Federal Reserve Bank of New York
Staff Reports

The Microstructure of Cross-Autocorrelations

Tarun Chordia
Asani Sarkar
Avanidhar Subrahmanyam

Staff Report no. 303
September 2007

The Microstructure of Cross-Autocorrelations
Tarun Chordia, Asani Sarkar, and Avanidhar Subrahmanyam
Federal Reserve Bank of New York Staff Reports, no. 303
September 2007
JEL classification: G10, G14

Abstract

This paper examines the mechanism through which the incorporation of information into prices leads to cross-autocorrelations in stock returns. The lead-lag relation between large and small stocks increases with lagged spreads of large stocks. Further, order flows in large stocks significantly predict the returns of small stocks when large stock spreads are high. This effect is consistent with the notion that trading on common information takes place first in the large stocks and is then transmitted to smaller stocks with a lag, suggesting that price discovery takes place in the large stocks.

Key words: lead-lag, returns, small stocks, large stocks, microstructure, information

---

Chordia: Goizueta Business School, Emory University (e-mail: tarun_chordia@emory.bus.edu). Sarkar: Federal Reserve Bank of New York (e-mail: asani.sarkar@ny.frb.org). Subrahmanyam: Anderson School of Management, University of California at Los Angeles (e-mail: subra@anderson.ucla.edu). We thank 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, Luboš Pástor, Sebastien Pouget, Joshua Rosenberg, Christoph Schenzler, Gordon Sick, Rob Stambaugh, Neal Stoughton, Marie Sushka, Jiang Wang, and seminar participants at Arizona State University, the University of Alberta, the University of Calgary, the "International Conference on New Financial Market Structures" at HEC Montréal, and the Federal Reserve Bank of New York for helpful comments.

1 Introduction

The manner in which information is impounded 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, it is clear that trading frictions impede the smooth incorporation of information into prices. The cross-autocorrelation effect, first documented by Lo and MacKinlay (1990) (henceforth LM), points to some friction that impedes the smooth flow of information into prices. LM document that 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. A number of explanations have been proposed for this lead-lag pattern in stock returns.

One set of explanations suggests that the leads and lags are a result of non-synchronous trading; the idea being that the information that has already been impounded 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 simply a restatement of portfolio autocorrelation and contemporaneous correlations (Boudoukh, Richardson and Whitelaw, 1994 and 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, 1988)).

Chordia and Swaminathan (2000) argue that the above explanations do not fully account for the lead-lag effect. Instead, they provide support for the speed of adjustment hypothesis,[1] which suggests that the lead-lag patterns arise due to differential speeds of adjustment of stocks to economy-wide information shocks. Some stocks adjust faster while others adjust more slowly. The following question then arises. Is it primarily the speedy adjustment of larger stocks or the slow adjustment of smaller stocks that gives rise to the lead-lag effect? Intuition suggests that information should first be incorporated into the prices of larger stocks due to their lower trading costs. However, Mech (1993) has suggested that it is the high trading costs of small stocks that leads to their slower adjustment. In this paper, we exploit the recent availability of microstructure data to

[1]This hypothesis was first proposed by Brennan, Jegadeesh and Swaminathan (1993). See also Badrinath, Kale and Noe (1995), McQueen, Pinegar and Morley (1996), Connolly and Stivers (1997), and Hou (2007).

search for the exact mechanism by which the incorporation of information into prices leads to the cross-autocorrelation patterns in stock returns. More specifically, we examine the impact of liquidity on the lead-lag patterns.

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, for example, Mitchell, Pulvino, and Stafford, 2002, Sadka and Scherbina, 2006, Hou and Moskowitz, 2005, Avramov, Chordia, and Goyal, 2006, and Chordia, Roll and Subrahmanyam, 2006). 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 on the relation between liquidity and return cross-autocorrelations. Our aim is to gain an understanding of the underlying causes of the lead-lag patterns by exploiting the dynamic links between liquidity, volatility, and returns as suggested by Chordia, Roll, and Subrahmanyam (2001).

Using vector autoregressions, 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, cross-autocorrelation patterns in returns are strongest when large stock spreads are high. Further, order flows in the large stocks play an important role in predicting small stock returns when large stock spreads widen. On the other hand, small stock order flows do not predict large or small stock returns. The hitherto undocumented role of order flow is consistent with price discovery taking place in the large stocks. If common information is first traded upon in the large stocks, then we would expect large stock spreads to widen first, followed by small stock spreads as information is subsequently incorporated into prices of small stocks with a lag.

In order to check that it is indeed trading in response to common information shocks that leads to the autocorrelations, we examine the impact of non-farm employment and Consumer Price Index announcements on the lead-lag effect. The arguments of Kim and Verrecchia (1994) indicate that information collection is stimulated after a public announcement as agents interpret the implication of the announcement for future cash flows. Consistent with this argument, we find that when the lagged order imbalance and lagged spreads of large stocks are higher, the returns of small stocks are also higher on the day of and the day following the announcement than during the rest of the sample period. This suggests that it is the trading in large stocks that follows economywide 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 no such role for small stock liquidity. 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 inefficiency should be broadened to include liquidity spillovers between asset classes. This point is implicitly recognized by researchers 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 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.

The rest of the paper is organized as follows. Section 2 provides a brief economic motivation for our analysis. Section 3 describes how the liquidity data is generated, while Section 4 presents basic time-series properties of the data and describes the adjustment process to stationarize the series. Section 5 presents the vector autoregressions involving order flows, liquidity, returns, and volatility across large and small-cap stocks. Section 6 considers the role of liquidity in the lead-lag relation between large and small-cap returns. Section 7 discusses robustness checks using a holdout sample of Nasdaq stocks and section 8 concludes.

