalance volatility less the lagged value of the six month moving average of monthly order imbalance volatility. The time trend line is also plotted. The first observation is in July, 1983.



Internet Appendix

Table A1: Determinants of *VOIB*

Each month, we regress *VOIB* on firm size and proxies for *N*, v_δ, and v_z. *OIB_SHR* is the monthly order imbalance defined as $(B-S)/(B+S)$, where *B* (*S*) is the number of shares traded initiated by buyers (sellers). *VOIB_SHR* is the standard deviation of daily order imbalance in a month. The order imbalance volatility calculated using number of trades is termed *VOIB_NUM*. The number of informed agents, *N*, is measured as the number of informed institutional investors. Following Abarbanell, Bushee and Raedy (2003), we break down institutional investors into informed and uninformed types, where the informed institutions are defined as investment companies and independent investment advisors because such institutions are more likely to be active investors. Other institutions, such as bank trusts, insurance companies, corporate/private pension funds, public pension funds, university and foundation endowments, have longer investment horizons and trade less actively. Following Chordia, Huh and Subrahmanyam (2009), we employ earnings volatility as a proxy for v_δ, where earnings volatility is the standard deviation of earnings per share (EPS) from the most recent eight quarters. Finally, we employ the average of daily dollar volume (in million dollars) as a proxy for v_z. All independent variables are standardized in the cross section. The time-series averages of coefficient estimates from monthly cross-sectional regressions are presented along with the associated Newey-West (1987) *t*-statistics. The coefficients are multiplied by 1000.

	VOIB_SHR	*VOIB_NUM*
Intercept	126.408***	121.566***
	(71.76)	(92.67)
N	0.459**	0.526***
	(2.46)	(3.28)
v_δ	0.178***	0.216***
	(6.17)	(8.05)
v_z	0.795***	1.133***
	(6.05)	(16.55)
SIZE	-7.123***	-7.203***
	(-36.58)	(-68.78)

Table A2: Summary Statistics before and after January 2001

Panel A (Panel B) presents the time-series averages of the cross-sectional statistics for common stocks listed on NYSE, AMEX and NASDAQ before (after) January 2001. The stock-month observation must have valid information to calculate the return, market capitalization, book-to-market ratio, and order imbalance, and must have the month-end price above one dollar. *OIB_SHR* is the monthly order imbalance defined *as (B−S)/(B+S),* where *B (S)* is the number of shares traded initiated by buyers (sellers). *VOIB_SHR* is the standard deviation of daily order imbalance in a month. *SVOIB_SHR* is the difference between *VOIB_SHR* in the current month and the six-month moving average of *VOIB_SHR* in the previous month. The variables calculated using the number of trades are termed as *OIB_NUM*, *VOIB_NUM*, and *SVOIB_NUM*. The Amihud illiquidity (*ILLIQ*) is calculated as the monthly average of the daily ratio of the absolute return to the dollar volume. *TURN* is the logarithm of the monthly average of the daily turnover ratio calculated as the number of shares traded divided by shares outstanding. *SPRD* is the spread measure using the cheap alternative solution by Holden and Jacobsen (2014). The shocks to the Amihud illiquidity (*SILLIQ*), turnover (*STURN*), spread (*SSPRD*), and return standard deviation (*SRet_Std*) are computed similarly to *SVOIB*.

Panel A: Descriptive statistics pre-2001						
Statistics	N	Mean	St. dev.	Median	Minimum	Maximum
OIB_SHR	2,444	-0.079	0.442	-0.119	-0.864	0.760
VOIB_SHR	2,444	0.414	0.180	0.419	0.064	0.831
SVOIB_SHR	2,444	0.001	0.125	0.001	-0.483	0.552
OIB_NUM	2,444	-0.071	0.384	-0.095	-0.790	0.695
VOIB_NUM	2,444	0.312	0.151	0.298	0.048	0.743
SVOIB_NUM	2,444	0.001	0.099	-0.003	-0.389	0.513
ILLIQ	2,444	0.070	0.335	0.003	0.000	7.823
SILLIQ	2,444	-0.006	0.375	0.000	-8.391	5.999
TURN	2,436	0.786	0.673	0.581	0.164	5.333
STURN	2,411	-0.025	0.481	-0.046	-2.297	2.466
SPRD	2,180	1.979	1.878	1.360	0.224	15.210
SSPRD	2,160	-0.001	0.007	0.000	-0.102	0.034

Panel B: Descriptive statistics post-2001						
Statistics	N	Mean	St. dev.	Median	Minimum	Maximum
OIB_SHR	3,687	-0.011	0.153	0.007	-0.662	0.469
VOIB_SHR	3,687	0.284	0.193	0.217	0.057	0.875
SVOIB_SHR	3,687	-0.004	0.069	-0.004	-0.434	0.327
OIB_NUM	3,687	0.003	0.125	0.014	-0.515	0.399
VOIB_NUM	3,687	0.251	0.182	0.187	0.045	0.847
SVOIB_NUM	3,687	-0.003	0.065	-0.003	-0.417	0.335
ILLIQ	3,687	0.072	0.445	0.001	0.000	9.135
SILLIQ	3,687	-0.004	0.453	0.000	-5.668	8.283
TURN	3,675	1.171	1.348	0.764	0.022	9.458
STURN	3,645	0.006	0.512	-0.026	-2.224	3.499
SPRD	3,667	1.178	1.765	0.446	0.038	12.483
SSPRD	3,666	0.000	0.000	0.000	-0.011	0.001

