JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 46, No. 3, June 2011, pp. 709–736
COPYRIGHT 2011, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S0022109011000081

Liquidity Dynamics and Cross-Autocorrelations

Tarun Chordia, Asani Sarkar, and Avanidhar Subrahmanyam*

Abstract

This paper examines the relation between information transmission and cross-autocorrelations. We present a simple model, where informed trading is transmitted from large to small stocks with a lag. In equilibrium, large stock illiquidity induced by informed trading portends stronger cross-autocorrelations. Empirically, we find that the lead-lag relation increases with lagged large stock illiquidity. Further, the lead from large stock order flows to small stock returns is stronger when large stock spreads are higher. In addition, this lead-lag relation is stronger before macro announcements (when information-based trading is more likely) and weaker afterward (when information asymmetries are lower).

I. Introduction

The manner in which information is incorporated into prices has long been of primary concern in financial economics. While asset pricing theories in the presence of frictionless markets allow for instantaneous diffusion of information, trading frictions impede the smooth incorporation of information into prices. Understanding the impact of such frictions on price dynamics is of fundamental importance. Lo and MacKinlay (LM) (1990) document an important pattern wherein the correlation between lagged large firm returns and current small firm returns is higher than the correlation between lagged small firm returns and current large firm returns. This points to some friction that prevents information from

*Chordia, tarun_chordia@bus.emory.edu, Goizueta Business School, Emory University, 1300 Clifton Rd., Atlanta, GA 30322; Sarkar, asani.sarkar@ny.frb.org, Federal Reserve Bank of New York, 33 Liberty St., New York, NY 10038; and Subrahmanyam, subra@anderson.ucla.edu, Anderson Graduate School of Management, University of California at Los Angeles, 110 Westwood Plz., Los Angeles, CA 90095. We thank an anonymous referee, Hendrik Bessembinder (the editor), Jeff Coles, Arturo Estrella, Michael Fleming, Akiko Fujimoto, Kenneth Garbade, Joel Hasbrouck, Charles Himmelberg, Mark Huson, Aditya Kaul, Spencer Martin, Vikas Mehrotra, Albert Menkveld, Federico Nardari, Christine Parlour, Ľuboš Pástor, Sebastien Pouget, Joshua Rosenberg, Christoph Schenzler, Gordon Sick, Rob Stambaugh, Neal Stoughton, Marie Sushka, Jiang Wang, and seminar participants at Arizona State University, University of Alberta, University of Calgary, the HEC (Montreal) Conference on New Financial Market Structures, and the Federal Reserve Bank of New York for helpful comments. The views stated here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

being incorporated into all stock prices simultaneously. Brennan, Jegadeesh, and Swaminathan (BJS) (1993) propose that the lead-lag patterns arise due to differential speeds of adjustment of stocks to economywide shocks.[1] Thus, systematic demand shocks should first be incorporated into the prices of the more actively traded larger stocks, resulting in the observed cross-autocorrelation patterns in stock returns.

The speed of adjustment hypothesis finds considerable support in the data (viz., BJS (1993), Chordia and Swaminathan (2000)).[2] However, the hypothesis permits multiple dynamic mechanisms. Specifically, differential speeds of adjustment could arise because of lags in transmission of common information, imperfect arbitrage across large and small stocks due to market frictions, or lags in transmission of informationless demand shocks with a systematic component. In this paper, we develop a simple model for cross-autocorrelations and conduct empirical tests to distinguish our economic rationale from alternative explanations for this phenomenon.

Our basic argument is as follows: We begin with the premise that common information is first traded upon in the more actively traded large stocks. This increases large stock illiquidity in equilibrium. As the informational event is transmitted to small stocks with a delay, low large stock liquidity implies a stronger lead-lag relation from large to small firms. Thus, our hypothesis indicates that greater illiquidity of large stocks is associated with stronger cross-autocorrelations.

A different rationale for leads and lags is that illiquidity causes differential frictions in the transmission of public information. Low liquidity of small stocks can create frictions for arbitrageurs that seek to close the pricing gap between large and small firms. This suggests that the lead-lag effect would increase when small stock illiquidity is high. Alternatively, informationless but random demands with a systematic component (as in Barberis and Shleifer (2003)) may affect the relatively liquid large firms first, and then affect small firms. Under this scenario, high large stock liquidity (because of informationless trading) would lead to an increase in the lead-lag relation. Thus, these mechanisms imply 3 possible relations between liquidity and the lead-lag patterns: i) lead-lag patterns are stronger when liquidity is low in the large stocks, ii) lead-lag patterns are stronger when liquidity is low in the small stocks, or iii) lead-lag patterns are stronger when liquidity is high in the large stocks. We use stock liquidity data to distinguish between these alternative mechanisms for cross-autocorrelations.

[1]Other important explanations have been proposed for lead and lag patterns. The cross-autocorrelations may be a result of nonsynchronous trading, the idea being that the information that has already been incorporated into prices of large firms is incorporated into small firms' prices only when they trade, which due to thin trading would, on average, occur with a lag. Another set of explanations suggests that the cross-autocorrelations are a restatement of portfolio autocorrelation and contemporaneous correlations (Boudoukh, Richardson, and Whitelaw (1994), Hameed (1997)). Yet another set of explanations suggests that the predictability of small firm returns could arise due to time-varying risk premia (Conrad and Kaul (1998)). LM (1990) as well as McQueen, Pinegar, and Thorley (1996) argue that these explanations are able to account partially, but not fully, for cross-autocorrelations.

[2]See also Badrinath, Kale, and Noe (1995), Connolly and Stivers (1997), and Hou (2007).

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

We note that in recent years, there has been interest in the general notion that liquidity is related to market efficiency; specifically, to the speed with which financial markets incorporate information (see, e.g., Mitchell, Pulvino, and Stafford (2002), Sadka and Scherbina (2007), Hou and Moskowitz (2005), Avramov, Chordia, and Goyal (2006), and Chordia, Roll, and Subrahmanyam (2008)). This work typically documents own-sector or own-stock linkages between market efficiency (as reflected in the time series of returns) and measures of liquidity. In contrast, we focus in this paper on the hitherto unexplored relation between liquidity and return cross-autocorrelations.

Using vector autoregressions (VARs) over a long time series of data encompassing more than 4,500 trading days, we establish that while there are persistent liquidity spillovers across size deciles, the returns of large stocks lead those of small stocks even after accounting for liquidity dynamics. Moreover, we document a new empirical result that validates our model: that cross-autocorrelation patterns in returns are strongest when large stock illiquidity (as measured by the bid-ask spread) is high. Further, order flows in the large stocks play an important role in predicting small stock returns when large stock liquidity is low. On the other hand, small stock order flows or spreads do not predict large stock returns. Our results hold for both midquote as well as transactions-based returns. The hitherto undocumented role of large stock order flow and large stock liquidity in predicting small firm returns is consistent with our theoretical setting and accords with our first hypothesis: that informed trading in the large stocks leads to the cross-autocorrelation patterns in stock returns. The other 2 hypotheses proposed previously are not supported by the data.

In order to verify that it is indeed trading based on common information that leads to the cross-autocorrelations, we examine the impact of 2 major macroeconomic events—namely, nonfarm employment and Consumer Price Index (CPI) announcements—on the lead-lag effect. Our consideration of macro events is motivated by recent research (Brandt, and Kavajecz (2004), Beber, Brandt, and Kavajecz (2008), and Evans and Lyons (2002), (2008)) that order flows contain information about the macroeconomy. We find that when the lagged order imbalance (OIB) and lagged spreads of large stocks are higher, the returns of small stocks are also higher prior to the announcement than during the rest of the sample period. Further, spread and order flows of large stocks predict small stock returns less strongly following the announcements, when information asymmetry is likely lower. These results point to the role of privately observed systematic information in causing the lead-lag pattern. Overall, the analysis indicates that it is the trading in large stocks that precedes common information shocks, which leads to the predictability of small stock returns.

While our results point to a significant role for large stock liquidity and order flow in predicting small stock returns, we find a limited role for small stock liquidity/order flows. This is consistent with Hou and Moskowitz (2005) who also find no role for own-asset liquidity in explaining cross-sectional return predictability. Our results demonstrate that discussions of the role of liquidity in explaining the sources of market efficiency should be broadened to include liquidity spillovers between asset classes. This point is implicitly recognized by researchers

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

in the context of financial crises, when discussions of liquidity spillovers become paramount.

We conduct a robustness check using a holdout sample of the relatively smaller NASDAQ stocks, whose liquidity is obtained by using closing bid and ask quotes available from the Center for Research in Security Prices (CRSP). This analysis indicates that the broad thrust of our spillover results obtains for NASDAQ stocks as well. Notably, large stock order flows strongly predict NASDAQ returns when large stock spreads are high.

We also examine whether our results hold at the industry level. Specifically, we test whether trading on information that is common to stocks within an industry (over and above marketwide information), and thus, industry-level order flow and liquidity play a role in causing cross-autocorrelations. To do this, we use 5 industry classifications proposed by Ken French (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/siccodes5.zip) and examine leads and lags across small and large firms within each industry, controlling for marketwide returns. We find that large firms lead small firms within each of the industries, and this lead-lag effect is accounted for by the interaction of large stock order flow with large stock spreads within each industry group. Overall, the results lend support to our proposed mechanism that information-based trading (at both the marketwide and industry levels) occurs first in the large stocks (widening spreads in that sector) and then spills over to the small stocks.

The rest of the paper is organized as follows. Section II provides a brief economic motivation for the specific hypothesis that is the focus of our study. Section III describes how the liquidity data are generated, while Section IV presents basic time-series properties of the data and describes the adjustment process to stationarize the series. Section V presents the VARs involving order flows, liquidity, returns, and volatility across large and small stocks. Section VI considers the role of liquidity in the lead-lag relation between large and small stock returns. Section VII discusses robustness checks using a holdout sample of NASDAQ stocks. Section VIII conducts industry-specific analyses, and Section IX concludes.

II. The Economic Setting

Our principal aim in this section is to establish an equilibrium link between liquidity and cross-autocorrelation via an informational channel. This analysis allows us to delineate a specific hypothesis that we test in subsequent sections. We start with the basic premise that agents with information about common factors choose to exploit their informational advantage in the relatively more liquid large stocks. Lagged quote updating by small stock market makers causes small stock returns to lag those of large stocks (viz., Chan (1993), Chowdhry and Nanda (1991), Gorton and Pennacchi (1993), and Kumar and Seppi (1994)). The informed trading that causes the lead-lag in the previous line of argument reduces liquidity temporarily in the large stocks. Liquidity decreases in the large stocks therefore portend a lagged adjustment of small firm returns to large firm returns. Further, because part of the equilibrium order flow arises due to informed

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

trading, large stock order flows and large stock returns both play important roles in predicting small stock returns.