2 Methodology and Economic Motivation

This paper's aim is to examine the microstructure of the lead-lag pattern by exploring the linkage between liquidity dynamics and the crossautocorrelation patterns in stock returns. Since there is an intimate link between liquidity and volatility (Chordia, Roll, and Subrahmanyam,, 2001), we consider the joint dynamics of liquidity, returns, and volatility using a structural model. We use the vector autoregression technique because bi-directional causalities across our variables are economically plausible. In particular, liquidity can affect volatility by attracting more trading, and can affect prices through the traditional channel of Amihud and Mendelson (1986). But, returns may also influence future trading behavior, which may, in turn, affect liquidity. In particular, both standard portfolio rebalancing arguments (Merton, 1973) as well as loss aversion (Odean, 1988) imply return-dependent investing behavior that, by creating an order imbalance, may affect liquidity. In addition, volatility may affect liquidity by affecting the inventory risk borne by market makers. Our vector autoregressions are motivated by the preceding arguments and allow us to examine whether the lead-lag patterns survive the consideration of liquidity dynamics.

We next explore the following question: Why specifically might movements in liquidity be related to return cross-autocorrelations? There are two possible reasons. First, arbitrageurs may choose to trade in small-cap stocks in order to profit from common information shocks that have already been incorporated into the prices of large firms. As suggested by Mech (1993) a widening of small-cap 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-cap spreads are high. The reasoning offers little role for large-cap spreads, since arbitrageurs' activity is initiated in the small firms whose returns lag those of the large firms.

Arbitrage, however, is not necessary for closing the lead-lag gap because market makers in the small-cap sector may directly use price quotes from the large-cap sector to update their own quotes. This leads us to our second line of argument, which indicates that *large-cap* spreads may play a role in determining leads and lags by signaling the occurrence of informational events.

To understand this second argument, note that if agents with information about common factors choose to exploit their informational advantage in the large-cap sector (which has a higher baseline level of liquidity than the small-cap sector), then lagged quote updating by small-cap market makers may cause 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).

Since the trading that causes the lead-lag in the above line of argument is expected to reduce liquidity temporarily in the large-cap sector (Glosten and Milgrom, 1985), spread increases in the large-cap sector may portend a lagged adjustment of small firm returns to large firm returns. Further, if it is the case that the content of information-based trades is reflected first in the large-cap sector, we would expect both large-cap order flows and large-cap returns to play important roles in predicting small-cap returns.

In the first line of argument, lagged small-cap spreads play a crucial role in determining the extent of the lead-lag relationship, whereas in the second, lagged large-cap spreads are relevant. Furthermore, the two arguments are not mutually exclusive. In order to distinguish between the two we analyze the link between the extent of the lead-lag relationship and the levels of large and small-cap spreads.

3 Data

Stock liquidity data were obtained for the period January 1, 1988 to December 31, 2002 (the data extends the sample of Chordia, Roll, and Subrahmanyam, 2001, by four additional years). The data sources are the Institute for the Study of Securities Markets (ISSM) and the New York Stock Exchange TAQ (trades and automated quotations). The ISSM data cover 1988-1992, inclusive, while the TAQ data are for 1993-2002.

We follow the filter rules and selection criteria in Chordia, Roll and Subrahmanyam (2001) to extract transaction-based measures of liquidity.[2] The measures we extract 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; and (iii) depth (DEP), measured as the average of the posted bid and ask dollar amounts offered for trade at the inside quotes.[3] The transactions based liquidity measures are averaged over the day to obtain daily liquidity measures for each stock. The daily order imbalance (OIB) (used in Section

[2]The following securities were not included in the sample since their trading characteristics might differ from ordinary equities: ADRs, shares of beneficial interest, units, companies incorporated outside the U.S., Americus Trust components, closed-end funds, preferred stocks and REITs.

[3]We have also performed alternative analyses using effective spreads, defined as twice the absolute difference between the transaction price and the mid-point of the prevailing quote. The results are largely unchanged from those for quoted spreads and so, for brevity, we do not report them in the paper.

6 and 7) 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 is 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 ranges from about $26 billion for the largest decile to about $47 million for the smallest one.[4] 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.

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 :

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

where D_j is a dummy variable for the day of the week and R_{it} (also the variable RET used below) represents the value-weighted average of individual stock CRSP returns for a particular decile.

4 Basic Properties of the Data

4.1 Summary Statistics

In Table 1, we present summary statistics associated with liquidity measures, together with information on the daily number of transactions for the two size deciles. Since previous studies such as Chordia, Roll, and Subrahmanyam (2001) suggest that the reduction in tick sizes likely had a major impact on bid-ask spreads, we provide separate statistics

[4] For the middle eight 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.

for the periods before and after the two changes to sixteenths and decimalization.[5] We find that spreads for large and small stocks are very close to each other (18.6 and 19.1 cents, respectively) before the shift to sixteenths, but they diverge considerably after the shift. Indeed, the average spread for large stocks is half that of small stocks (5.0 versus 10.2 cents) in the period following decimalization. While we have verified that both of these differences are statistically significant,[6] the point estimates indicate that decimalization has been accompanied by a substantial reduction in the spreads of large stocks, which is consistent with the prediction of Ball and Chordia (2001). The increasing popularity of exchange-traded funds (ETFs), would likely result in still lower spreads for broad portfolios (Subrahmanyam, 1991).

The difference in mean inside depths of large and small stocks has narrowed in recent times, and the differences are statistically distinguishable from zero. The depth of large stocks is on average double that of small ones in the pre-sixteenths period, but it is about 50% higher than depth in the small-cap sector in the post-decimalization period. Depths have decreased after decimalization relative to the eighths regime. This is consistent with the prediction of Harris (1994), and an unreported t-test indicates that these decreases are also statistically significant for both small and large-cap stocks. The average daily number of transactions has increased substantially in recent years for both large and small-cap stocks. For example, the average daily number of transactions for the large stocks increased from 580 in the first subperiod (before the shift to sixteenths) to 3,984 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 two changes in the tick size and also demonstrates how large stock spreads have diverged from those of small stocks towards 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

[5] 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.