Table A3: Fama-MacBeth Regression Estimates Using an ARMA(1,1) Model for *VOIB* to Extract Shocks

This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012. The order imbalance is calculated using the number of shares traded in Column 1 and using the number of trades in Column 2. We use an ARMA(1,1) model for *VOIB* to extract shocks. The dependent variable is the risk-adjusted return calculated using the Fama-French (1993) factors as well as the momentum factor and the liquidity factor of Pastor and Stambaugh (2003) with loadings conditional on the size and book-to market ratio. All independent variables (except *R1* and *R212*) are lagged one month. *OIB* is the monthly order imbalance defined as *(B-S)/(B+S),* where *B (S)* is the trades initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *POIB* is the logistic transform of the ratio of number of days with positive order imbalance and total number of trading days in the month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SIZE* represents the logarithm of market capitalization. BM is the logarithm of the book-to-market ratio. *R1* is the lagged one month return. *R212* is the cumulative returns over the second through the twelfth months prior to the current month. *TURN* is the logarithm of the monthly average turnover ratio calculated as the trading volume divided by shares outstanding. *StdTURN* is the standard deviation of *TURN* in the past 36 months. *ILLIQ* represents the Amihud measure of illiquidity. *ACC* represents accruals, measured as in Sloan (1996). *AG* is the asset growth computed in Cooper, Gulen and Shill (2008). *ISSUE* represents new issues as in Pontiff and Woodgate (2008). *IVOL* is the idiosyncratic volatility computed as in Ang, Hodrick, Xing, and Zhang (2006). *PROFIT* is the profitability variable as in Fama and French (2006). *SUE* is the standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. *MAX* is the maximum daily return in the last month. *DISP* is the analyst dispersion in earnings forecasts and *DISPD* is a dummy that equals to one if the stock is covered by at least two analysts and zero otherwise. *SSTT* is the small size trade imbalance as in Hvidkjaer (2008). *HiloSprd* is the high-low spread estimate of Corwin and Schultz (2012). *INSTV* is the average of the eight most recent quarterly absolute institutional ownership percentage changes. *ILLIQV* is the idiosyncratic volatility of liquidity in Akbas, Armstrong and Petkova (2011). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). Ret_Std is the standard deviation of daily returns in a month. *SOIB, SPOIB, STURN, SStdTURN, SIVOL, SILLIQ, SDISP*, and *SRet_Std* are defined similarly as *SVOIB*. All variables are winsorized at the 0.5% and 99.5% levels. *N* is the average number of stocks per month. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	SHR	*NUM*
Model	1	2
Intercept	3.354***	3.178***
	(5.28)	(5.10)
VOIB	1.601***	1.783***
	(2.86)	(3.15)
SVOIB	-2.155***	-2.316***
	(-5.42)	(-5.55)
ILLIQ	2.292***	2.271***
	(4.95)	(4.96)
SILLIQ	-1.052**	-1.060**
	(-2.37)	(-2.41)
OIB	-0.527	-0.215
	(-1.42)	(-0.45)
SOIB	0.694*	0.824*
	(1.94)	(1.82)
POIB	0.102	-0.077
	(1.55)	(-1.19)
SPOIB	-0.064	0.079
	(-0.92)	(1.28)
SIZE	-0.225***	-0.208***
	(-6.32)	(-5.89)
BM	-0.010	-0.008
	(-0.21)	(-0.17)

Table A3 (continued)

R212	0.012	0.025
	(0.06)	(0.13)
R1	-0.048***	-0.048***
	(-10.37)	(-10.36)
TURN	0.068	0.114
	(0.80)	(1.34)
STURN	0.909***	0.851***
	(10.29)	(9.91)
STDTURN	-0.180***	-0.192***
	(-3.90)	(-4.15)
SSTDTURN	-0.111	-0.094
	(-0.81)	(-0.69)
IVOL	28.281	30.114
	(1.50)	(1.60)
SIVOL	-26.758*	-27.691*
	(-1.79)	(-1.85)
ACC	-1.132***	-1.111***
	(-3.42)	(-3.37)
AG	-0.212***	-0.211***
	(-3.19)	(-3.25)
ISSUE	-1.179***	-1.198***
	(-3.98)	(-4.03)
PROFIT	0.067	0.055
	(1.05)	(0.87)
SUE	0.031**	0.030**
	(2.54)	(2.46)
Max	4.710***	4.725***
	(4.14)	(4.13)
DISP	-0.365***	-0.373***
	(-3.49)	(-3.58)
SDISP	0.102	0.100
	(0.87)	(0.87)
DISPD	-0.074	-0.094
	(-1.02)	(-1.29)
SSTT	9.781	1.055
	(0.86)	(0.08)
HiLoSprd	25.442***	25.850***
	(3.26)	(3.29)
INSTV	-0.088*	-0.099*
	(-1.68)	(-1.90)
ILLIQV	-0.026	-0.016
	(-0.45)	(-0.26)
PIN	-2.908***	-3.019***
	(-2.74)	(-2.84)
Ret_Std	-54.655***	-56.535***
	(-3.32)	(-3.46)
SRet_Std	18.378	19.435
	(1.46)	(1.53)
Adj. R-sq	0.055	0.055
N	1379	1379