We consider 2 securities, labeled L and S, which are traded over T periods in a Kyle (1985) and Admati and Pfleiderer (1988) setting. Each security's final payoff is given by

$$F_k = \bar{F}_k + \sum_{t=1}^{T} (\beta_k \delta_t + \epsilon_{kt}),$$

for $k = L, S$. Here $\bar{F}_k$ represents the nonstochastic unconditional means of the assets, whereas β_k represents the security's beta. Further, each innovation δ_t and ϵ_{kt} is revealed after trade at time t and prior to trade at time $t+1$. Each innovation is independent of all other random variables, and all random variables are mutually independent with mean 0. Each ϵ_{kt} has a variance v_ϵ.

We propose that each δ_t can be subdivided into M independent and identically distributed subinnovations such that $\delta_t = \sum_{m=1}^{M} \delta_{mt}$. Each of these subinnovations has a variance v_δ. Further, during any period t, I_t agents acquire factor information. Each of these I_t agents acquires information about exactly one subinnovation that is different from that acquired by the other $I_t - 1$ agents (where $I_t < M$). Further, a single agent in each market observes the realization of the firm-specific component ϵ_{kt}.[3] There is liquidity trading in the amount of z_{kt}, $k = L, S$ in each subperiod. The liquidity trade innovations are also mean 0 and independent of all other random variables in the model, with $\text{var}(z_{kt}) = v_{kz}$.

We use the well-known stylized fact (e.g., Gompers and Metrick (2001)) that institutions are most heavily invested in large stocks. These agents are those most likely to trade on factor information. We therefore assume that the factor-informed agents trade only in the large stocks.[4] Prices in each sector are set by competitive, risk-neutral market makers who observe the order flows in their own market contemporaneously but in the other market with a lag. We have that

$$P_{kt} = \text{E}(F_k | \phi_{kt}),$$

where ϕ_{kt} is the information set of the market maker in security k at time t. Let ω_{kt} denote the total order flow in security k at time t. We propose that the price at each date takes the linear form

$$P_{kt} = \bar{F}_k + \sum_{\tau=1}^{t-1} (\beta_k \delta_\tau + \epsilon_{k\tau}) + \lambda_{kt} \omega_{kt}.$$

Given this pricing function, each factor-informed agent i_t at time t maximizes

$$\text{E}[x_{it}(F_L - P_{Lt}) | \delta_{it}],$$

[3] The dichotomization of informed agents into marketwide and firm-specific agents is for notational simplicity and has no significant impact on our economic intuition.

[4] Bernhardt and Mahani (2007) point out the importance of market segmentation in information-based explanations for cross-autocorrelations.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

and each firm-specific informed agent maximizes

$$\mathrm{E}[y_{kt}(F_k - P_{kt})|\epsilon_{kt}].$$

Thus, each factor-informed agent submits an order

$$x_{it} = [\beta_L \delta_{it}]/[2\lambda_{Lt}]$$

in the large stocks, whereas the agent with firm-specific information submits an order

$$y_{kt} = \epsilon_{kt}/(2\lambda_{kt}).$$

In each market, the liquidity parameter is given by

$$\lambda_{kt} = \frac{\mathrm{cov}(F_k, \omega_{kt})}{\mathrm{var}(\omega_{kt})}. \tag{1}$$

Note that $\omega_{Lt} = \sum_{i=1}^{I_t} x_{it} + y_{Lt} + z_{Lt}$ and $\omega_{St} = y_{St} + z_{St}$.

Observing that a positive λ_{kt} is necessary for satisfying the 2nd-order condition, substituting for the order flows in equation (1) implies that

$$\lambda_{Lt} = 0.5\sqrt{\frac{I_t \beta_L^2 v_\delta + v_\epsilon}{v_{Lz}}}, \tag{2}$$

$$\lambda_{St} = 0.5\sqrt{\frac{v_\epsilon}{v_{Sz}}}. \tag{3}$$

Note that the illiquidity parameter in the large firm sector increases in the number of agents with factor information.

It is obvious that there is no lead from small firms to large firms, because there is no useful information in small firm order flow that is useful for forecasting the cash flows of the large firms. The lead-lag coefficient from large to small firms in this market, denoted by γ_t, and expressed in terms of price changes, is

$$\gamma_t = \frac{\mathrm{cov}(P_{St+1} - P_{St}, P_{Lt} - P_{Lt-1})}{\mathrm{var}(P_{Lt} - P_{Lt-1})}.$$

It is easy to verify that the previous covariance reduces to

$$\gamma_t = \frac{\mathrm{cov}(\beta_S \delta_t, \lambda_{Lt}\omega_{Lt})}{\mathrm{var}(\lambda_{Lt}\omega_{Lt})}. \tag{4}$$

Now, for notational convenience, let us denote $\sum_{j=1}^{I_t} \delta_{jt} \equiv \eta_t$, with $v_\eta^t = \mathrm{var}(\eta_t)$. We know that

$$\omega_{Lt} = [\beta_L \eta_t + \epsilon_{Lt}]/[2\lambda_{Lt}] + z_{Lt}.$$

Thus, we have

$$\gamma_t = \frac{\mathrm{cov}\{\beta_S \delta_t, 0.5(\beta_L \eta_t + \epsilon_{Lt}) + \lambda_{Lt} z_{Lt}\}}{\mathrm{var}\{0.5(\beta_L \eta_t + \epsilon_{Lt}) + \lambda_{Lt} z_{Lt}\}}.$$

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

Substituting for λ_{Lt} from equation (2), the previous equation reduces to

$$\gamma_t = \frac{\beta_S}{\beta_L} \times \frac{\beta_L^2 v_\eta^t}{\beta_L^2 v_\eta^t + v_\epsilon}$$

or, equivalently,

$$\gamma_t = \frac{\beta_S}{\beta_L} \times \frac{\beta_L^2 I_t v_\delta}{\beta_L^2 I_t v_\delta + v_\epsilon}.$$

It can be seen that γ_t is increasing in I_t, and so is λ_{Lt}. This leads us to the following proposition.

Proposition 1. There is dynamic variation in the lead-lag relationship between large and small firms. The strength of this relation at a point in time is inversely related to large firm liquidity at that time.

Another coefficient of interest is the relation between small firm returns and lagged large firm order flows. Let us denote this coefficient by κ_t. It is easy to verify that

$$\kappa_t = \frac{\text{cov}(P_{St+1} - P_{St}, \omega_{Lt})}{\text{var}(\omega_{Lt})} = \lambda_{Lt}\gamma_t.$$

Since we already know that γ_t is increasing in I_t and so is λ_{Lt}, κ_t will also vary positively with market illiquidity. Overall, the basic notion is that high illiquidity implies that informed agents are active in the large stocks, which in turn increases the informativeness of the lagged order flow (and thus lagged large stock price changes). In turn, this increases the predictive ability of large firm order flows and returns for small firm returns.

The previous economic argument is not the only possible link between liquidity and cross-autocorrelations. For example, our framework does not allow for arbitrage. However, if market makers update quotes with a lag, arbitrageurs may choose to trade in small stocks in order to profit from common information shocks that have already been incorporated into the prices of large firms. An exogenous widening of small firm spreads can create greater frictions for arbitrageurs that seek to close the pricing gap between large and small firms. This simple argument suggests that the lead-lag effect would increase when small firm spreads are high.

While the previous argument suggests that lagged small stock spreads play a crucial role in determining the extent of the lead-lag relationship, our model suggests that lagged large stock spreads are relevant for the lead-lag effect. The two arguments are not mutually exclusive. Moreover, yet another hypothesis is that random liquidity demands with a systematic component (as in Barberis and Shleifer (2003)) are realized in large stocks first and then in small firms. In this case, the liquidity trading in large firms would increase large firm liquidity, so that high large firm liquidity would portend increased cross-autocorrelations.

In order to distinguish between these explanations, we analyze the link between the extent of the lead-lag relationship and the levels of large and small firm spreads.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

III. Data and Basic Methodology

Stock liquidity data were obtained for the period January 1, 1988 to December 31, 2007. The data sources are the Institute for the Study of Securities Markets (ISSM) and the New York Stock Exchange (NYSE) TAQ (trades and automated quotations). The ISSM data cover 1988–1992, inclusive, while the TAQ data are for 1993–2007.[5]

A relevant issue in our analysis is the liquidity measure that is appropriate for our purposes. We analyze data at the daily frequency for both large and small stocks. Therefore, thin trading in the small stocks precludes estimating the regression-based Kyle (1985) lambdas, which require high transaction frequency for reliable estimation. Instead, we extract measures whose reliability is less susceptible to trading frequency. These measures are: i) quoted spread (QSPR), measured as the difference between the inside bid and ask quote; ii) relative or proportional quoted spread (RQSPR), measured as the quoted spread divided by the midpoint of the bid-ask spread; iii) effective spread (ESPR), measured as twice the absolute value of the difference between the transaction price and the midpoint of the bid and ask prices; and iv) relative effective spread (RESPR), which is the effective spread divided by the midpoint of the spread. The transactions-based liquidity measures are averaged over the day to obtain daily liquidity measures for each stock. The daily OIB (used in Sections VI and VII) is defined as the dollar value of shares bought less the dollar value of shares sold divided by the total dollar value of shares traded. Imbalances are calculated using the Lee and Ready (1991) algorithm to sign buy and sell trades.

Once the individual stock liquidity data are assembled, in each calendar year the stocks are divided into deciles by their market capitalization on the last trading day of the previous year (obtained from CRSP). Value-weighted daily averages of liquidity are then obtained for each decile, and daily time series of liquidity are constructed for the entire sample period. The largest firm group is denoted decile 9, while decile 0 denotes the smallest firm group. The average market capitalizations across the deciles range from about $26 billion for the largest decile to about $47 million for the smallest decile.[6] Since any cross-sectional differences in liquidity dynamics would be most manifest in the extreme deciles, we mainly present results for deciles 9 and 0, allowing us to present our analysis parsimoniously. When relevant, however, we also discuss results for other deciles.

Our economic arguments center around the relation between liquidity and cross-autocorrelations. However, our liquidity measures are necessarily imperfect. Since there is an intimate link between liquidity and volatility (Benston and Hagerman (1974)), we also consider volatility in our analysis. Volatility potentially

[5]The following securities were not included in the sample since their trading characteristics might differ from ordinary equities: American Depositary Receipts (ADRs), shares of beneficial interest, units, companies incorporated outside the U.S., Americus Trust components, closed-end funds, preferred stocks, and real estate investment trusts (REITs). Stocks with prices over $999 were removed from the sample. We also apply the filter rules in Chordia, Roll, and Subrahmanyam (2001) while extracting the transactions data.