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

focus on quoted spreads, rather than depth. In any case, our principal results are robust to using the proportional spread series as well as depth.

4.2 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 time-series variations. Each time-series, returns, spreads, depths, 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 calendar regularities in quoted spreads. We do not present results for other variables, nor for the variance equation (3). These results are available upon request.

We are interested in differences in the adjustment regression coefficients between the different size sectors. A readily noticeable finding is that the nature of calendar regularities in liquidity is different across large and small stocks. For example, January spreads are higher for large stocks than spreads in other months (all dummy coefficients from February to December are negative and significant for large-cap stocks). This regularity is much less apparent for small-cap stocks since only the November and December coefficients are negative and significant in the regressions. To confirm a January effect in large-cap spreads, we compare the mean difference in January spreads across the two sectors and find that large-cap spreads are significantly higher than small-cap spreads at the 5% level. In addition, omitting all of the monthly dummy coefficients and including only the January dummy, we find that this dummy is not significant for small-cap stocks. However, it is significant with a t statistic of 12.17 for large-cap stocks. Thus, overall the evidence indicates that large-cap spreads are significantly higher in January, but the same is not true for small-cap stocks.[7]

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

The January behavior in spreads may be due to the fact that portfolio managers shift out of the large-cap sector following window-dressing in December. In addition, there has been a strong negative trend in spreads since decimalization for both small and large companies. The higher spreads on Mondays for small-cap stocks are to be understood in conjunction with a day-of-the-week effect in returns. In unreported results, we find that (consistent with French, 1980 and Gibbons and Hess) returns are higher on Fridays than on other days, while order flow is tilted significantly to the sell-side for small stocks on Mondays (relative to Fridays). Thus, agents appear to trade in order to countervail the buying pressure on Fridays. This "rebound" selling on Mondays following high returns towards the end of the week can contribute to increased spreads, as market makers struggle to offload the increased inventory.[8] The adjustment results for stock depths are not reported but are generally consistent with those for spreads.

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 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.

5 Vector Autoregresssion: NYSE Stocks

Section 2 indicates that there are sound economic reasons to expect cross-sector effects and bi-directional causalities. Thus, in the spirit of Hasbrouck (1991) and Chordia, Sarkar, and Subrahmanyam (2005) we adopt an six-equation vector autoregression (VAR) that incorporates six variables, (three each - i.e., measures of liquidity, returns, and volatility - from large and small-cap stocks).[9] The VAR thus includes the endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, and QSPR9, whose suffixes 0 and 9

[8]Chordia, Roll, and Subrahmanyam (2001, 2002) show that down markets and high levels of absolute order imbalances are accompanied by decreased liquidity.

[9]Inclusion of order flows (OIB) as an additional variable in the VAR does not alter any of the major conclusions.

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. Typically, the slope of the information criterion (as a function of lags) is quite flat for larger lag lengths, so the choice of smaller lag lengths is justified. The suggested lag length by this procedure is two. The VAR is therefore estimated with this lag length together with a constant term and uses 3782 observations. Since the key contribution of our study is to examine persistent spillovers across different stock market sectors, we first present Chi-square 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 zero when j is the 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.

Consistent with LM, large-cap returns cause small-cap returns but the reverse is not true; thus, large-cap returns lead small-cap returns even after accounting for liquidity dynamics. Further, within each market, there is two-way Granger causation between quoted spreads and volatility. While large-cap returns do Granger-cause large-cap spreads, small-cap spreads are not Granger-caused by small-cap returns. An economic interpretation is that order flow shocks have larger magnitudes in the large-cap sector than in the small-cap sector, possibly because of more herding (and thus more extreme imbalances) in large-cap stocks, that tend to be owned more often by institutions (Sias and Starks, 1997, Dennis and Strickland, 2003). Hence, price movements induced by inventory imbalances may have a greater persistent effect on large-cap liquidity than on small-cap liquidity.

Overall, there is compelling evidence that both own- and cross-volatilities are relevant in forecasting liquidity in a given sector so that volatility shifts in either sector play a key role in liquidity dynamics in both sectors. 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. An IRF traces the impact of a one 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. Results from the IRFs are generally sensitive to the specific ordering of the endogenous variables.[10] 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 three variables. Also, our qualitative results remain mostly unchanged if we reverse the ordering of small and large-cap stocks; we note instances when this is not the case.

Panels A and B of Figure 2 illustrate the response of endogenous variables in a particular (large- or small-cap) sector to a unit standard deviation orthogonalized shock in the endogenous variables in the other sector for a period of ten days. Monte Carlo two-standard-error bands are provided to gauge the statistical significance of the responses. Period 1 in the impulse response functions 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). We find that return spillovers persist even after accounting for liquidity, in that large-cap stock returns are useful in forecasting small-cap returns for a period of one day (this finding is explored further in the next section). There also is clear evidence that shocks to large-cap volatility are useful in forecasting small-cap volatility and vice versa. In addition, large-cap spread responds negatively to an innovation in small-cap stock returns and positively to a shock to small-cap volatility. In both cases, the response persists for at least ten days, illustrating the strength of the cross-market effects. Finally, small-cap spreads respond to large-cap spreads, as well as to large-cap volatility and returns.

Are these results robust to the relative ordering of the small and large-cap sectors? 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. In unreported own-sector impulse response functions, we find that volatility shocks in a sector result in a persistent decline in liquidity in that sector. As suggested by Stoll (1978) increased volatility, by increasing inventory risk, tends to decrease liquidity. We also examine the impulse responses of large-cap or decile 9 stocks to other deciles (e.g. decile 5) and find that the results are qualitatively similar to the previously reported

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

responses of large-cap stocks to decile 0 stocks.