Table A4: Fama-MacBeth Regression Estimates with Other Control Variables

This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012. In this table, we add two more control variables: *SHiloSprd* and *O/S*. *SHiloSprd* is *HiloSprd* shock, where *HiloSprd* is the high-low spread estimate of Corwin and Schultz (2012). *O/S* is the ratio of option trading volume and stock trading volume, measured as in Roll, Schwartz and Subrahmanyam (2010). The order imbalance is calculated using the number of shares traded in Columns 1 to 3 and using the number of trades in Columns 4 to 6. The dependent variable is the risk-adjusted return calculated using the Fama-French (1993) factors as well as the momentum factor and the liquidity factor of Pastor and Stambaugh (2003) with loadings conditional on the size and book-to market ratio. All independent variables (except *R1* and *R212*) are lagged one month. *OIB* is the monthly order imbalance defined as *(B-S)/(B+S),* where *B (S)* is the trades initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *POIB* is the logistic transform of the ratio of number of days with positive order imbalance and total number of trading days in the month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SIZE* represents the logarithm of market capitalization. *BM* is the logarithm of the book-to-market ratio. *R1* is the lagged one month return. *R212* is the cumulative returns over the second through the twelfth months prior to the current month. *TURN* is the logarithm of the monthly average turnover ratio calculated as the trading volume divided by shares outstanding. *StdTURN* is the standard deviation of *TURN* in the past 36 months. *ILLIQ* represents the Amihud measure of illiquidity. *ACC* represents accruals, measured as in Sloan (1996). *AG* is the asset growth computed in Cooper, Gulen and Shill (2008). ISSUE represents new issues as in Pontiff and Woodgate (2008). *IVOL* is the idiosyncratic volatility computed as in Ang, Hodrick, Xing, and Zhang (2006). *PROFIT* is the profitability variable as in Fama and French (2006). *SUE* is the standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. *MAX* is the maximum daily return in the last month. *DISP* is the analyst dispersion in earnings forecasts and *DISPD* is a dummy that equals to one if the stock is covered by at least two analysts and zero otherwise. *SSTT* is the small size trade imbalance as in Hvidkjaer (2008). *INSTV* is the average of the eight most recent quarterly absolute institutional ownership percentage changes. *ILLIQV* is the idiosyncratic volatility of liquidity in Akbas, Armstrong and Petkova (2011). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). Ret_Std is the standard deviation of daily returns in a month. *SOIB*, *SPOIB*, *STURN*, *SStdTURN*, *SIVOL, SILLIQ, SDISP, SRet_Std* and SHiloSprd are defined similarly as *SVOIB*. All variables are winsorized at the 0.5% and 99.5% levels. *N* is the average number of stocks per month. Newey-West t-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

	SHR			*NUM*		
Model	1	2	3	4	5	6
Intercept	2.332***	1.683	1.513	2.099***	0.658	0.540
	(3.84)	(1.56)	(1.40)	(3.55)	(0.56)	(0.46)
VOIB	3.781***	9.000***	9.299***	4.262***	11.200***	11.537***
	(4.18)	(3.65)	(3.79)	(5.14)	(4.04)	(4.17)
SVOIB	-4.916***	-6.780***	-7.065***	-5.801***	-9.138***	-9.441***
	(-6.05)	(-3.07)	(-3.21)	(-7.53)	(-3.69)	(-3.80)
ILLIQ	2.265***	71.651*	70.904*	2.236***	65.402*	65.302*
	(4.94)	(1.79)	(1.77)	(4.92)	(1.78)	(1.77)
SILLIQ	-0.988**	-70.687*	-69.403*	-0.987**	-60.830*	-59.513*
	(-2.33)	(-1.85)	(-1.88)	(-2.37)	(-1.69)	(-1.70)
OIB	-0.424	-1.893*	-1.963*	-0.468	-2.728	-2.794
	(-1.15)	(-1.73)	(-1.79)	(-0.97)	(-1.53)	(-1.56)
SOIB	0.566	2.142*	2.183*	1.044**	4.362**	4.441**
	(1.57)	(1.73)	(1.76)	(2.29)	(2.36)	(2.40)
POIB	0.067	0.001	0.013	-0.038	0.093	0.105
	(1.03)	(0.01)	(0.10)	(-0.59)	(0.81)	(0.92)
SPOIB	-0.038	-0.060	-0.068	0.035	-0.142	-0.155
	(-0.54)	(-0.47)	(-0.53)	(0.55)	(-1.19)	(-1.31)
SIZE	-0.172***	-0.058	-0.051	-0.152***	-0.008	-0.002
	(-4.95)	(-1.04)	(-0.91)	(-4.48)	(-0.12)	(-0.04)
BM	-0.016	-0.119**	-0.120**	-0.011	-0.129**	-0.129**
	(-0.33)	(-2.00)	(-2.04)	(-0.22)	(-2.17)	(-2.19)

Table A4 (continued)