[6]For the middle 8 deciles, the average market capitalizations (in billions of dollars) are 5.05, 2.56, 1.48, 0.94, 0.61, 0.39, 0.24, and 0.13.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

captures elements of liquidity not captured by our measures, and it also allows us to ascertain whether our conclusions survive after accounting for the risk-return relationship. Specifically, we consider the joint dynamics of liquidity, returns, and volatility using a VAR model.

Following Schwert (1990), Jones, Kaul, and Lipson (1994), Chan and Fong (2000), and Chordia, Sarkar, and Subrahmanyam (2005), daily return volatility (VOL) is obtained as the absolute value of the residual from the following regression for decile i on day t:

$$R_{it} = a_1 + \sum_{j=1}^{4} a_{2j} D_j + \sum_{j=1}^{12} a_{3j} R_{it-j} + e_{it}, \tag{5}$$

where D_j is a dummy variable for the day of the week, and R_{it} (the return series used in the analysis) represents the value-weighted average of individual stock returns for a particular decile (using market capitalization as of the end of the previous year as weights, as in the case of spreads). The returns are measured using the midpoint of bid-ask quotes matched to the last transaction of the day.[7] We use midquotes instead of transaction prices to avoid bid-ask bounce issues.

IV. Properties of the Data

A. Summary Statistics

In Table 1, we present summary statistics associated with liquidity measures, together with information on the daily number of transactions for the 2 size deciles. Since previous studies such as Chordia et al. (2001) suggest that the reduction in tick sizes likely had a major impact on bid-ask spreads, we provide separate statistics for the periods before and after the 2 changes to sixteenths and decimalization.[8] We find that the mean quoted spreads for large and small stocks are very close to each other (20.0 and 20.3 cents, respectively) before the shift to sixteenths, but they diverge considerably after the shift. Indeed, the average spread for large stocks is less than half that of small stocks (2.3 vs. 6.9 cents) in the period following decimalization. The point estimates indicate that decimalization has been accompanied by a substantial and statistically significant[9] reduction in the spreads of large stocks, which is consistent with the prediction of Ball and Chordia (2001).

The effective spreads follow a pattern similar to the quoted spreads. Before the shift to the sixteenths, the average small (large) stock effective spreads were

[7] Following Lee and Ready (1991), from 1993–1998 inclusive, the matching quote is the first quote at least 5 seconds prior to the trade. Due to a generally accepted decline in reporting errors in recent times (see, e.g., Madhavan, Ming, Straser, and Wang (2002)), after 1998, the matching quote is simply the first quote prior to the trades.

[8] Chordia, Shivakumar, and Subrahmanyam (2004) provide similar statistics for size-based quartiles, but they do not present statistics for the post-decimalization period, since their sample ends in 1998.

[9] Unless otherwise stated, "significant" is construed as "significant at the 5% level or less" throughout the paper.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

TABLE 1

Levels of Stock Market Liquidity

The stock liquidity series in Table 1 are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. QSPR stands for quoted spread, RQSPR represents the relative quoted spread computed as the quoted spread divided by the midpoint of the spread, ESPR denotes the effective spread and is computed as twice the absolute value of the difference between the transaction price and the midpoint of the bid and ask prices, RESPR is the relative effective spread computed as the effective spread divided by the spread midpoint, and NTRADE represents the number of shares traded. The suffixes "0" and "9" represent the smallest and largest size deciles, respectively. The stock data series excludes September 4, 1991, on which no trades were reported in the transactions database. The mean, median, and standard deviation (SD) are reported for each measure. The sample spans the period January 4, 1988 to December 31, 2007; the number of observations is 5,041 for each decile.

	1/4/1988–6/23/1997			6/24/1997–1/28/2001			1/29/2001–12/31/2007		
Stock Market Liquidity	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
QSPR0	0.203	0.199	0.019	0.167	0.166	0.009	0.076	0.069	0.021
ESPR0	0.146	0.133	0.026	0.111	0.111	0.008	0.055	0.053	0.013
RQSPR0	0.037	0.036	0.010	0.030	0.030	0.007	0.011	0.007	0.008
RESPR0	0.028	0.027	0.006	0.021	0.020	0.004	0.008	0.005	0.006
NTRADE0	11.447	10.033	5.289	29.300	27.790	9.541	222.218	151.937	245.443
QSPR9	0.200	0.200	0.028	0.128	0.124	0.013	0.030	0.023	0.014
ESPR9	0.140	0.138	0.023	0.083	0.083	0.009	0.021	0.017	0.009
RQSPR9	0.004	0.004	0.001	0.002	0.002	0.000	0.001	0.000	0.000
RESPR9	0.003	0.003	0.001	0.001	0.001	0.000	0.000	0.000	0.000
NTRADE9	447.213	444.840	306.250	2,471.252	2,368.591	1,025.346	10,500.120	5,052.976	12,820.280

14.6 (14.0) cents; during the sixteenths period the effective spreads were 11.1 (8.3) cents; and during the decimal period they were 5.5 (2.1) cents. The relative quoted and the relative effective spreads for the larger stocks are generally orders of magnitude smaller than those for small stocks. For instance, the average relative quoted (effective) spread for small and large stocks are 3.7% (2.8%) and 0.4% (0.3%) cents, respectively. Thus, all else being equal, higher priced stocks have higher spreads. The average daily number of transactions has increased substantially in recent years for both large and small stocks. For example, the average daily number of transactions for the large stocks increased from 447 in the first subperiod (before the shift to sixteenths) to 10,500 in the last subperiod (post decimalization), and this difference, not surprisingly, is statistically significant.

Figure 1 plots the time series for quoted spreads for the largest and smallest deciles. The figure clearly documents the declines caused by 2 changes in the tick size and also demonstrates how large stock spreads have diverged from those of small stocks toward the end of the sample period. In the remainder of the paper, we focus primarily on raw spreads that are not scaled by price because we do not want to contaminate our inferences by attributing movements in stock prices to movements in liquidity. We also focus on quoted spreads, because results for effective spreads are very similar to those for quoted spreads.

B. Adjustment of Time-Series Data on Liquidity, Returns, and Volatility

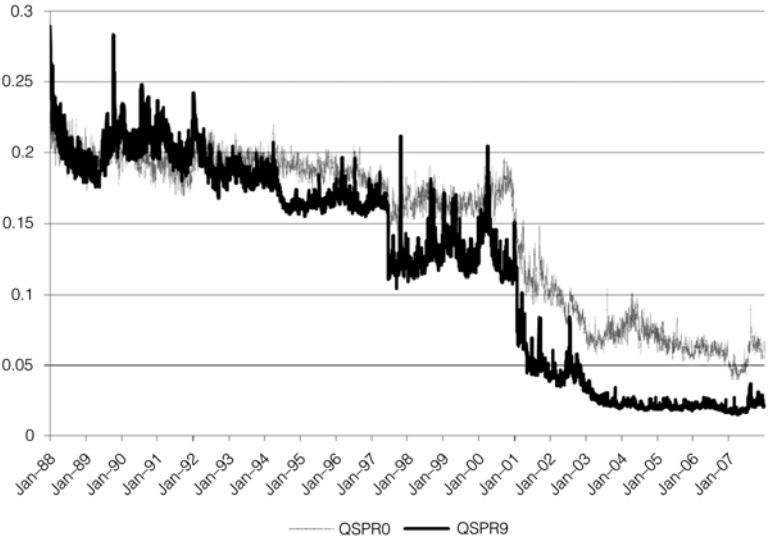
Liquidity across stocks may be subject to deterministic movements such as time trends and calendar regularities. Since we do not wish to pick up such predictable effects in our time-series analysis, we adjust the raw data for deterministic

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

FIGURE 1

Quoted Spreads for Large and Small Stocks

In Figure 1 quoted bid-ask spreads for NYSE stocks are plotted for the period 1988–2007. QSPR9 is the spread for the largest firm decile, and QSPR0 is the spread for the smallest firm decile.



time-series variations. The quoted spreads, returns, and volatility are transformed by the method proposed by Gallant, Rossi, and Tauchen (1992). Details of the adjustment process are available in the Appendix.

Table 2 presents selected regression coefficients for the quoted spread. For the sake of brevity, we only present the coefficients for selected calendar regularities in quoted spreads. We do not present results for other variables.

A readily noticeable finding is that the nature of calendar regularities in liquidity is different across large and small stocks. For example, spreads for larger stocks are higher at the beginning of the year than in other months (all dummy coefficients from March to December are negative and significant for large firms). This regularity is reversed for small stocks.[10] The January behavior in spreads may be due to the fact that portfolio managers shift out of the large stocks following window dressing in December. In addition, there has been a strong negative trend in spreads since decimalization for both small and large companies.

To examine the presence of unit roots in the adjusted series, we conduct augmented Dickey-Fuller and Phillips-Perron tests. We allow for an intercept under the alternative hypothesis, and we use the information criteria to guide selection of the augmentation lags. We easily reject the unit-root hypothesis for every series (including those for volatility and imbalances), generally with p-values less than 0.01. For the remainder of the paper, we analyze these adjusted series, and all references to the original variables refer to the adjusted time series of the variables.

[10]Clark, McConnell, and Singh (1992) document a decline in spreads from the end of December through the end of January, but they do not compare seasonals for large and small stocks explicitly.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

TABLE 2

Adjustment Regressions for Liquidity

The stock liquidity series in Table 2 are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. QSPR stands for quoted spread. The sample spans the period January 4, 1988–December 31, 2007; the number of observations is 5,041 for all deciles. The stock data series excludes September 4, 1991, on which no trades were reported in the transactions database. The suffixes "0" and "9" represent the smallest and largest size deciles, respectively. Holiday: a dummy variable that equals 1 if a trading day satisfies the following conditions, i) if Independence Day, Veterans' Day, Christmas, or New Year's Day falls on a Friday, then the preceding Thursday; ii) if any holiday falls on a weekend or on a Monday then the following Tuesday; iii) if any holiday falls on a weekday then the preceding and the following days, and 0 otherwise. Monday–Thursday: equals 1 if the trading day is Monday, Tuesday, Wednesday, or Thursday, and 0 otherwise. February–December: equals 1 if the trading day is in 1 of these months, and 0 otherwise. The tick size change dummy variable equals 1 for the period June 24, 1997–January 28, 2001. The decimalization dummy variable equals 1 for the period January 29, 2001–December 31, 2007. Estimation is done using the ordinary least squares (OLS). All coefficients are multiplied by a factor of 100. Estimates marked ** and * are significant at the 1% and 5% levels or better, respectively.