6 Impact of Liquidity on Cross-Autocorrelations

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 2 suggest that we should expect such interactions. For example, informational shocks impact liquidity, and if these also influence return spillovers, then we would expect the strength of spillovers to be related to liquidity.

The Granger-causality results of Section 5 indicate that large-cap returns are informative in predicting small-cap returns. This is consistent with LM who document that large-cap returns lead small-cap returns at short horizons. Brennan, Jegadeesh and Swaminathan (1993) suggest that the cross-autocorrelations may be caused by differences in the speed of adjustment to common information. If informed traders trade first in the large stocks (due to lower trading costs) then spreads in the large stocks should increase first and this increase in spreads should presage an impact on small stock returns as the information is incorporated into small stocks with a lag.

Given that Chordia, Roll, Subrahmanyam (2007) have shown that liquidity influences return predictability, it is possible that illiquidity may impact the speed of adjustment. This provides another motivation for examining whether liquidity dynamics are related to the lead-lag relationship between large firm and small firm returns. We follow Brennan, Jegadeesh and Swaminathan (1993) in conducting the lead-lag analysis at the daily frequency.[11]

We capture the influence of liquidity levels and order-flow dynamics on the lead-lag relationship between small and large-cap stocks by adding a number of interaction variables

[11] In an exploratory investigation, we considered a weekly horizon similar to that used by Lo and MacKinlay (1990) and subsequently in studies by Mech (1993), Badrinath, Kale, and Noe (1995), McQueen, Pinegar, and Thorley (1996), and Sias and Starks (1997). 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 one 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.

to the equation for RET0 within the VAR framework. These interaction variables include the first lags of QRET09, QRET99, and QOIB99, where QRET09=QSPR0*RET9, QRET99=QSPR9*RET9, and QOIB99= QSPR9*OIB9. If, as discussed above, 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 OLS method is no longer efficient. Thus, we use the Seemingly Unrelated Regression (SUR) method to estimate the system of equations.

Table 4 presents the results of these regressions. The coefficient of lagged return alone (which already is part of the main VAR) is a statistically significant 0.088 consistent with the lead-lag result of LM. The second column interacts the spread in large and small-cap stocks with the lagged large-cap return. The magnitude of the coefficient on the lagged large-cap return is considerably reduced and becomes insignificant after including the interaction variables. The coefficient on QRET99 (large-cap spread interacted with returns) is a significant 0.368, suggesting that the lead-lag relation between lagged returns of large-cap stocks and the current returns of small-cap stocks is strongest when the large-cap sector is illiquid. Thus, the evidence is consistent with the notion that a widening of large-cap spreads signals an information event that is transmitted to small-cap stocks with a lag.

6.1 Is the Lead-Lag Effect due to Informed Trades?

In this section, we further test the notion that information gets transmitted to prices in either sector by way of informed order flows. We interact order imbalance (OIB) with spreads in the large-cap market (to construct the variable QOIB99) and include the lag of this interaction variable in the regression. The results, shown in the third column of Table 6, indicate that large-cap order flow interacted with large-cap spreads is strongly predictive of small-cap returns, whereas the return interaction variable becomes insignificant and its magnitude diminishes in the presence of the OIB. We also present the chi-square statistics and p-values associated with the Wald test for the null hypothesis that the coefficients of all exogenous variables are jointly zero. We reject the null hypothesis that the coefficients of the imbalance interaction term OIB99 and the spread-return interaction terms QRET09 and QRET99 are jointly zero at a p-value below 0.001. Small cap order flow is not significantly related to future large or small cap returns, consistent with the notion that informational events first occur in large cap stocks and then spillover to small cap stocks, and not vice versa. Overall, the evidence supports the hypothesis that large-cap order flows induced by informational events drive the lead and lag relationship between large-cap and small-cap firms.

In order to provide additional insight into the results of Table 6, we calculate the correlation between small-cap returns on day t and large-cap returns on day $t-1$ for days $t-1$ where the quoted large-cap spread is one standard deviation above and below its sample mean. The estimates obtained for the two cases are 0.20 ($p = 0.00$) and 0.05 ($p = 0.10$). The corresponding correlations when the large-cap order imbalance is used in place of returns are 0.15 ($p = 0.00$) and 0.08 ($p = 0.06$). These correlations clearly confirm our basic result that the lead-lag from large-cap returns to small-cap returns is strongest when large-cap spreads are high.

Of course, the information-based trading that causes large-cap spreads to widen may spill over to small-cap stocks for two reasons. First, some investors may receive information later than others (Hirshleifer, Subrahmanyam, and Titman, 1994), or, alternatively, the informed traders may prefer to exploit their information by first trading in the more liquid large stocks. Second, small-cap market makers may not be able to update their quotes to fully reflect the information content of large-cap trades, owing to camouflage provided by liquidity trades (Kyle, 1985). This would leave some profit potential for late informed traders in small-cap stocks. If large-cap informed trading does indeed spill over to small-cap stocks with a lag, we expect small-cap order flows to exhibit an increased correlation with lagged large-cap order flows when large-cap spreads are high. Additionally, a greater small-cap spread on day t should be associated with a greater correlation between small-cap returns at time t and large-cap returns at time $t-1$. This is what we find.

First, the correlation between OIB0 on day t and OIB9 on day $t-1$ is 0.14 ($p < 0.01$) when QSPR9 on day $t-1$ is one standard deviation above its mean and 0.09 ($p = 0.05$) otherwise. Second, we sort the sample by days when the small-cap spread is one standard deviation above and below its sample mean. The correlation between day t small-cap returns and day $t-1$ large-cap returns is 0.15 (0.05) when the small-cap spread is above (below) its sample mean on day t. Only the correlation of 0.15 is significantly different from zero at the 5% level. To clarify, these correlations document a link between cross-autocorrelations and small stock spreads at time t. As Table 6 demonstrates, there is no significant link between small stock return predictability from large stocks and small stock spreads at time $t-1$. This is consistent with information events widening spreads first in large stocks at time $t-1$, and then in small stocks at time t. When the order imbalance replaces returns, the corresponding correlations are 0.09 and 0.07, respectively; again, only the first correlation is significantly different from zero at the 5% level. Thus, the evidence is consistent with the cross-autocorrelations being caused by liquidity-straining trades that occur first in the large-cap sector and then in the small-cap sector.