R212	-0.029	-0.559*	-0.549	-0.027	-0.584*	-0.575*
	(-0.15)	(-1.65)	(-1.62)	(-0.14)	(-1.78)	(-1.76)
R1	-0.049***	-0.039***	-0.040***	-0.049***	-0.042***	-0.042***
	(-10.62)	(-5.81)	(-5.86)	(-10.59)	(-6.03)	(-6.07)
TURN	0.161*	0.059	0.085	0.206**	0.157	0.183
	(1.95)	(0.41)	(0.59)	(2.46)	(1.03)	(1.19)
STURN	0.750***	0.405***	0.387**	0.675***	0.284*	0.268*
	(8.60)	(2.69)	(2.54)	(8.22)	(1.92)	(1.78)
STDTURN	-0.176***	0.090	0.085	-0.187***	0.049	0.043
	(-3.83)	(0.87)	(0.82)	(-4.08)	(0.46)	(0.41)
SSTDTURN	-0.171	0.237	0.253	-0.145	0.345	0.361
	(-1.24)	(0.89)	(0.96)	(-1.06)	(1.30)	(1.38)
IVOL	22.517	-8.279	-6.122	27.368	-2.172	0.165
	(1.20)	(-0.24)	(-0.18)	(1.46)	(-0.06)	(0.00)
SIVOL	-20.612	5.009	4.121	-23.852	-1.045	-2.139
	(-1.40)	(0.18)	(0.15)	(-1.62)	(-0.04)	(-0.08)
ACC	-1.119***	-0.727	-0.688	-1.094***	-0.655	-0.622
	(-3.40)	(-1.46)	(-1.37)	(-3.32)	(-1.32)	(-1.24)
AG	-0.216***	-0.173***	-0.174***	-0.217***	-0.166***	-0.167***
	(-3.21)	(-2.77)	(-2.80)	(-3.30)	(-2.73)	(-2.75)
ISSUE	-1.210***	-0.993**	-0.999**	-1.240***	-1.094***	-1.100***
	(-4.07)	(-2.53)	(-2.53)	(-4.16)	(-2.80)	(-2.80)
PROFIT	0.068	-0.048	-0.049	0.050	-0.053	-0.054
	(1.05)	(-0.55)	(-0.56)	(0.79)	(-0.62)	(-0.63)
SUE	0.030**	-0.020	-0.020	0.030**	-0.025	-0.025
	(2.44)	(-0.96)	(-0.93)	(2.40)	(-1.23)	(-1.19)
Max	4.856***	6.045***	6.113***	4.973***	6.468***	6.544***
	(4.28)	(3.68)	(3.75)	(4.39)	(3.81)	(3.87)
DISP	-0.355***	-0.155	-0.148	-0.368***	-0.158	-0.152
	(-3.40)	(-0.93)	(-0.89)	(-3.54)	(-0.97)	(-0.93)
SDISP	0.089	0.228	0.233	0.094	0.218	0.224
	(0.76)	(1.19)	(1.23)	(0.81)	(1.16)	(1.21)
DISPD	-0.102	-0.151	-0.142	-0.125*	-0.146	-0.138
	(-1.39)	(-0.89)	(-0.83)	(-1.70)	(-0.85)	(-0.80)
SSTT	8.644	-11.359	-10.384	-0.775	-24.751**	-23.653**
	(0.76)	(-1.28)	(-1.19)	(-0.06)	(-2.45)	(-2.40)
HiLoSprd	31.431***	17.356*	29.528**	34.376***	17.122	28.864**
	(3.42)	(1.65)	(2.24)	(3.72)	(1.58)	(2.15)
INSTV	-0.093*	-0.184	-0.187	-0.107**	-0.193	-0.196
	(-1.74)	(-1.55)	(-1.57)	(-2.03)	(-1.53)	(-1.54)
ILLIQV	-0.040	0.044	0.040	-0.024	0.086	0.084
	(-0.69)	(0.25)	(0.23)	(-0.40)	(0.51)	(0.50)
PIN	-3.969***	-13.300***	-13.442***	-4.011***	-12.747***	-12.950***
	(-3.05)	(-4.36)	(-4.39)	(-3.25)	(-4.50)	(-4.56)
Ret_Std	-51.637***	-26.074	-32.129	-57.176***	-33.926	-40.365
	(-3.09)	(-0.86)	(-1.07)	(-3.41)	(-1.11)	(-1.33)
SRet_Std	15.399	-3.570	0.682	19.359	1.980	6.654
	(1.20)	(-0.15)	(0.03)	(1.51)	(0.08)	(0.27)
SHiLoSprd	-7.768		-17.980	-9.871		-18.021
	(-1.07)		(-1.50)	(-1.37)		(-1.50)
O/S		-0.015***	-0.015***		-0.015***	-0.015***
		(-2.94)	(-2.89)		(-2.93)	(-2.86)
Adj. R-sq	0.056	0.062	0.063	0.056	0.063	0.063
N	1560	1159	1159	1560	1159	1159