Independent Variables	QSPR0	QSPR9
Intercept	20.073**	21.816**
Holiday	0.297**	–0.010
Month		
February	0.151**	–0.118
March	0.237**	–0.290**
April	0.297**	–0.197**
May	0.092	–0.633**
June	0.067	–0.773**
July	0.184**	–0.722**
August	0.363**	–0.755**
September	0.280**	–0.945**
October	0.385**	–0.474**
November	0.169**	–0.809**
December	–0.084	–0.742**
Tick size change dummy	–3.706**	–10.734**
Decimalization dummy	–8.274**	–14.713**
Trend, pre-tick size change		
Time	–0.001**	–0.001**
$Time^2$	0.000**	0.000**
Trend, pre-decimalization		
Time	–0.001	0.011**
$Time^2$	0.000**	0.000**
Trend, post-decimalization		
Time	–0.008**	–0.008**
$Time^2$	0.000**	0.000**

V. Cross-Autocorrelations

We begin by examining the nature of the lead-lag relationship between small and large firms within our sample. We adopt a 6-equation VAR that incorporates 6 variables (3 each, i.e., measures of liquidity, returns, and volatility, from large and small stocks). The VAR thus includes the endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9, whose suffixes 0 and 9 denote the size deciles, with 0 representing the smallest size decile and 9 the largest.

We choose the number of lags in the VAR on the basis of the Akaike information criterion (AIC) and the Schwarz information criterion (SIC). Where these two criteria indicate different lag lengths, we choose the lesser lag length for the sake of parsimony. The suggested lag length by this procedure is 4. The VAR is therefore estimated with this lag length together with a constant term and uses 5,041 observations. To document return spillovers, we first present χ^2 statistics for the null hypothesis that variable i does not Granger-cause variable j. Specifically, in Table 3 we test whether the lag coefficients of i are jointly 0 when j is the

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

dependent variable in the VAR. The cell associated with the *i*th row variable and the *j*th column variable shows the statistic associated with this test.

TABLE 3

Granger Causality Results

Table 3 presents Granger causality results from a VAR with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9, with the smallest decile being "0" and the largest being "9." The VAR is estimated with 4 lags, includes a constant term, and uses 5,041 observations. Cell *ij* shows χ^2 statistics of pairwise Granger causality tests between the *i*th row variable and the *j*th column variable. The null hypothesis is that all lag coefficients of the *i*th row variable are jointly 0 when *j* is the dependent variable in the VAR. QSPR stands for quoted spread. The stock liquidity series are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. RET is the decile return, and VOL is the return volatility. The sample spans the period January 4, 1988–December 31, 2007. ** and * denote significance at the 1% and 5% levels, respectively.

Endogenous Variables	VOL0	VOL9	RET0	RET9	QSPR0	QSPR9
VOL0		7.876	28.574**	3.631	10.905*	5.541
VOL9	11.529*		8.163	14.449**	13.853**	93.350**
RET0	25.771**	30.547**		2.285	16.179**	1.530
RET9	24.361**	48.144**	36.367**		1.777	39.622**
QSPR0	7.232	21.859**	2.233	2.245		9.137
QSPR9	56.205**	75.922**	15.714**	5.551	40.194**	

Consistent with LM (1990), large firm returns Granger-cause small firm returns. Thus, large firm returns strongly lead small firm returns even after accounting for liquidity dynamics. Further, large firm volatility Granger-causes large and small firm liquidity. Finally, large firm liquidity strongly Granger-causes small firm liquidity. These results establish the role of the large-cap sector as the leader in price discovery as well as liquidity dynamics.

We now estimate the impulse response functions (IRFs) in order to examine the joint dynamics of liquidity, volatility, and returns implied by the full VAR system.[11] Results from the IRFs are generally sensitive to the specific ordering of the endogenous variables.[12] In our base IRFs, we fix the ordering for endogenous variables as follows: VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9. While the rationale for the relative ordering of returns, volatility, and liquidity is ambiguous, we find that the impulse response results are robust to the ordering of these 3 variables.

Figures 2 and 3 illustrate the response of endogenous variables in a particular (large or small) sector to a unit standard deviation orthogonalized shock in the endogenous variables in the other sector for a period of 10 days. Monte Carlo 2-standard-error bands are provided to gauge the statistical significance of the responses.[13] There is clear evidence that shocks to large firm spreads are useful in forecasting small firm spreads but not vice versa.

[11] An IRF traces the impact of a 1-standard-deviation innovation to a specific variable on the current and future values of the chosen endogenous variable. We use the inverse of the Cholesky decomposition of the residual covariance matrix to orthogonalize the impulses.

[12] However, the VAR coefficient estimates (and, hence, the Granger causality tests) are unaffected by the ordering of variables.

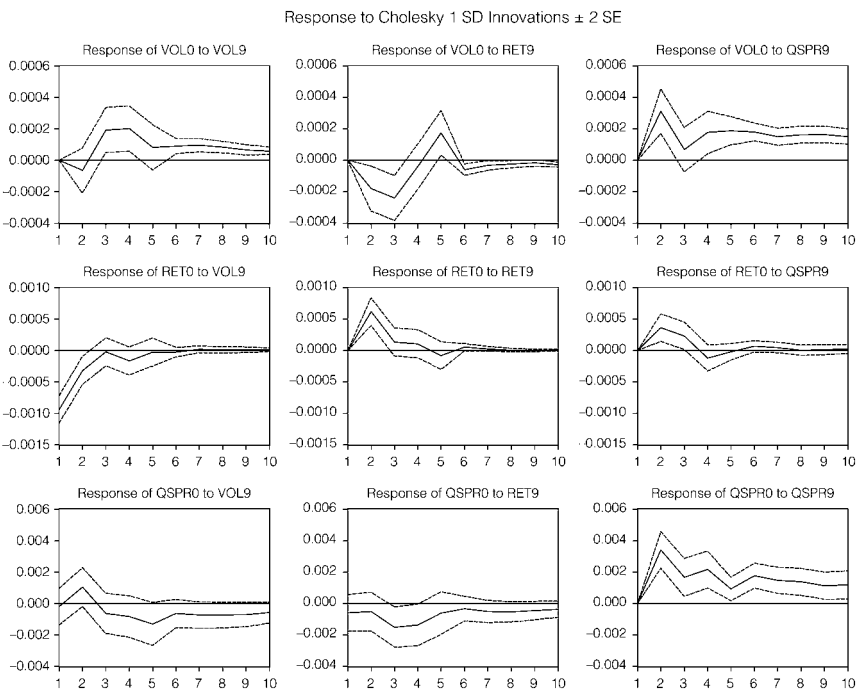
[13] Period 1 in the IRFs represents the contemporaneous response, and the units on the vertical axis are in actual units of the response variable (e.g., dollars in the case of spreads).

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

FIGURE 2

Impulse Response Functions:
Response of Decile 0 (smallest firms) to Decile 9 (largest firms)

Figure 2 presents impulse response functions from the VARs with endogenous variables representing volatility (VOL), returns (RET), and quoted bid-ask spreads (QSPR). The ordering is VOL9, VOL0, RET9, RET0, QSPR9, and QSPR0, with the smallest decile being "0" and the largest being "9." The figure presents impulse responses of decile 0 to decile 9. The 1st period (indicated by "1" on the horizontal axis) is the contemporaneous response.



We also find that return spillovers persist even after accounting for liquidity, in that large stock returns are useful in forecasting small stock returns for a period of 1 day. However, the impulse responses do not show a statistically significant lead from small stocks to large stocks.[14] We also have verified that in the equation for RET9, the coefficient of the lag of RET0 is statistically insignificant. Thus, there is no statistically significant lead from small stocks to large stocks. These results are consistent with our theoretical model. Our model also suggests a specific mechanism that relates cross-autocorrelations to liquidity dynamics. In the next section, we test for this mechanism by examining small-cap return dynamics in more detail; specifically, by adding additional variables to the equation for RET0.[15]

[14]To examine whether these results are robust to the relative ordering of the small and large stocks, we reestimate the IRFs after reversing the VAR ordering as follows: VOL9, VOL0, RET9, RET0, QSPR9, and QSPR0. The results are generally unchanged in the reordering. We also examine the impulse responses of large or decile 9 stocks to other deciles (e.g., decile 5) and find that the results are qualitatively similar to the previously reported responses of large stocks to decile 0 stocks.

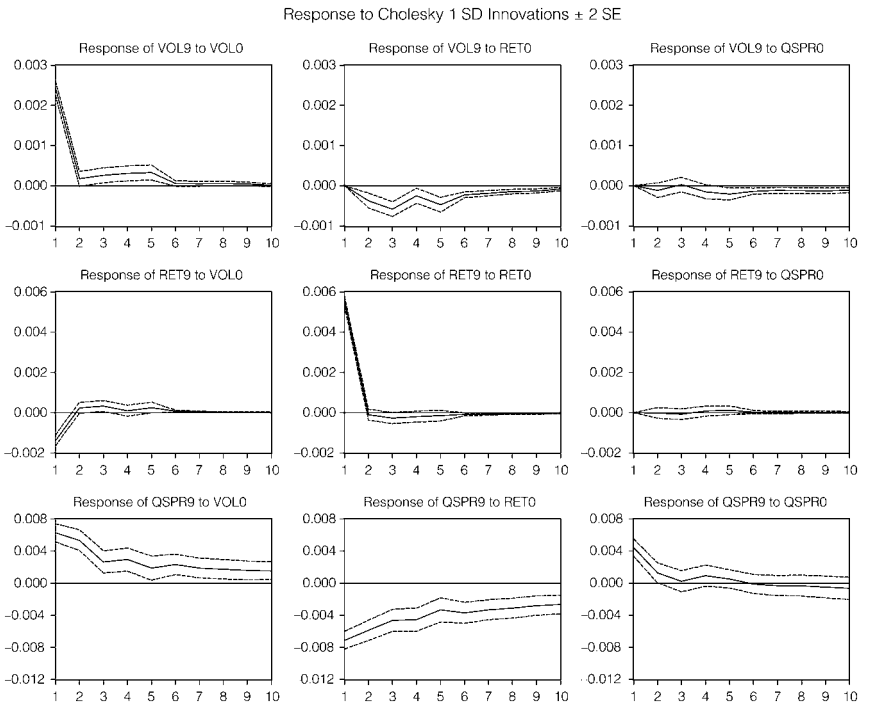
[15]Since there is no lead from small- to large-cap returns, the RET9 equation is not analyzed further.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

FIGURE 3

Impulse Response Functions:
Response of Decile 9 (largest firms) to Decile 0 (smallest firms)

Figure 3 presents impulse response functions from the VARs with endogenous variables representing volatility (VOL), returns (RET), and quoted bid-ask spreads (QSPR). The ordering is VOL9, VOL0, RET9, RET0, QSPR9, and QSPR0, with the smallest decile being "0" and the largest being "9." The figure presents impulse responses of decile 0 to decile 9. The 1st period (indicated by "1" on the horizontal axis) is the contemporaneous response.