We next report results for all other deciles in Table 5. We use the same interaction variables as above, except that we replace QRET09 with QRETN9=QRETN*RET9, where N represents the size decile. We make a similar replacement for the OIB variable. The table shows that the large-cap order flow variable interacted with large-cap spreads is informative in predicting returns in every size decile. With the exception of decile 1, the coefficient magnitudes on the order flow variable generally decline with size decile, and the magnitudes for the four largest firm deciles is about 40% smaller than for the four smallest ones.

Our 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-cap sector and in the form of lagged quote updates by small-cap market makers. Our return computations are based on transaction prices and account for transaction-induced lags. However, small stocks often do not trade for several hours within a day. Thus, if the last transaction in a stock is at 10:00 am, for instance, then the transaction price would not incorporate information shocks that occur later in the day.

To address the above issue, we perform an alternative analysis by computing mid-quote returns using the last available quote for each firm on a given day. We do this for the 1993-2002 period because the timing of quotes and trades is more reliable on the TAQ dataset which begins in 1993. One benefit of using the post-1993 sample is that it allows us to assess whether the lead-lag relation between small and large firm returns is particular to the earlier part of our sample. The results appear in the last two columns 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 but one case.[12] Thus, our transaction price-based results on predictability extend to mid-quote returns as well, and our earlier results continue to hold for the post-1993 sample.

The midquote return results shed additional light on the economic causes of the lead-lag effect. Specifically, one possible interpretation of Table 5 is that secular decreases in liquidity can reduce trading activity in small-cap stocks and this reduction can affect leads and lags. Our results point to the notion that this effect is not the predominant driver of lead-lag between the large and small-cap sectors. To see this, observe that the mid-quote 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. Since our results are robust to both transaction returns as well as mid-quote returns, they 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-cap stocks. Overall, our findings underscore the role of order flow in the lead-lag relationship between the large-cap sector and other 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-cap 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 (1993), 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-cap stocks. Moreover, unlike Mech (1993), we consider the role of large-cap spreads in addition to small-cap spreads in determining the extent of the lead-lag relationship, and we find that large-cap spreads are most relevant to the lead-lag effect.

The economic significance of our results is material, though it does not suggest a gross violation of market efficiency. For example, the standard deviation of the return

[12]The Wald test shows that, as before, all p values except the one for decile 7 (where none of the variables are significant) are less than 0.05.

and spread-based interaction variable in Table 4 is about 0.002. Based on the relevant coefficient (0.368) in Table 4, we find that a one standard deviation move in the interaction variable changes small-cap returns by 0.073%. However, assuming 83 such events in an year (i.e., on about 33% of days forming a typical 250 trading-day year), this works out to 6.08% on an annual basis. Based on the midquote return coefficient of 0.054 (for the smallest firm decile) in Table 5, a one standard deviation move (equaling about 0.015) in the order-flow based interaction variable has a daily effect of 0.068%, aggregating to about 6.52% across 83 events. Trading frictions including market impact costs and brokerage commissions, however, could nullify the profitability of such strategies.

6.2 Common Information and Cross-Autocorrelations

In this section we show that it is indeed trading in response to common information that causes the lead-lag effect. We study two major macroeconomic announcements, viz., those for non-farm employment and the Consumer Price Index (CPI). Fleming and Remolona (1999) and Chordia, Roll, and Subrahmanyam (2001) suggest that amongst all major macroeconomic news releases, these announcements have the strongest impact on financial market liquidity and price formation. Consequently, we analyze the strength of the lead-lag effect around these announcements.

Table 6 presents the results on the impact of trading around the announcements. Specifically, we define a dummy variable, MACRO, that takes the value one over the two day period including and following the announcements. We interact MACRO and (1-MACRO) with the variables in Table 4, i.e., lagged values of RET9, QRET9, and QOIB9. 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.

We find that the coefficient on lagged QOIB99 when interacted with MACRO is a significant 0.084 and when interacted with (1-MACRO) it is 0.063 and is also significant. Moreover the difference between the two is statistically significant. This suggests that the impact of large cap liquidity and trading on small cap returns is greater on the day of and the day following the announcements than during other days of the sample period.[13] These results are consistent with Kim and Verrecchia (1994) who indicate that trading on information intensifies after a public information shock as agents interpret the implication of the announcement for future cash flows.

Existing literature also suggests examining the lead-lag effect before the announcement. For example, Beber and Brandt (2006) find that market participants wait to trade until macroeconomic uncertainty is resolved (the so-called "calm before the storm" effect). Consistent with this notion, Balduzzi, Elton, and Green (2001) find that trading activity is lower prior to public announcements. Fleming and Remolona (1999) suggest that this phenomenon may be due to dealer's reluctance to make markets before a large resolution of uncertainty. These studies indicate that the extent of information-based trading may be lower prior to public announcements. To check this we examine the cross-autocorrelation effect just before the non-farm employment and CPI announcements. Thus, we now define MACRO to take value one over a two day period just prior to the announcements.[14] The results in last two columns of Table 6 show that the coefficient on the interaction term of MACRO and QOIB99 is 0.061 and that on the interaction term of (1-MACRO) and QOIB99 is 0.068, suggesting that the impact of the interaction variable is smaller during the pre-announcement period. Hence, consistent with our conjecture, large cap order flow has less impact on the lead-lag relation before announcements compared to the remaining sample. It is worth noting, however, that while the difference in the two QOIB coefficients is statistically significant, the economic significance of the difference (0.007) is not as strong as that of the corresponding post-announcement ones.