Table A5: Fama-MacBeth Regressions for Robustness Checks Using *VOIB_NUM*

This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012. Model 1 (Model 2) uses raw return (mid quote return from open to close) as the dependent variable. In Models 3 (Model 4), all shock variables are calculated using the three-month (twelve-month) moving averages accordingly. Model 5 excludes the great financial crisis period of 2008 and 2009. Models 6 and 7 use data before and after January 2001 only. Model 8 uses the Weighted Least Squares regressions in cross-sectional estimation following Asparouhova, Bessembinder and Kalcheva (ABK, 2010). In Model 9, we form decile portfolios sorted by OIB every day, replace each individual firm's OIB with the average OIB of the decile portfolio to which the firm belongs ($\widehat{\mathrm{OIB}}$), and construct other order flow variables using $\widehat{OIB}$. All independent variables (except *R*1 and *R212*) are lagged one month. *OIB* is the monthly order imbalance defined as *(B-S)/(B+S),* where *B (S)* is the number of trades initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *POIB* is the logistic transform of the ratio of number of days with positive order imbalance and total number of trading days in the month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SIZE* represents the logarithm of market capitalization. *BM* is the logarithm of the book-to-market ratio. *R1* is the lagged one month return. *R212* is the cumulative returns over the second through the twelfth months prior to the current month. *TURN* is the logarithm of the monthly average turnover ratio calculated as the trading volume divided by shares outstanding. *StdTURN* is the standard deviation of *TURN* in the past 36 months. *ILLIQ* represents the Amihud measure of illiquidity. *ACC* represents accruals, measured as in Sloan (1996). *AG* is the asset growth computed in Cooper, Gulen and Shill (2008). *ISSUE* represents new issues as in Pontiff and Woodgate (2008). *IVOL* is the idiosyncratic volatility computed as in Ang, Hodrick, Xing, and Zhang (2006). *PROFIT* is the profitability variable as in Fama and French (2006). SUE is the standardized unexpected earnings, computed as the most recent quarterly earnings less the earnings four quarters ago, standardized by its standard deviation estimated over the prior eight quarters. *MAX* is the maximum daily return in the last month. *DISP* is the analyst dispersion in earnings forecasts and *DISPD* is a dummy that equals to one if the stock is covered by at least two analysts and zero otherwise. *SSTT* is the small size trade imbalance as in Hvidkjaer (2008). *HiloSprd* is the high-low spread estimate of Corwin and Schultz (2012). *INSTV* is the average of the eight most recent quarterly absolute institutional ownership percentage changes. *ILLIQV* is the idiosyncratic volatility of liquidity in Akbas, Armstrong and Petkova (2011). *PIN* is the probability of informed trade measured by Easley, Kiefer, O'Hara, and Paperman (1996). *Ret_Std* is the standard deviation of daily returns in a month. *SOIB*, *SPOIB, STURN, SStdTURN, SIVOL, SILLIQ, SDISP*, and *SRet_Std* are defined similarly as *SVOIB*. All variables are winsorized at the 0.5% and 99.5% levels. *N* is the average number of stocks per month. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Model	1	2	3	4	5	6	7	8	9
	raw ret	o-c ret	MA=3	MA=12	ex-crisis	post2001	pre2001	ABK	$\widehat{OIB}$
Intercept	3.625***	4.248***	2.719***	2.194***	2.252***	4.472***	0.748	2.271***	2.492***
	(4.07)	(3.50)	(4.83)	(3.48)	(3.51)	(6.01)	(0.94)	(3.73)	(4.15)
VOIB	4.434***	5.490***	3.294***	3.514***	3.612***	7.054***	1.961***	4.020***	2.695***
	(4.94)	(4.55)	(4.45)	(4.68)	(4.45)	(4.02)	(3.86)	(4.88)	(4.30)
SVOIB	-6.246***	-7.081***	-4.850***	-4.427***	-5.061***	-9.316***	-3.084***	-5.609***	-4.724***
	(-7.62)	(-6.28)	(-7.04)	(-6.28)	(-6.86)	(-6.35)	(-5.61)	(-7.33)	(-7.81)
ILLIQ	2.615***	3.489***	2.258***	1.981***	2.232***	2.494***	2.057***	2.233***	-0.249
	(5.66)	(5.64)	(5.15)	(3.88)	(4.68)	(4.15)	(3.23)	(4.97)	(-1.64)
SILLIQ	-1.254***	-2.506***	-1.243***	-0.764	-0.861*	-2.461***	0.037	-0.980**	1.002***
	(-3.04)	(-6.02)	(-3.35)	(-1.51)	(-1.96)	(-4.02)	(0.07)	(-2.35)	(3.02)
OIB	-0.308	-1.165*	0.028	-0.877	-0.613	-2.278***	0.750	-0.486	0.005
	(-0.66)	(-1.76)	(0.08)	(-1.40)	(-1.26)	(-2.64)	(1.61)	(-1.01)	(0.16)
SOIB	0.882**	1.843**	0.555	1.550***	1.078**	2.616***	-0.016	1.055**	0.001
	(2.04)	(2.57)	(1.64)	(2.72)	(2.43)	(3.15)	(-0.04)	(2.34)	(0.02)
POIB	-0.033	0.164	-0.068	-0.057	-0.015	0.055	-0.094	-0.033	2.293***
	(-0.49)	(1.60)	(-1.28)	(-0.67)	(-0.23)	(0.50)	(-1.19)	(-0.51)	(4.96)

Table A5 (Continued)