VI. Impact of Liquidity on Cross-Autocorrelations

A. Basic Results

The persistence of return cross-autocorrelations documented in the previous section raises the issue of how liquidity interacts with these spillovers. Economic arguments laid out in Section II suggest that we should expect such interactions. We follow BJS (1993) in conducting the lead-lag analysis at the daily frequency.[16] We capture the influence of liquidity levels and order-flow dynamics on the

[16]In an exploratory investigation, we considered a weekly horizon similar to that used by LM (1990). We find that the lead-lag relation from large to small stocks has weakened in recent years. Indeed, a quick check using the CRSP size decile returns indicates that from July 1962 to December 1988 (defining a week as starting Wednesday and ending Tuesday), the correlation between weekly small-cap returns and 1 lag of the weekly large-cap return is as high as 0.210, whereas from 1988 to 2002, this correlation drops to 0.085. This is perhaps not surprising; we would expect technological improvements in trading to contribute to greater market efficiency. Since the baseline lead-lag effect is weak over the weekly horizon within our sample period, we desist from an analysis of weekly returns and liquidity in this paper.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

lead-lag relationship from RET9 to RET0 by adding a number of liquidity-related interaction variables to the equation for RET0. These interaction variables include the first lags of the 4 variables defined by QRETij = QSPRi × RETj, with $i, j = 0$ or 9. If, as discussed previously, information events occur exogenously, the interaction variables can be treated as exogenous to the VAR system. With the addition of these interaction terms, the VAR no longer conforms to the standard form, and so the ordinary least squares (OLS) method is no longer efficient. Thus, we use the seemingly unrelated regression (SUR) method to estimate the system of equations.

Panel A of Table 4 presents the results of these regressions where RET0 is the left-hand variable. The coefficient of lagged large-cap return alone is a statistically significant 0.083. We also note that small-cap returns are positively autocorrelated, consistent with Mech (1993), Boudoukh et al. (1994), and Chordia and Swaminathan (2000). Column 2 interacts the spread in large and small stocks with the lagged large and small firm returns. The coefficient on QRET99 (large firm spread interacted with returns) is strongly significant, suggesting that the lead-lag relation between lagged returns of large stocks and the current returns of small stocks is stronger when the large stocks are illiquid. Thus, the evidence is consistent with the notion that a widening of large stock spreads signals an information event that is transmitted to small stocks with a lag.

We interact OIB with spreads of the large stocks (to construct the variable QOIB99) and include the lag of this interaction variable in the regression. As suggested by our theory, the results, shown in column 6 of Table 4, indicate that large firm order flow interacted with large firm spreads is strongly predictive of small firm returns, whereas the return interaction variable becomes insignificant and its magnitude diminishes in the presence of OIB.[17] We also present the χ^2 statistics and p-values associated with the Wald test for the null hypothesis that the coefficients of all exogenous variables are jointly 0. We reject the null hypothesis that the coefficients of the imbalance interaction term OIB99 and the spread-return interaction terms QRET00, QRET09, QRET90, and QRET99 are jointly 0 at a p-value below 0.001. The overall evidence supports the hypothesis that large firm order flows induced by informational events drive the lead-lag relationship between large and small firms. The results also show that the interaction of the large firm order flow with the large firm spread forms a stronger predictor of the small firm return than the large firm return itself. Thus, trading in large stocks in the presence of low liquidity is what drives the cross-autocorrelation phenomenon.

We also note that there is an intriguing negative coefficient on the variable that interacts small-cap spreads with small-cap returns, QRET00. We know that the small stocks experience positive autocorrelation overall (from the first regression in Panel A of Table 4). The negative coefficient on QRET00 suggests that this positive autocorrelation is attenuated in the presence of frictions when spreads are high.

[17]To disentangle the role of OIB in the interaction of spreads with OIB, we include 1 lag of OIB separately as an exogenous variable in the VAR. This variable is not significant, and the interaction term remains positive and significant.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

TABLE 4

VAR Results with Interaction Terms, for the Smallest Decile

Table 4 presents results from VARs with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9, where $N = 0$ and 9 refer to size deciles. The deciles are numbered in order of increasing size, with the smallest decile being "0" and the largest being "9." In addition, 1 lag of the exogenous variables QRET09, QRET99, QRET00, QRET90, and QOIB99 are included in the equation for RET0, where QRETij = QSPRi × RETj, for i, j = 0, 9, and QOIB99 = QSPR9 × OIB9. The VAR is estimated with 4 lags, include a constant term, and uses 5,041 observations. The seemingly unrelated regression (SUR) method is used to estimate the system of equations. QSPR stands for quoted spread. The stock liquidity series are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. RET is the decile return and VOL is the return volatility. OIB is measured as the dollar value of shares bought minus the dollar value of shares sold, divided by the total dollar value of trades. The sample spans the period January 4, 1988–December 31, 2007. The Wald test reports the χ^2 statistics for the null hypothesis that the coefficients of all exogenous variables are jointly 0. ** and * denote significance at the 1% and 5% levels, respectively.

	Estimate	*t*-Stat.	Estimate	*t*-Stat.	Estimate	*t*-Stat.
Panel A. January 4, 1988–December 31, 2007						
Endogenous Variable: RET0						
RET9(–1)	0.083**	5.856	0.055	1.697	0.030	0.903
RET0(–1)	0.124**	7.078	0.239**	5.900	0.244**	6.024
QRET99(–1)	——	——	0.297**	2.642	0.185	1.621
QRET90(–1)	——	——	–0.108	–0.744	–0.126	–0.866
QRET09(–1)	——	——	–0.135	–0.713	–0.110	–0.585
QRET00(–1)	——	——	–0.628**	–2.820	–0.614**	–2.762
QOIB99(–1)	——	——	——	——	0.050**	4.868
Wald Test						
χ^2	——	——	21.454		45.283	
Probability	——	——	0.000		0.000	
Panel B. January 29, 2001–December 31, 2007						
Endogenous Variable: RET0						
RET9(–1)	0.083**	5.881	0.019	0.334	0.012	0.217
RET0(–1)	–0.053**	–3.009	0.258**	3.579	0.258**	3.585
QRET99(–1)	——	——	0.395*	2.228	0.388*	2.186
QRET90(–1)	——	——	0.000	0.001	–0.020	–0.079
QRET09(–1)	——	——	–0.720*	–2.157	–0.727*	–2.177
QRET00(–1)	——	——	–0.802*	–1.965	–0.774	–1.891
QOIB99(–1)	——	——	——	——	0.013	1.003
Wald Test						
χ^2	——	——	31.052		32.115	
Probability	——	——	0.000		0.000	

Motivated by the evidence (Chordia et al. (2008)) that the market has become more efficient in recent years, particularly after decimalization, in Panel B of Table 4 we report results for the decimal period. In this period, there is still a lead-lag effect, and the coefficient of QRET99 is significant but slightly less so than in Panel A. However, the point estimate of the coefficient is actually higher than its corresponding value in Panel A, suggesting that a power problem induced by the decrease in the number of observations influences the decrease in significance of QRET99. We also note that QOIB99 in this case is not significant. Overall, the results are indicative of greater efficiency in the post-decimal period, which is consistent with intuition.

We next report results for all other deciles in Table 5 (columns 2 and 3). We use the same interaction variables as previously, except that we replace the small firm (decile 0) variables with the variables corresponding to deciles 1–8. We make a similar replacement for the OIB variable. The table shows that the large firm order flow variable interacted with large firm spreads is informative in predicting returns in every size decile, even though the coefficient magnitudes are generally smaller for the larger deciles.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

TABLE 5

VAR Results with Interaction Terms, for all deciles

Table 5 presents results from VARs with endogenous variables VOLN, VOL9, RETN, RET9, QSPRN, and QSPR9, where $N = 1$ through 8 refers to size deciles. RET denotes the decile return, VOL the return volatility, and QSPR the quoted spread. The deciles are numbered in order of increasing size, with the smallest decile being "0" and the largest being "9." In addition, 1 lag of the exogenous variables QRETN9, QRET99, QRETNN, QRET9N, and QOIB99 are included in the equation for RETN, where $N = 1$ through 8, and QRETij = QSPRi × RETj, for i, $j = N$, 9, and QOIB99 = QSPR9 × OIB9. OIB is the order imbalance, measured as the dollar value of shares bought minus the dollar value of shares sold, divided by the total dollar value of trades. For columns 2 and 3, RET is measured based on the quote midpoint corresponding to the last trade of the day; for columns 4 and 5, RET is measured using the closing quote midpoints on each day; for columns 6 and 7, RET is the midpoint return adjusted for nonsynchronous trading by the method of Jokivuolle (1995). All VARs are estimated with 4 lags and include a constant term. The seemingly unrelated regression (SUR) method is used to estimate the system of equations. The stock liquidity series are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. ** and * denote significance at the 1% and 5% levels, respectively. The sample spans the period January 4, 1988–December 31, 2007 (5,041 observations).

	Coefficient of QOIB99(−1) for Midpoint Returns		Coefficient of QOIB99(−1) for Adjusted Midpoint Returns		Coefficient of QOIB99(−1) for Nonsynchronized Returns	
Size Decile	Estimate	*t*-Stat.	Estimate	*t*-Stat.	Estimate	*t*-Stat.
0	0.050**	4.868	0.035**	3.501	0.039*	2.143
1	0.050**	4.857	0.041**	4.327	0.035**	4.832
2	0.055**	5.574	0.050**	5.372	0.042**	5.545
3	0.052**	5.458	0.042**	4.779	0.039**	5.416
4	0.052**	5.747	0.041**	4.984	0.041**	5.772
5	0.039**	4.776	0.030**	4.104	0.032**	4.833
6	0.034**	4.693	0.027**	4.156	0.029**	4.736
7	0.025**	3.652	0.017**	2.833	0.023**	3.762
8	0.023**	3.962	0.011*	2.359	0.021**	4.031

The rationale for the lead-lag patterns is based on the notion that transactions in response to informational events occur first in large stocks and then spill over to small stocks, partially in the form of lagged transactions in the small stocks and in the form of lagged quote updates by small stock market makers. Our return computations are based on midquote prices matched to transactions and account for transaction-induced lags. However, small stocks often do not trade for several hours within a day. Thus, there still may be a residual effect of nonsynchronous trading.

To address this issue, we compute midquote returns using the last available quote for each firm on a given day. The results appear in columns 4 and 5 of Table 5. As can be seen, the coefficients of the imbalance interaction variables are positive in every case and significant at the 5% level in all cases. Thus, our transaction price-based results on predictability extend to midquote returns as well, and our earlier results continue to hold for the post-1993 sample.