Overall, the results support the hypothesis that it is trading in response to common information shocks that causes the lead-lag effect.

[13] We also perform an analysis where the interactions of MACRO with lagged RET9 are the only exogenous variables in the regression (corresponding to the first two columns of Table 4). We find that the estimated coefficient of lagged RET9 is significantly greater on post-announcement days compared to other days, confirming that the lead lag effect is stronger during the post-announcement period.

[14] 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. Also, in a few cases, the employment and CPI announcements were close together, and the "before" days of one coincided with the "after" days of the other. There were 11 such days. To prevent ambiguity, we delete such observations.

7 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.[15] 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 3; 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 (though results are substantively similar for transaction return series).

As before, we adjust the Nasdaq series of spreads, returns, and volatility to account for regularities.[16] 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 3 lags and a constant term. The Granger causality results across large-cap NYSE stocks and Nasdaq stocks are presented in Panel A of Table 7. There is evidence of a liquidity spillover from NYSE to Nasdaq, but not from Nasdaq to NYSE stocks. Recall that in Table 5, there is evidence of bivariate causality from large-cap to small-cap stocks and vice-versa. The Granger causality results indicate that liquidity discovery may take place in the larger exchange market.

In Panel B of Table 7, we present the analog of the lead-lag regression in Table 5.

[15]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).

[16]For Nasdaq stocks, the dummy variable for the change to sixteenths equals 1 for the period June 12, 1997 to March 11, 2001. Further, there are 3 decimalization dummies 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 from April 9, 2001 to December 31, 2002 (see, for example, Chung and Chuwonganant, 2003). Consistent with Bessembinder (2003), the value-weighted average spread on Nasdaq of 44.2 cents in our sample period prior to June 12,1997 drops to 12.5 cents in the period June 12, 1997 to March 24, 2001, and remains at just 3.6 cents for the remainder of our sample period. The high spread in the pre-sixteenth period relative to that of NYSE stocks is consistent with Huang and Stoll (1996).

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-cap imbalance interacted with large-cap spreads is significant, and its magnitude is greater than the corresponding coefficient magnitude in Table 5. Overall, these results confirm the role of large-cap order flows and liquidity in predicting returns in both small-cap NYSE firms as well as Nasdaq firms. Thus, our key results continue to obtain even after inclusion of Nasdaq stocks in our analysis.

8 Concluding Remarks

Our principal aim 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 use a structural model to examine the joint time-series of liquidity, returns, and volatility for NYSE size decile portfolios and the value-weighted Nasdaq portfolio over the period 1988 through 2002. Our analysis of size-based portfolios is motivated by Lo and MacKinlay (1990) who document the lead-lag effect in returns.

We find that return cross-autocorrelations survive even after accounting for liquidity dynamics. Our results also show that order flows in the large-cap sector predict small-cap returns when large-cap spreads are high. This result holds for both transaction returns as well as the end of day mid-quote returns, demonstrating that the finding is not due to stale prices. The role of liquidity in cross-autocorrelations intensifies during major macroeconomic announcements. These findings are consistent with our hypothesis that common informational events impact the large-cap sector first, causing large-cap spreads to widen, and are subsequently incorporated with a lag into the prices of small-cap stocks. Price discovery takes place in the large stocks.

As a by-product of our analysis, we document some interesting differences in calendar regularities across market cap-based deciles. For instance, within our sample period, there is a distinct upward pressure on large-cap NYSE spreads in January, relative to other months, that is not as strongly evident in small-cap stocks. This finding is consistent with portfolio managers withdrawing from the large-cap sector following window-dressing at the turn of the year. Spreads of large-cap stocks are lowest at the beginning of the week but those of small-cap stocks appear to be highest at this time, and small-cap order imbalances are tilted towards the sell side at the beginning of the week. This pattern accords with the notion that arbitrageurs indulge in net selling activity in small-cap stocks at the beginning of the week following the upward pressure on small-cap returns at the end of the week.

From the standpoint of asset pricing, our 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, it would be important 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

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:

$$w = x'\beta + u \quad \text{(mean equation).} \tag{2}$$

In equation (2), w is the series to be adjusted and x contains the adjustment variables. The residuals are used to construct the following variance equation:

$$log(u^2) = x'\gamma + v \quad \text{(variance equation).} \tag{3}$$

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

$$w_{adj} = a + b(\hat{u}/exp(x'\gamma/2)), \tag{4}$$

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) four dummies for Monday through Thursday, (ii) 11 month of the year dummies for February through December, and (iii) a dummy for non-weekend 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, 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 dummies for financial market events that could affect the liquidity of both small and large-cap stocks. Specifically, we include (v) 3 crisis dummies, where the crises are: the Bond Market crisis (March 1 to May 31, 1994), the Asian financial crisis (July 2 to December 31, 1997) and the Russian default crisis (July 6 to December 31, 1998);[17] (vi) dummies for the day of and the two days prior to macroeconomic

[17]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, October 1999, available at http://www.bis.org/publ/cgfspubl.htm).

announcements about GDP, employment and inflation (intended to capture informed trading and portfolio balancing around public information releases); (vii) a dummy for the period between the shift to sixteenths and the shift to decimals, and another for the period after the shift to decimals; (viii) a dummy for the week after 9/11/01, when we expect liquidity to be unusually low, and (ix) a dummy for 9/16/91 where, for some reason (most likely a recording error) only 248 firms were recorded as having been traded on the ISSM dataset whereas the number of NYSE-listed firms trading on a typical day in the sample is over 1,100.

References

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

Admati, A. and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 1, 3-40.

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

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

Avramov, D., T. Chordia, and A. Goyal, 2006, Liquidity and autocorrelations in individual stock returns,*Journal of Finance* 61, 2365-2394.