	raw ret	o-c ret	MA=3	MA=12	ex-crisis	post2001	pre2001	ABK	$\widehat{OIB}$
SPOIB	0.033	-0.085	0.079	0.033	0.021	-0.139	0.148*	0.032	-1.016**
	(0.46)	(-0.81)	(1.54)	(0.38)	(0.32)	(-1.36)	(1.95)	(0.50)	(-2.36)
SIZE	-0.183***	-0.218***	-0.193***	-0.159***	-0.163***	-0.261***	-0.094**	-0.163***	-0.178***
	(-3.63)	(-3.31)	(-5.80)	(-4.48)	(-4.50)	(-5.76)	(-2.05)	(-4.68)	(-5.26)
BM	0.126*	0.086	-0.013	0.010	-0.013	-0.013	-0.006	-0.009	-0.011
	(1.84)	(0.96)	(-0.27)	(0.18)	(-0.27)	(-0.20)	(-0.08)	(-0.18)	(-0.24)
R212	0.125	0.131	0.069	-0.205	0.210*	-0.627	0.381**	-0.030	-0.028
	(0.60)	(0.48)	(0.37)	(-1.02)	(1.76)	(-1.64)	(2.51)	(-0.16)	(-0.15)
R1	-0.042***	-0.021***	-0.048***	-0.048***	-0.048***	-0.037***	-0.057***	-0.049***	-0.049***
	(-9.31)	(-4.10)	(-10.22)	(-10.08)	(-9.95)	(-5.84)	(-9.14)	(-10.51)	(-10.52)
TURN	0.139	0.142	0.288***	0.120	0.193**	-0.044	0.347***	0.188**	0.129
	(1.56)	(1.18)	(3.33)	(1.28)	(2.23)	(-0.44)	(2.95)	(2.22)	(1.51)
STURN	0.849***	0.728***	0.593***	0.674***	0.669***	0.782***	0.634***	0.695***	0.725***
	(8.39)	(6.06)	(7.22)	(6.69)	(7.78)	(7.30)	(5.29)	(8.30)	(8.06)
STDTURN	-0.160**	-0.159*	-0.263***	-0.164***	-0.198***	-0.149*	-0.212***	-0.187***	-0.180***
	(-2.46)	(-1.66)	(-5.80)	(-3.29)	(-4.31)	(-1.95)	(-3.77)	(-4.06)	(-3.91)
SSTDTURN	-0.359**	-0.393*	0.118	-0.210*	-0.109	-0.199	-0.131	-0.158	-0.153
	(-2.27)	(-1.91)	(0.65)	(-1.86)	(-0.79)	(-1.06)	(-0.67)	(-1.14)	(-1.10)
IVOL	18.331	35.611	16.049	-0.005	26.321	56.337**	10.596	29.356	26.367
	(0.86)	(1.37)	(1.07)	(0.00)	(1.33)	(2.04)	(0.42)	(1.56)	(1.40)
SIVOL	-19.828	-36.941*	-11.480	-0.755	-24.221	-53.769**	-7.314	-26.318*	-24.728*
	(-1.19)	(-1.90)	(-1.09)	(-0.04)	(-1.55)	(-2.33)	(-0.39)	(-1.77)	(-1.66)
ACC	-1.148***	-1.070**	-1.162***	-1.160***	-1.094***	-0.757*	-1.334***	-1.098***	-1.117***
	(-3.14)	(-2.26)	(-3.42)	(-3.52)	(-3.14)	(-1.85)	(-2.80)	(-3.32)	(-3.37)
AG	-0.218***	-0.205***	-0.211***	-0.191***	-0.216***	-0.115*	-0.280***	-0.213***	-0.214***
	(-3.63)	(-2.81)	(-3.19)	(-3.30)	(-3.13)	(-1.67)	(-2.87)	(-3.27)	(-3.24)
ISSUE	-1.249***	-1.744***	-1.234***	-1.248***	-1.275***	-1.371***	-1.139***	-1.233***	-1.245***
	(-4.05)	(-3.72)	(-4.11)	(-4.02)	(-4.14)	(-3.81)	(-2.62)	(-4.15)	(-4.19)
PROFIT	0.002	0.086	0.055	0.095	0.077	-0.114	0.164*	0.052	0.064
	(0.02)	(0.98)	(0.83)	(1.33)	(1.18)	(-1.31)	(1.92)	(0.81)	(1.00)
SUE	0.029**	0.021	0.031**	0.024*	0.037***	0.033*	0.027*	0.030**	0.030**
	(2.11)	(1.17)	(2.49)	(1.90)	(3.15)	(1.74)	(1.66)	(2.41)	(2.47)
Max	5.430***	3.048**	4.913***	5.467***	5.145***	3.942**	5.530***	4.880***	4.818***
	(4.61)	(1.98)	(4.43)	(4.55)	(4.48)	(2.41)	(3.57)	(4.28)	(4.21)
DISP	-0.345***	-0.156	-0.365***	-0.395***	-0.376***	-0.370***	-0.358**	-0.363***	-0.356***
	(-2.92)	(-0.88)	(-3.91)	(-3.24)	(-3.41)	(-3.02)	(-2.32)	(-3.50)	(-3.44)
SDISP	0.025	0.020	0.128	0.082	0.088	0.324**	-0.070	0.090	0.093
	(0.20)	(0.09)	(1.13)	(0.64)	(0.71)	(1.99)	(-0.45)	(0.77)	(0.80)
DISPD	-0.251***	-0.165	-0.079	-0.121	-0.098	-0.119	-0.120	-0.118	-0.079
	(-2.74)	(-1.30)	(-1.09)	(-1.60)	(-1.32)	(-0.99)	(-1.31)	(-1.60)	(-1.11)
SSTT	-2.585	-32.408*	-2.384	-2.509	1.061	-6.857	4.797	0.215	4.495
	(-0.20)	(-1.92)	(-0.18)	(-0.19)	(0.08)	(-0.88)	(0.23)	(0.02)	(0.38)
HiLoSprd	29.964***	36.377***	30.024***	30.186***	29.648***	13.565	36.498***	27.039***	24.017***
	(4.08)	(3.56)	(3.65)	(4.55)	(3.61)	(1.52)	(3.20)	(3.46)	(3.09)
INSTV	-0.093	-0.173	-0.125**	-0.058	-0.108*	-0.183***	-0.058	-0.109**	-0.097*
	(-1.52)	(-0.96)	(-2.31)	(-0.78)	(-1.91)	(-3.19)	(-0.72)	(-2.05)	(-1.78)
ILLIQV	-0.047	0.030	-0.006	-0.017	-0.018	0.132	-0.130*	-0.024	-0.021
	(-0.72)	(0.29)	(-0.11)	(-0.29)	(-0.30)	(1.34)	(-1.87)	(-0.40)	(-0.35)
PIN	-4.362***	-7.153***	-3.166***	-3.572***	-2.972**	-11.840***	1.571*	-3.881***	-3.233***
	(-3.13)	(-3.94)	(-2.83)	(-3.05)	(-2.49)	(-5.83)	(1.96)	(-3.15)	(-2.87)
Ret_Std	-47.724**	-55.036**	-50.033***	-29.748	-57.849***	-67.172***	-48.985**	-56.40***	-51.814***
	(-2.39)	(-2.14)	(-3.73)	(-1.49)	(-3.32)	(-2.85)	(-2.18)	(-3.44)	(-3.15)
SRet_Std	10.716	21.395	13.285	-7.352	20.679	38.670**	6.143	19.369	16.768
	(0.70)	(1.17)	(1.45)	(-0.43)	(1.56)	(2.04)	(0.37)	(1.55)	(1.34)
Adj. R-sq	0.081	0.084	0.055	0.056	0.055	0.048	0.081	0.055	0.055
N	1561	1497	1576	1514	1528	2082	1198	1560	1560