We perform a further check to address any possible influence of asynchronous trading activity by adopting the approach of Jokivuolle (1995), who devises a method for estimating the true value of a stock index level when some of its constituents are subject to stale pricing. He shows that the log of the true index value can be represented as the permanent component of a Beveridge and Nelson (BN) (1981) decomposition of the series of log index first differences. We implement the Jokivuolle method as follows: First, we compute the observed index level for each of the midquote return series by normalizing the index at the beginning of the sample period to 100 and then updating it based on observed returns. We then

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

extract the BN permanent component for each these index series.[18] Finally, we use returns implied by these BN-adjusted indices in place of the original midquote returns in the VAR. Results obtained from using these adjusted returns appear in the last 2 columns of Table 5. As can be seen, the significance of the coefficients remains largely unchanged, and the coefficient magnitudes increase in several cases, underscoring the robustness of our results.[19]

The last midquote return results shed additional light on the economic causes of the lead-lag effect. Specifically, one possible interpretation of Table 4 is that secular decreases in liquidity can reduce trading activity in small stocks, and this reduction can affect leads and lags. Our results point to the notion that this effect is not the driver of lead-lag between the large and small firms. To see this, observe that the last midquote series only captures the quote updating activity of market makers. The frequency of *quote updating* is not likely to be affected directly by liquidity, because specialists can continuously update quotes even in the absence of trading. Thus, our results are consistent with the view that market makers' opportunity costs of continuously monitoring order flow in other markets play a pivotal role in the lead-lag relationship across small and large stocks.

Our results are new and differ from those in the literature. For instance, Mech (1993) tests the hypothesis that the lead from large to small stock returns is greater when the small firm spread is high relative to the profit potential (proxied by the absolute return). He does not find support for this hypothesis. From a conceptual standpoint, in contrast to Mech, we do not view the spread as an inverse measure of profit potential but as an indicator that private information traders are active in large stocks. Moreover, unlike Mech, we consider the role of large firm spreads in addition to small stock spreads in determining the extent of the lead-lag relationship, and we find that large firm spreads are most relevant to the lead-lag effect.

The economic significance of our results does not suggest an overwhelming violation of market efficiency.[20] Based on the relevant QOIB coefficient (0.050) in Panel A of Table 4, and the sample standard deviation of QOIB of 0.0123, we find that a 1-standard-deviation increase in the interaction variable results in a daily small firm return increment of about 0.06%. Trading frictions including market impact costs and brokerage commissions are likely to render these magnitudes unprofitable.

B. Common Information and Cross-Autocorrelations

In this section we study the lead-lag relation around macroeconomic announcements. Brandt and Kavajecz (2004), Beber et al. (2008), and Evans and

[18]The extended sample autocorrelation function method (i.e., the extended sample autocorrelation function (ESACF) option in SAS) is used to pick the order of the autoregressive moving average (ARMA) process for each series.

[19]Using the BN (1981) decomposition on the transaction price series also does not alter the significance of the QOIB coefficients. Since the results are unaltered by using the Jokivuolle (1995) adjustment, we use the unadjusted series in the remainder of the paper.

[20]More generally, Bessembinder and Chan (1998) find that profits from a broad set of technical rules are not overwhelming after accounting for transaction costs.

Lyons (2002), (2008) indicate that order flows in financial markets contain information about the macroeconomy. Thus, we would expect trading based on systematic information to be stronger around macroeconomic news events. We study 2 major announcements, viz., those for nonfarm employment and the CPI. We do not use the gross domestic product (GDP) announcements because they are at a quarterly frequency, while the employment and inflation announcements are available at a monthly frequency. Moreover, Fleming and Remolona (1999) suggest that among all major macroeconomic news releases, these announcements have the strongest impact on financial market liquidity and price formation. Consequently, we analyze how our previous results are affected around these announcements.

In our sample, the announcements all occurred prior to the commencement of trading on the NYSE on the corresponding day, indicating that the day of the announcement corresponded to a period after the announcement. To analyze how our results change prior to the announcement, we thus define an indicator variable MACRO to take a value of 1 on the day just prior to the day of the announcements.[21] We interact MACRO and (1 − MACRO) with the variables in Table 4 (i.e., lagged values of QRET99 and QOIB99). We then include these interactions as explanatory variables in place of the original ones in the VAR. As usual, we report only the results for the equation with the small stock returns as the dependent variable.

The results in Table 6 show that the coefficient on the interaction term of MACRO and QOIB99 is 0.069 and that on the interaction term of (1 − MACRO) and QOIB99 is 0.046, suggesting that the impact of the interaction variable is larger during the preannouncement period. Hence, consistent with our conjecture, large firm order flow has greater impact on the lead-lag relation before announcements compared to the remaining sample. It is worth noting, however, that while the difference in the 2 QOIB coefficients is statistically significant, the size of the difference is about half that of the baseline effect of the QOIB variable.

We would expect information asymmetry to be temporarily lower following public announcements that resolve information uncertainty.[22] If leads and lags are caused by private information flows (as suggested by our theory), we would expect them to be weaker after public announcements than on other days. To address this, we alternatively define the dummy variable, MACRO, to take on the value of unity on the day of the announcement. In this case (column 4 of Table 6), we find that the coefficient on lagged QOIB99 when interacted with MACRO is an insignificant 0.033 and when interacted with (1 − MACRO) it is 0.054, and the latter is significant. Moreover the difference between the two is statistically significant. This suggests that the impact of large firm liquidity and that trading on small firm returns is smaller on the day of the announcements than during other days of the sample period. Overall, the results support the hypothesis that trading

[21] In one case, the employment and CPI announcements were close together, and the "before" day of one coincided with the "after" day of the other. To prevent ambiguity, we delete this observation.

[22] Pasquariello and Vega (2007) show that public macroeconomic announcements lower adverse selection and thus increase liquidity.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

TABLE 6

VAR Results with Interaction Terms, for the Smallest Decile, with Interaction Terms for Days around Macroeconomic Announcements

Table 6 presents results from VARs with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9, where $N = 0$ and 9 refer to size deciles. The deciles are numbered in order of increasing size, with the smallest decile being "0" and the largest being "9." In addition, 1 lag of the exogenous variables QRET09, QRET99, QRET90, QRET00, and QOIB99 is included in the equation for RET0, where QRETij = QSPRi × RETj, for i, j = 0, 9, and QOIB99 = QSPR9 × OIB9. QOIB99 is further interacted with the dummy variables MACRO and (1 – MACRO), where MACRO indicates the day before or of announcements of the core Consumer Price Index (CPI), the Producer Price Index (PPI) and nonfarm employment. The VAR is estimated with 4 lags, includes a constant term, and uses 5,041 observations. The seemingly unrelated regression (SUR) method is used to estimate the system of equations. QSPR stands for quoted spread. The stock liquidity series are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. RET is the decile return, and VOL is the return volatility. OIB is measured as the dollar value of shares bought minus the dollar value of shares sold, divided by the total dollar value of trades. The sample spans the period from January 4, 1988 to December 31, 2007. The Wald test reports the χ^2 statistics for the following null hypotheses: i) the coefficients of QRET99(–1) × MACRO and QRET99(–1) × (1 – MACRO) are equal; or ii) the coefficients of QOIB99(–1) × MACRO and QOIB99(–1) × (1 – MACRO) are equal. ** and * denote significance at the 1% and 5% levels, respectively.

	MACRO = 1			
	Day before Announcement		Day of Announcement	
	Estimate	*t*-Stat.	Estimate	*t*-Stat.
Endogenous Variable: RET0				
RET9(–1)	0.026	0.795	0.030	0.928
QRET99(–1) × MACRO (D1)	0.401*	1.974	–0.025	–0.126
QRET99(–1) × (1 – MACRO) (D2)	0.183	1.594	0.193	1.681
QOIB99(–1) × MACRO (D3)	0.069**	3.534	0.033	1.620
QOIB99(–1) × (1 – MACRO) (D4)	0.046**	4.307	0.054**	5.027
Wald Test				
χ^2 (D1 = D2)	4.444		4.012	
Probability	0.108		0.135	
χ^2 (D3 = D4)	24.872		25.386	
Probability	0.000		0.000	

in response to common information shocks plays a material role in the lead-lag effect.

VII. NASDAQ Stocks: A Robustness Check

As a robustness check on our basic results, we now consider spillovers between the liquidity of NYSE stocks and a holdout sample of the relatively smaller NASDAQ stocks.[23] Our analysis also allows us to consider the potential effects of the different market structures across NYSE and NASDAQ on liquidity spillovers.

We use daily NASDAQ closing bid and ask prices on the CRSP database in order to compute daily bid-ask spreads for NASDAQ. The NASDAQ spread index is constructed by using the value-weighted average of the spread in a manner similar to that used for the NYSE indices described in Section III; return and volatility measures are also constructed in the corresponding manner. Due to the potentially more severe problem of stale prices among the relatively smaller NASDAQ stocks, however, we report results using quote midpoint return series to compute returns and volatility.

[23]The average market capitalization of NASDAQ stocks is $0.93 billion, which is comparable to the average market capitalization of stocks in NYSE decile 5 (about $0.94 billion).

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

As before, we adjust the NASDAQ series of spreads, returns, and volatility to account for regularities.[24] We estimate a VAR for which the endogenous variables are returns, volatilities, and quoted spreads for NASDAQ stocks and NYSE decile 9 stocks. The VAR includes 4 lags and a constant term. In Table 7, we present the analog of the lead-lag regression in Table 5. Specifically, we examine the response of NASDAQ stock returns to the interaction of decile 9 order flow with decile 9 spreads. As before, the interaction terms are treated as exogenous variables in the VAR. Again, the large firm OIB interacted with large firm spreads is significant, and its magnitude is about 30% higher than the corresponding coefficient in Table 4. The lagged own NASDAQ return is also positive and significant, like in the case of small-cap returns in Table 4. Overall, these results confirm the role of large firm order flows and liquidity in predicting returns in both small NYSE firms as well as NASDAQ firms. Thus, our key results continue to obtain for NASDAQ stocks.

TABLE 7

VARs Results with Interaction Terms, including NASDAQ Stocks

Table 7 presents the results from a VAR with endogenous variables VOLN, VOL9, RETN, RET9, QSPRN, and QSPR9 with the smallest NYSE decile being "0" and the largest being "9"; the subscript "*N*" represents NASDAQ stocks. The VAR uses 5,025 observations, has 4 lags, and includes a constant term. QSPR stands for quoted spread. RET is the decile return, VOL is the return volatility, and OIB is the buy dollar volume less sell dollar volume normalized by the total dollar volume. One lag of the exogenous variables QRETN9, QRET99, QRETNN, QRET9N, and QOIB99 is included in the equation for RETN, and QRETij = QSPRi × RETj, for $i, j = N$, 9, and QOIB99 = QSPR9 × OIB9. OIB is the order imbalance, measured as the dollar value of shares bought minus the dollar value of shares sold, divided by the total dollar value of trades. The seemingly unrelated regression (SUR) method is used to estimate the system of equations. The sample spans the period from January 4, 1988 to December 31, 2007. ** and * denote significance at the 1% and 5% levels, respectively.