Badrinath, S., J. Kale, and T. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, *Review of Financial Studies* 8, 401-430.

Balduzzi, P., E. Elton, and C. Green, 2001, Economic news and bond prices: Evidence from the U.S. Treasury market, *Journal of Financial and Quantitative Analysis* 36, 523-543.

Ball, C., and T. Chordia, 2001, True spreads and equilibrium prices, *Journal of Finance* 56, 1801-1836.

Beber, A., and M. Brandt, 2006, The effect of macroeconomic news on beliefs and preferences: Evidence from the options market, *Journal of Monetary Economics* 53, 1997-2039.

Benston, G., and R. Hagerman, 1974, Determinants of bid-asked spreads in the over-the-counter market, *Journal of Financial Economics* 1, 353-364.

Bessembinder, H., 2003, Trade execution costs and market quality after decimalization, *Journal of Financial and Quantitative Analysis* 38, 747-777.

Borio, C., and R. McCauley, 1996, The economics of recent bond yield volatility, BIS Economic Paper No. 44, Bank for International Settlements, Basel, Switzerland.

Boudoukh, J., M. Richardson, and R. Whitelaw, 1994, A tale of three schools: Insights on autocorrelations of short-horizon stock returns, *Review of Financial Studies* 7, 539-573.

Branch, B., and W. Freed, 1977, Bid-asked spreads on the AMEX and the big board, *Journal* *of Finance* 32, 159-163.

Brennan, M., N. Jegadeesh, and B. Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *Review of Financial Studies* 6, 799-824.

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

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

Chan, K., 1993, Imperfect information and cross-autocorrelation among stock prices, *Journal of Finance* 48, 1211-1230.

Chan, K., and W. Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, *Journal of Financial Economics* 57, 247-273.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3-28.

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

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

Chordia, T., R. Roll, and A. Subrahmanyam, 2006, Liquidity and market efficiency, forthcoming, *Journal of Financial Economics.*

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, An empirical analysis of stock and bond market liquidity, *Review of Financial Studies* 18, 85-129.

Chordia, T., L. Shivakumar, and A. Subrahmanyam, 2004, Liquidity dynamics across small and large firms, *Economic Notes* 33, 111-143.

Chordia, T., and B. Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, *Journal of Finance* 55, 913-935.

Chowdhry, B., and V. Nanda, 1991, Multimarket trading and market liquidity, *Review of*

Financial Studies 4, 483-511.

Chung K., and C. Chuwonganant, 2004, Tick size, order handling rules, and trading costs, *Financial Management* 33, 47-62.

Clark, R., J. McConnell, and M. Singh, 1992, Seasonalities in NYSE bid-ask spreads and stock returns in January, *Journal of Finance* 47, 1999-2014.

Connolly, R., and C. Stivers, 1997, Time-varying lead-lag of equity returns in a world of incomplete information, working paper, University of North Caroline at Chapel Hill.

Conrad, J., and G. Kaul, 1998, An anatomy of trading strategies, *Review of Financial Studies* 11, 489-519.

Conrad, J., M. Gultekin, and G. Kaul, 1991, Asymmetric predictability of conditional variances, *Review of Financial Studies* 4, 597-622.

Davis, P., 1999, A reappraisal of market liquidity risk in the light of the Russia/LTCM global securities market crisis, working paper, Bank of England, London, U.K.

Dennis, P., and D. Strickland, 2003, Who blinks in volatile markets? Individuals or Institutions? *Journal of Finance* 53, 1923-1949.

Eisfeldt, A., 2004, Endogenous liquidity in asset markets, *Journal of Finance* 59, 1-30.

Emmons, W., and F. Schmid, 2000, The Asian crisis and the exposure of large U.S. firms, Federal Reserve Bank of St. Louis Review, January 2000, 15-34.

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

Fleming, M. and E. Remolona, 1999, Price formation and liquidity in the U.S. Treasury market: The response to public information, *Journal of Finance* 54, 1901-1915.

French, K., 1980, Stock returns and the weekend effect, *Journal of Financial Economics* 8, 55-69.

Gallant, A., P. Rossi, and G. Tauchen, 1992, Stock prices and volume, *Review of Financial Studies* 5, 199-242.

Gibbons, M., and R. Hess, 1981, Day of the week effects and asset returns, *Journal of Business* 54, 579-596.

Glosten, L., and P. Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.

Glosten, L., 1989, Insider trading, liquidity, and the role of the monopolist specialist, *Journal of Business* 62, 211-235.

Gorton, G., and G. Pennacchi, 1993, Security baskets and index-linked securities, *Journal of Business* 66, 1-27.

Hameed, A., 1997, Time-varying factors and cross-autocorrelations in short-horizon stock returns, *Journal of Financial Research* 20, 435-458.

Harris, L., 1995, Minimum price variations, discrete bid-ask spreads, and quotation sizes, *Review of Financial Studies* 7, 149-178.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 56, 179-207.

Hasbrouck, J., and D. Seppi, 2001, Common factors in prices, order flows and liquidity, *Journal of Financial Economics* 59, 383-411.

Haugen, R., and J. Lakonishok, 1987, *The Incredible January Effect*, Dow Jones-Irwin, Homewood, IL.

Hiemstra, C., and J. Jones, 1994, Testing for linear and nonlinear Granger causality in the stock price-volume relation, *Journal of Finance* 49, 1639-1664.

Hirshleifer, D., A. Subrahmanyam, and S. Titman, 1994, Security analysis and trading patterns when some investors receive information before others, *Journal of Finance* 49, 1665-1698.

Hirshleifer, D., A. Subrahmanyam, and S. Titman, 2006, Feedback and the success of irrational investors, *Journal of Financial Economics* 81, 311-338.

Ho, T., and H. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance* 38, 1053-1074.

Hou, K., 2007, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20, 1113-1138.