Table A6: Fama-MacBeth Regressions Using Randomly Formed Portfolios

We perform an alternative test that accounts for measurement error in *OIB*. In this test, we use 20 randomly formed portfolios as test assets every month, and use the portfolios' average order flows to compute their *VOIB* and *SVOIB*. We then run Fama-MacBeth regressions for the 20 portfolios, using equally-weighted open-close quote midpoint returns to account for non-synchronous trading. We repeat this procedure 100 times. This table presents the time-series averages of individual stock cross-sectional OLS regression coefficient estimates between July 1983 and December 2012.

	SHR	*NUM*
Intercept	0.210	0.354
	(0.57)	(1.05)
VOIB	2.684***	2.894***
	(6.82)	(7.30)
SVOIB	-5.988***	-7.026***
	(-9.39)	(-10.57)

Table A7: Bivariate Portfolio Sorts Controlling for Limits to Arbitrage Proxies (Contemporaneous Relationship)

Each month between July 1983 and December 2012, stocks are first sorted into high and low groups based on one of the control variables, and then into *SVOIB_SHR* or *SVOIB_NUM* quintile portfolios within each control variable group. Then the quintile portfolio returns, the return differences between high and low quintile *SVOIB* portfolios, and the alphas using the Fama and French (1993) model along with the momentum factor and the Pastor and Stambaugh (2003) liquidity factor are reported with the Newey-West *t*-statistics in parentheses. *VOIB* is the standard deviation of daily order imbalance in a month, where order imbalance is defined *as (B−S)/(B+S)* with *B (S)* being the trades initiated by buyers (sellers). *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. The order imbalance is calculated using the number of shares traded in Panel A and using the number of trades in Panel B. *SIZE* represents the logarithm of market capitalization. *INST* is the percentage of shares held by institutional investors. *IVOL* is the idiosyncratic stock return volatility. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: *SVOIB_SHR*

	Controlling for *SIZE*			Controlling for *INST*			Controlling for *IVOL*		
SVOIB	Small	Large	Small-Large	Low	High	Low-High	Low	High	Low-High
Low	3.652	2.776		3.887	1.967		1.027	4.906	
2	1.043	1.704		1.694	1.292		0.869	2.209	
3	0.423	1.423		1.163	1.066		0.778	1.427	
4	0.146	1.378		0.763	0.929		0.647	1.067	
High	0.157	1.402		0.604	0.797		0.389	0.782	
High-Low	-3.495***	-1.374***	-2.121***	-3.283***	-1.170***	-2.112***	-0.638***	-4.124***	3.486***
	(-8.24)	(-6.75)	(-8.02)	(-8.12)	(-9.25)	(-6.41)	(-6.31)	(-8.70)	(8.24)
Alpha	-3.369***	-1.359***	-2.010***	-3.163***	-1.199***	-1.964***	-0.651***	-3.980***	3.329***
	(-8.45)	(-7.38)	(-7.23)	(-8.25)	(-9.69)	(-6.01)	(-6.55)	(-9.11)	(8.60)

Panel B: *SVOIB_NUM*

	Controlling for *SIZE*			Controlling for *INST*			Controlling for *IVOL*		
SVOIB	Small	Large	Small-Large	Low	High	Low-High	Low	High	Low-High
Low	3.204	2.655		3.498	1.756		0.943	4.363	
2	1.035	1.596		1.749	1.215		0.848	2.099	
3	0.341	1.407		1.030	1.106		0.838	1.387	
4	0.173	1.422		0.829	1.010		0.677	1.153	
High	0.668	1.603		1.005	0.965		0.404	1.390	
High-Low	-2.535***	-1.052***	-1.484***	-2.493***	-0.791***	-1.702***	-0.540***	-2.973***	2.433***
	(-5.83)	(-5.62)	(-4.93)	(-6.08)	(-5.55)	(-4.91)	(-5.35)	(-6.36)	(5.80)
Alpha	-2.449***	-1.062***	-1.387***	-2.432***	-0.815***	-1.617***	-0.539***	-2.861***	2.322***
	(-5.89)	(-6.00)	(-4.58)	(-6.20)	(-5.38)	(-4.67)	(-5.11)	(-6.51)	(5.88)