	Estimate	*t*-Stat.	Estimate	*t*-Stat.	Estimate	*t*-Stat.
Endogenous Variable: RETN						
RET9(–1)	–0.020	–0.638	–0.032	–0.626	–0.054	–1.047
RETN(–1)	0.110**	5.363	0.131**	3.892	0.127**	3.764
QRET99(–1)	——	——	0.106	0.448	–0.048	–0.200
QRET9N(–1)	——	——	–0.060	–0.339	–0.054	–0.307
QRETN9(–1)			–0.007	–0.101	–0.010	–0.147
QRETNN(–1)			–0.048	–1.153	–0.041	–0.995
QOIB99(–1)	——	——	——	——	0.065**	3.733
Wald Test						
χ^2	——	——	2.619		16.576	
Probability	——	——	0.624		0.005	

VIII. Industry-Specific Information

Our analysis indicates that the lead-lag relation is driven by private informational events that are traded upon first in the large stocks and then spill over to the small stocks. However, up to this point we have only conducted analyses at the overall market level. As the analyses of Barclay, Litzenberger, and Warner

[24]For NASDAQ stocks, the dummy variable for the change to sixteenths equals 1 for the period from June 12, 1997 to March 11, 2001. Further, there are 3 decimalization dummy variables to reflect the gradual introduction of decimalization for various subsets of stocks over the following periods: March 12, 2001 to March 25, 2001; March 26, 2001 to April 8, 2001; and April 9, 2001 to December 31, 2007 (see, e.g., Chung and Chuwonganant (2004)).

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

(BLW) (1990) and Hou (2007) suggest, at least some information-based trading may also occur at the industry level. Motivated by this observation, we now consider whether our results hold within industries. To keep our sample size per industry large enough in order to draw reliable inferences, we use the 5 broad industry classifications proposed by Kenneth French on his Web site.[25]

We explore the lead-lag relation within industries by first categorizing all NYSE-listed stocks into 5 groups by the French industry classifications. The sample is then sorted into 2 equally sized subsamples by market capitalization as of the end of the previous month.[26] We then calculate the value-weighted daily midquote returns, volatility, OIB, and adjusted quoted spreads for the 2 size groups within each of these industry groups, just as in Tables 4 and 5. Finally, we run VARs similar to those in Table 4 for each of the 5 industry classifications. In order to control for marketwide effects, and for the notion that stocks may have differential adjustment speeds to marketwide information, we include the contemporaneous, 1 lead, and 1 lag of the NYSE Composite Index return in the equations for large/small firm returns.

The results, by industry group, appear in Table 8 and are reported in the same format as Table 4 (the coefficients of the market return and lagged small-cap return are suppressed for brevity). We find that for each industry, there is a lead-lag relation between small and large firms. As before, this lead-lag relation is at least partially accounted for by the interaction of large firm order flows with large firm spreads. Specifically, these interaction terms are strongly significant for each of the 5 industry groups. Further, inclusion of the interaction variable accounts for the significance of the lagged large firm return as an explanatory variable for the small firm return.

To further ensure that our results are not being driven by a marketwide effect, in addition to controlling for market returns we also include the lagged interaction of the quoted spread with the order flow of the largest decile of all NYSE-listed stocks (the variable of focus in the last section). Results from the inclusion of this variable appear in the last 2 columns of Table 8. The marketwide interaction variable reduces the magnitude of the industry coefficients, but they remain positive in all 5 cases, and significant in 4 of 5 cases. Thus, there remains evidence of industrywide information trading in a manner consistent with our model even after controlling for overall market effects.

Our findings are broadly consistent with those of BLW (1990). BLW use an interesting natural experiment in the context of the Tokyo Stock Exchange. During their sample period, the exchange remained open on Saturdays for only 3 Saturdays a month. Thus, the ratios of weekend variances when the exchange is open and, in turn, closed on Saturdays form a metric for private information revealed through trading. Table 3 of BLW indicates that these ratios are largest for manufacturing firms and smallest for financial and communication firms. The point estimates of the interaction variable in our Table 8 are largest for the manufacturing

[25]The specific URL is http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/siccodes5.zip.

[26]Note that we do not form size deciles in this section but split the sample evenly within each industry. This is to ensure large enough samples of both small and large stocks within each of the 5 industry groups.

TABLE 8

VAR Results with Interaction Terms, by Industry, for NYSE Stocks

Table 8 presents results from VARs with endogenous variables VOL0, VOL1, RET0, RET1, QSPR0, and QSPR1, where $N = 0$ and 1 refer to 2 size groups. The smaller firm group is labeled "0" and the larger firm group "1." The variable IND refers to the 5 industry groups from Ken French's Web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/siccodes5.zip). One lag of the exogenous variables QRET01, QRET11, and QOIB11 is included in the equation for RET0, where QRET01 = QSPR0 × RET1, QRET11 = QSPR1 × RET1, and QOIB11 = QSPR1 × OIB1. In order to control for marketwide effects, we include the contemporaneous, 1 lead, and 1 lag of the NYSE Composite index return in the equations for large-/small-cap returns. QOIB99 = QSPR9 × OIB9, where QSPR9 and OIB9 are the value-weighted order flows and quoted spreads for the largest firm decile of NYSE-listed stocks. The VAR is estimated with 4 lags, includes a constant term, and uses 5,041 observations. The seemingly unrelated regression (SUR) method is used to estimate the system of equations. QSPR stands for quoted spread. The stock liquidity series are constructed by first averaging all transactions for each individual stock on a given trading day and then constructing value-weighted averages for all individual stock daily means that satisfy the data filters described in the text. RET is the decile return and VOL is the return volatility. OIB is measured as the dollar value of shares bought minus the dollar value of shares sold, divided by the total dollar value of trades. The sample spans the period from January 4, 1988 to December 31, 2007. The Wald test reports the χ^2 statistics for the null hypothesis that the coefficients of all exogenous variables are jointly 0. ** and * denote significance at the 1% and 5% levels, respectively.

	Estimate	*t*-Stat.	Estimate	*t*-Stat.	Estimate	*t*-Stat.	Estimate	*t*-Stat.
Panel A. IND = 1 (consumer durables, nondurables, wholesale, retail, and some services)								
Endogenous Variable: RET0								
RET1(–1)	0.077**	4.115	0.024	0.802	0.000	–0.004	0.003	0.107
QRET11(–1)	——	——	0.489**	3.355	0.196	1.298	0.184	1.220
QOIB11(–1)	——	——	——	——	0.074**	7.172	0.088**	6.353
QOIB99(–1)	——	——	——	——	——	——	–0.018	–1.454
Wald Test								
χ^2	——		16.072		67.779		69.960	
Probability	——		0.003		0.000		0.000	
Panel B. IND = 2 (manufacturing, energy, and utilities)								
Endogenous Variable: RET0								
RET1(–1)	0.080**	3.551	0.013	0.362	–0.021	–0.613	–0.022	–0.640
QRET11(–1)	——	——	0.208	1.393	0.007	0.044	0.002	0.013
QOIB11(–1)	——	——	——	——	0.077**	8.178	0.067**	6.007
QOIB99(–1)	——	——	——	——	——	——	0.014	1.592
Wald Test								
χ^2	——		16.682		83.820		86.160	
Probability	——		0.002		0.000		0.000	
Panel C. IND = 3 (business equipment, telephone and television transmission)								
Endogenous Variable: RET0								
RET1(–1)	0.052**	2.801	0.015	0.430	0.006	0.187	0.006	0.185
QRET11(–1)	——	——	0.232	1.796	0.076	0.563	0.070	0.521
QOIB11(–1)	——	——	——	——	0.045**	4.022	0.032*	2.410
QOIB99(–1)	——	——	——	——	——	——	0.027	1.930
Wald Test								
χ^2	——		4.907		21.117		24.877	
Probability	——		0.297		0.001		0.000	
Panel D. IND = 4 (healthcare, medical equipment, and drugs)								
Endogenous Variable: RET0								
RET1(–1)	0.065**	4.174	–0.022	–0.717	–0.033	–1.063	–0.034	–1.082
QRET11(–1)	——	——	0.276*	2.225	0.215	1.710	0.223	1.769
QOIB11(–1)	——	——	——	——	0.034**	2.851	0.022	1.640
QOIB99(–1)	——	——	——	——	——	——	0.026	1.787
Wald Test								
χ^2	——		21.204		29.323		32.524	
Probability	——		0.000		0.000		0.000	
Panel E. IND = 5 (other: mines, construction, building materials, transportation, hotels, business services, entertainment)								
Endogenous Variable: RET0								
RET1(–1)	0.078**	4.450	0.041	1.492	0.018	0.664	0.017	0.613
QRET11(–1)	——	——	0.355**	2.955	0.206	1.699	0.206	1.696
QOIB11(–1)	——	——	——	——	0.064**	7.209	0.059**	5.343
QOIB99(–1)	——	——	——	——	——	——	0.009	0.901
Wald Test								
χ^2	——		16.756		68.885		69.662	
Probability	——		0.002		0.000		0.000	

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

and consumer durables groups and lowest for the health care and communication groups, which broadly accords with BLW.

Overall, the results of this section, as those in the previous section, provide considerable support to our model. The findings also suggest that there is private information trading at the industry level. Finally, the analysis proposes a new predictive variable for small firm industry returns, namely, the (lagged) interaction of large firm order flow with large firm spreads.

IX. Concluding Remarks

Our main goal in this paper is to identify the precise mechanism by which the incorporation of information into prices leads to the cross-autocorrelation patterns in stock returns. We develop a model to examine the linkage between liquidity and cross-autocorrelations, and test its predictions. In equilibrium, cross-autocorrelations arise because systematic information-based trading impacts the relatively liquid, large firms first. This trading causes large stock spreads to widen, and the information is subsequently incorporated with a lag into the prices of small stocks. Our framework therefore predicts stronger leads and lags when large firm spreads are high.

Empirically, we find that return cross-autocorrelations survive even after accounting for liquidity and volatility dynamics. Our results also show that large firm returns and order flows predict small firm returns more strongly when large firm spreads are high. This result holds for both transaction returns as well as the end-of-day midquote returns, demonstrating that the finding is not due to stale prices. The role of liquidity in cross-autocorrelations intensifies just prior to major macroeconomic announcements (when informed trading is likely to be intense) and weakens immediately following such announcements (when information asymmetries are lower). The results hold at both marketwide and industry levels. Overall, the data analysis supports our economic hypothesis.

From the standpoint of asset pricing, our ancillary results indicate that the liquidity of a stock is not an exogenous attribute, but its dynamics are sensitive to movements in financial market variables, such as returns and volatility, in both its own and other markets. We further show that return dynamics, in turn, are sensitive to liquidity. Developing a general equilibrium model that prices liquidity while recognizing these endogeneities is a challenging task but is worthy of future investigation. In particular, one interesting issue is to tease out the direct impact of volatility on expected returns through the traditional risk-reward channel as well as to understand its indirect impact by way of its effect on liquidity.

Appendix. Adjustment Process for Liquidity, Returns, and Volatility

In this Appendix, we provide details about the adjustment process for the different time series. Specifically, we regress the series on a set of adjustment variables:

$$w = x'\beta + u \qquad \text{(mean equation)}. \tag{A-1}$$

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

In equation (A-1), w is the series to be adjusted, and x contains the adjustment variables. The residuals are used to construct the following variance equation:

$$\log(u^2) = x'\gamma + v \quad \text{(variance equation)}. \tag{A-2}$$

The variance equation is used to standardize the residuals from the mean equation, and the adjusted w is calculated in the equation

$$w_{\text{ADJ}} = a + b(\hat{u}/\exp(x'\gamma/2)), \tag{A-3}$$

where a and b are chosen so the sample means and variances of the adjusted and the unadjusted series are the same.

The adjustment variables used are as follows: First, to account for calendar regularities in liquidity, returns, and volatility, we use i) 4 dummy variables for Monday–Thursday, ii) 11 month-of-the-year dummy variables for February–December, and iii) a dummy variable for nonweekend holidays set such that if a holiday falls on a Friday then the preceding Thursday is set to 1, if the holiday is on a Monday then the following Tuesday is set to 1, and if the holiday is on any other weekday then the day preceding and following the holiday is set to 1. This captures the fact that trading activity declines substantially around holidays. We also include iv) a time trend and the square of the time trend to remove deterministic trends that we do not seek to explain.

We further consider dummy variables for financial market events that could affect the liquidity of both small and large stocks. Specifically, we include v) 4 crisis dummy variables, where the crises are the bond market crisis (March 1–May 31, 1994), the Asian financial crisis (July 2–December 31, 1997), the Russian default crisis (July 6–December 31, 1998),[27] and the recent credit crunch, encompassing August 1–December 31, 2007; vi) a dummy variable for the period between the shift to sixteenths and the shift to decimals, and another for the period after the shift to decimals; vii) a dummy variable for the week after September 11, 2001, when we expect liquidity to be unusually low; and viii) a dummy variable for September 16, 1991 where, for some reason (most likely a recording error) only 248 firms were recorded as having been traded on the ISSM data set, whereas the number of NYSE-listed firms trading on a typical day in the sample is over 1,100.

References

Admati, A. R., and P. Pfleiderer. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies*, 1 (1988), 3–40.

Avramov, D.; T. Chordia; and A. Goyal. "Liquidity and Autocorrelations in Individual Stock Returns." *Journal of Finance*, 61 (2006), 2365–2394.

Badrinath, S. G.; J. R. Kale; and T. H. Noe. "Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns." *Review of Financial Studies*, 8 (1995), 401–430.

Ball, C. A., and T. Chordia. "True Spreads and Equilibrium Prices." *Journal of Finance*, 56 (2001), 1801–1835.

Barberis, N., and A. Shleifer. "Style Investing." *Journal of Financial Economics*, 68 (2003), 161–199.

Barclay, M. J.; R. H. Litzenberger; and J. B. Warner. "Private Information, Trading Volume, and Stock-Return Variances." *Review of Financial Studies*, 3 (1990), 233–253.

Beber, A.; M. W. Brandt; and K. A. Kavajecz. "What Does Equity Sector Orderflow Tell Us about the Economy?" Working Paper, Duke University (2008).

Benston, G. J., and R. L. Hagerman. "Determinants of Bid-Asked Spreads in the Over-the-Counter Market." *Journal of Financial Economics*, 1 (1974), 353–364.

Bernhardt, D., and R. S. Mahani. "Asymmetric Information and Stock Return Cross-Autocorrelations." *Economics Letters*, 96 (2007), 14–22.

[27]The dates for the bond market crisis are from Borio and McCauley (1996). The starting date for the Asian crisis is the day that the Thai baht was devalued; dates for the Russian default crisis are from the Bank for International Settlements (see "A Review of Financial Market Events in Autumn 1998," CGFS Reports No. 12 (Oct. 1999), available at http://www.bis.org/publ/cgfspubl.htm).

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

Bessembinder, H., and K. Chan. "Market Efficiency and the Returns to Technical Analysis." *Financial Management*, 27 (1998), 5–17.
Beveridge, S., and C. R. Nelson. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'." *Journal of Monetary Economics*, 7 (1981), 151–174.
Borio, C. E. V., and R. N. McCauley. "The Economics of Recent Bond Yield Volatility." BIS Economic Paper No. 45, Bank for International Settlements, Basel, Switzerland (1996).
Boudoukh, J.; M. P. Richardson; and R. F. Whitelaw. "A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Stock Returns." *Review of Financial Studies*, 7 (1994), 539–573.
Brandt, M. W., and K. A. Kavajecz. "Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve." *Journal of Finance*, 59 (2004), 2623–2654.
Brennan, M. J.; N. Jegadeesh; and B. Swaminathan. "Investment Analysis and the Adjustment of Stock Prices to Common Information." *Review of Financial Studies*, 6 (1993), 799–824.
Chan, K. "Imperfect Information and Cross-Autocorrelation among Stock Prices." *Journal of Finance*, 48 (1993), 1211–1230.
Chan, K., and W.-M. Fong. "Trade Size, Order Imbalance, and the Volatility-Volume Relation." *Journal of Financial Economics*, 57 (2000), 247–273.
Chordia, T.; R. Roll; and A. Subrahmanyam. "Market Liquidity and Trading Activity." *Journal of Finance*, 56 (2001), 501–530.
Chordia, T.; R. Roll; and A. Subrahmanyam. "Liquidity and Market Efficiency." *Journal of Financial Economics*, 87 (2008), 249–268.
Chordia, T.; A. Sarkar; and A. Subrahmanyam. "An Empirical Analysis of Stock and Bond Market Liquidity." *Review of Financial Studies*, 18 (2005), 85–129.
Chordia, T.; L. Shivakumar; and A. Subrahmanyam. "Liquidity Dynamics across Small and Large Firms." *Economic Notes*, 33 (2004), 111–143.
Chordia, T., and B. Swaminathan. "Trading Volume and Cross-Autocorrelations in Stock Returns." *Journal of Finance*, 55 (2000), 913–935.
Chowdhry, B., and V. Nanda. "Multimarket Trading and Market Liquidity." *Review of Financial Studies*, 4 (1991), 483–511.
Chung K. H., and C. Chuwonganant. "Tick Size, Order Handling Rules, and Trading Costs." *Financial Management*, 33 (2004), 47–62.
Clark, R. A.; J. J. McConnell; and M. Singh. "Seasonalities in NYSE Bid-Ask Spreads and Stock Returns in January." *Journal of Finance*, 47 (1992), 1999–2014.
Connolly, R. A, and C. T. Stivers. "Time-Varying Lead-Lag of Equity Returns in a World of Incomplete Information." Working Paper, University of North Carolina at Chapel Hill (1997).
Conrad, J., and G. Kaul. "An Anatomy of Trading Strategies." *Review of Financial Studies*, 11 (1998), 489–519.
Evans, M. D. D., and R. K. Lyons. "Order Flow and Exchange Rate Dynamics." *Journal of Political Economy*, 110 (2002), 170–180.
Evans, M. D. D., and R. K. Lyons. "How Is Macro News Transmitted to Exchange Rates?" *Journal of Financial Economics*, 88 (2008), 26–50.
Fleming, M. J., and E. M. Remolona. "Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information." *Journal of Finance*, 54 (1999), 1901–1915.
Gallant, A. R.; P. E. Rossi; and G. Tauchen. "Stock Prices and Volume." *Review of Financial Studies*, 5 (1992), 199–242.
Gompers, P. A., and A. Metrick. "Institutional Investors and Equity Prices." *Quarterly Journal of Economics*, 116 (2001), 229–259.
Gorton, G. B., and G. G. Pennacchi. "Security Baskets and Index-Linked Securities." *Journal of Business*, 66 (1993), 1–27.
Hameed, A. "Time-Varying Factors and Cross-Autocorrelations in Short-Horizon Stock Returns." *Journal of Financial Research*, 20 (1997), 435–458.
Hou, K. "Industry Information Diffusion and the Lead-Lag Effect in Stock Returns." *Review of Financial Studies*, 20 (2007), 1113–1138.
Hou, K., and T. J. Moskowitz. "Market Frictions, Price Delay, and the Cross-Section of Expected Returns." *Review of Financial Studies*, 18 (2005), 981–1020.
Jokivuolle, E. "Measuring True Stock Index Value in the Presence of Infrequent Trading." *Journal of Financial and Quantitative Analysis*, 30 (1995), 455–464.
Jones, C. M.; G. Kaul; and M. L. Lipson. "Transactions, Volume, and Volatility." *Review of Financial Studies*, 7 (1994), 631–651.
Kumar, P., and D. J. Seppi. "Information and Index Arbitrage." *Journal of Business*, 67 (1994), 481–509.
Kyle, A. S. "Continuous Auctions and Insider Trading. *Econometrica*, 53 (1985), 1315–1335.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press

Lee, C. M. C., and M. J. Ready. "Inferring Trade Direction from Intraday Data." *Journal of Finance*, 46 (1991), 733–747.
Lo, A. W., and A. C. MacKinlay. "When Are Contrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies*, 3 (1990), 175–205.
Madhavan, A.; K. Ming; V. Straser; and Y. Wang. "How Effective Are Effective Spreads? An Evaluation of Trade Side Classification Algorithms." Working Paper, ITG, Inc. (2002).
McQueen, G.; M. Pinegar; and S. Thorley. "Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns." *Journal of Finance*, 51 (1996), 889–919.
Mech, T. S. "Portfolio Return Autocorrelation." *Journal of Financial Economics*, 34 (1993), 307–344.
Mitchell, M.; T. Pulvino; and E. Stafford. "Limited Arbitrage in Equity Markets." *Journal of Finance*, 57 (2002), 551–584.
Pasquariello, P., and C. Vega. "Informed and Strategic Order Flow in the Bond Markets." *Review of Financial Studies*, 20 (2007), 1975–2019.
Sadka, R., and A. Scherbina. "Analyst Disagreement, Mispricing, and Liquidity." *Journal of Finance*, 62 (2007), 2367–2403.
Schwert, G. W. "Stock Market Volatility and the Crash of '87." *Review of Financial Studies*, 3 (1990), 77–102.

https://doi.org/10.1017/S0022109011000081 Published online by Cambridge University Press