Table A8: Dynamic Effects of Liquidity Shocks: Robustness Checks

This table presents robustness checks for the effects of SVOIB on returns in the next 1, 2-3, 4-6, 7-9, and 10-12 months. We calculate *OIB* as the monthly order imbalance defined as *(B-S)/(B+S),* where *B (S)* is the number of shares traded initiated by buyers (sellers). *VOIB* is the standard deviation of daily order imbalance in a month. *SVOIB* is the difference between *VOIB* in the current month and the six-month moving average of *VOIB* in the previous month. *SILLIQ* is the shock to the Amihud illiquidity measure defined similarly as *SVOIB*. Panel A presents the results for the effects of *SVOIB_NUM*. We report the average Fama-French three-factor along with the momentum and liquidity factors adjusted returns of portfolios that buy stocks in the highest quintile of illiquidity shocks and sell stocks in the lowest quintile of illiquidity shocks, and the time-series averages of the coefficient estimates for the illiquidity shocks from cross-sectional regressions. In Panel B, we divide the sample into two groups based on the institutional holding (*INST*) or idiosyncratic volatility (*IVOL*), and report for each group the time-series averages of the coefficient estimates of *SVOIB_SHR* and *SILLIQ* from cross-sectional regressions. In Panel C, we divide the sample into two groups based on *SIZE*, the institutional holding or idiosyncratic volatility, and report the results for *SVOIB_NUM* All cross-sectional regressions control for the same variables as those in Column 3 of Table 5. For brevity, coefficient estimates are reported only for the illiquidity shock variables. All variables are winsorized at the 0.5% and 99.5% levels. Newey-West *t*-statistics are reported in parentheses. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Price impact in the long run

	Alpha		FM coefficient	
	SVOIB_NUM	*SILLIQ*	*SVOIB_NUM*	*SILLIQ*
Month1	-0.945***	-1.534***	-5.619***	-0.979**
	(-7.49)	(-9.08)	(-7.34)	(-2.30)
Month2-3	0.058	-0.639***	0.286	-1.232***
	(0.50)	(-4.47)	(1.22)	(-3.45)
Month4-6	0.234**	-0.178	0.163	-1.655***
	(2.20)	(-1.29)	(0.89)	(-4.84)
Month7-9	0.241***	-0.094	0.465**	-1.496*
	(2.69)	(-0.62)	(2.05)	(-1.86)
Month10-12	0.339***	0.148	0.557***	-0.816
	(4.22)	(0.96)	(3.24)	(-0.94)

Panel B: Fama-MacBeth regression coefficient estimates conditioning on *INST* and *IVOL*

	INST				*IVOL*			
	Low		High		Low		High	
	SVOIB_SHR	*SILLIQ*	*SVOIB_SHR*	*SILLIQ*	*SVOIB_SHR*	*SILLIQ*	*SVOIB_SHR*	*SILLIQ*
Month1	-4.401***	-0.465*	-3.110***	-25.156	-1.655***	-0.893	-3.941***	-0.331
	(-4.89)	(-1.87)	(-4.57)	(-1.41)	(-6.55)	(-0.50)	(-8.83)	(-1.53)
Month2-3	-1.289***	-0.528***	0.887*	-11.310***	0.483**	-0.121	-1.286***	-0.612***
	(-2.94)	(-2.71)	(1.79)	(-3.35)	(2.41)	(-0.10)	(-3.84)	(-3.43)
Month4-6	0.912***	-0.456*	1.053**	-1.186	0.469**	-0.726	0.920***	-0.145
	(3.51)	(-1.83)	(1.97)	(-0.59)	(2.52)	(-0.72)	(3.31)	(-0.70)
Month7-9	1.344***	-0.158	0.836**	-3.583	0.442**	-0.251	1.294***	-0.041
	(4.89)	(-0.73)	(2.18)	(-1.13)	(2.50)	(-0.17)	(4.49)	(-0.23)
Month10-12	1.321***	0.317	0.694**	-6.289**	0.494***	0.213	1.861***	0.611
	(4.92)	(1.31)	(2.04)	(-2.05)	(2.77)	(0.16)	(6.51)	(1.47)

A8 (continued)

Panel C: Fama-MacBeth regression coefficient estimates for *SVOIB_NUM* conditioning on *SIZE, INST* and *IVOL*

	SIZE		*INST*		*IVOL*	
	Small	Large	Low	High	Low	High
Month1	-4.826***	-2.993***	-5.170***	-3.415***	-1.249***	-4.826***
	(-5.60)	(-2.79)	(-5.89)	(-4.33)	(-3.63)	(-9.90)
Month2-3	-0.665**	0.137	-0.473	-1.056	-0.334	-2.279***
	(-2.13)	(0.27)	(-1.48)	(-1.47)	(-1.61)	(-4.80)
Month4-6	0.735***	0.653	0.846***	1.192*	0.344*	1.120***
	(2.77)	(1.55)	(2.99)	(1.69)	(1.86)	(3.63)
Month7-9	1.211***	0.695**	1.446***	1.109**	0.733***	1.374***
	(4.86)	(2.00)	(4.78)	(2.52)	(3.77)	(4.41)
Month10-12	1.177***	0.729**	1.176***	0.571	0.540***	1.861***
	(4.80)	(2.10)	(4.41)	(1.29)	(2.74)	(6.25)