Hou, K., and T. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981-1020.

Huang, R., and H. Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, ıJournal of Financial Economics 41, 313-357.

Huberman, G., and D. Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161-178.

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

Johnson, T., 2006, Dynamic liquidity in endowment economies, *Journal of Financial Economics* 80, 531-562.

Jones, C., 2001, A century of stock market liquidity and trading costs, working paper, Columbia University, New York, NY.

Jones, C., G. Kaul, and M. Lipson, 1994, Transactions, volume, and volatility, *Review of Financial Studies* 7, 631-651.

Keim, D., 1983, Size-related anomalies and stock return seasonality: Further empirical evidence, *Journal of Financial Economics* 12, 13-32.

Kim, O., Verrecchia, R., 1991a. Trading volume and price reactions to public announcements. Journal of Accounting Research 29, 302321.

Kim, O., Verrecchia, R., 1991b. Market reactions to anticipated announcements. Journal of Financial Economics, 30, 273309.

Kim, O. and Verrecchia, R., 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 4167.

Kumar, A., and C. Lee, 2006, Retail investor sentiment and return comovements, *Journal of Finance* 61, 2451-2586.

Kumar, P., and D. Seppi, 1994, Information and index arbitrage, *Journal of Business* 67, 481-509.

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

Leach, C., and A. Madhavan, 1993, Price experimentation and security market structure, *Review of Financial Studies* 6, 375-404.

Lee, C., and B. Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, *Journal of Financial Markets* 3, 83-111.

Lee, Y., Y. Liu, R. Roll, and A. Subrahmanyam, 2004, Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange, *Journal of Financial and Quantitative Analysis* 39, 327-341.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-747.

Lee, C., A. Shleifer, and R. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75-109.

Lehmann, B., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1-28.

Lo, A., and C. MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* 3, 175-205.

Lo, A., and J. Wang, 2000, Trading volume: definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257-300.

Longin, F., 1997, The threshold effect in expected volatility: A model based on asymmetric information, *Review of Financial Studies* 10, 837-869.

McQueen, G., M. Pinegar, and S. Thorley, 1996, Delayed reaction to good news and the cross-autocorrelation of portfolio returns, *Journal of Finance* 51, 889-919.

Mech, T., 1993, Portfolio return autocorrelation, *Journal of Financial Economics* 34, 307-344.

Merton, R., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

Merton, R., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Mitchell, M., T. Pulvino, and E. Stafford, 2002, Limited arbitrage in equity markets, *Journal of Finance* 57, 551-584.

Muller, A., and W. Verschoor, 2004, The financial crisis exchange rate risk of U.S. multinationals, working paper, Limburg Institute of Financial Economics, Maastricht, Netherlands.

Odders-White, E., 2000, On the occurrence and consequences of inaccurate trade classification, *Journal of Financial Markets* 3, 205-332.

Odean, T., 1998, Are investors reluctant to realize their losses?, *Journal of Finance* 53, 1775-1798.

O'Hara, M., and G. Oldfield, 1986, The microeconomics of market making, *Journal of Financial and Quantitative Analysis* 21, 361-376.

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

Pesaran, H., and Y. Shin, 1998, Generalised impulse response analysis in linear multivariate models, *Economics Letters* 58, 17-29.

Rozeff, M., and W. Kinney, 1976, Capital market seasonality: The case of stock returns, *Journal of Financial Economics* 3, 379-402.

Sadka, R., and A. Scherbina, 2006, Analyst disagreement, mispricing, and liquidity, forthcoming, *Journal of Finance.*

Schwert, W., 1990, Stock market volatility and the crash of 87, *Review of Financial Studies* 3, 77-102.

Sias, R., and L. Starks, 1997, Return autocorrelation and institutional investors, *Journal of Financial Economics* 46, 103-131.

Stoll, H., 1978, The pricing of security dealer services: An empirical study of NASDAQ stocks, *Journal of Finance* 33, 1153-1172.

Stoll, H., 2000, Presidential address: Friction, *Journal of Finance* 55, 1479-1514.

Subrahmanyam, A., 1991, A theory of trading in stock index futures, *Review of Financial Studies* 4, 17-51.

Subrahmanyam, A., 1994, Circuit breakers and market volatility: A theoretical perspective, *Journal of Finance* 49,237-254.

Table 1: Levels of Stock Market Liquidity

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. QSPR stands for quoted spread, NTRADE for the number of shares traded, and DEP for depth measured as the average of the posted bid and ask dollar amounts offered for trade. 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 (S.D.) is reported for each measure. The sample spans the period January 4, 1988 to December 31, 2002; the number of observations is 3782 for all deciles.

	1/4/1988-6/23/1997			6/24/1997-1/28/2001			1/29/2001-12/31/2002		
	Mean	Median	S.D.	Mean	Median	S.D.	Mean	Median	S.D.
QSPR0	0.191	0.191	0.009	0.167	0.166	0.009	0.102	0.101	0.016
DEP0	6.373	6.277	1.036	4.378	4.387	1.206	2.169	2.211	0.502
NTRADE0	13.168	12.488	4.269	31.866	28.162	13.985	47.052	43.558	16.324
QSPR9	0.186	0.185	0.019	0.127	0.124	0.013	0.050	0.047	0.012
DEP9	13.215	12.865	3.515	7.524	7.081	1.788	3.420	3.278	0.633
NTRADE9	579.7	551.2	222.3	2401.1	2295.4	982.1	3984.3	3836.6	1002.5

Table 2: Adjustment Regressions for Liquidity

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. QSPR stands for quoted spread. The sample spans the period January 4, 1988 to December 31, 2002; the number of observations is 3782 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 one if a trading day satisfies the following conditions, (1) if Independence day, Veterans’ Day, Christmas or New Year’s Day falls on a Friday, then the preceding Thursday, (2) if any holiday falls on a weekend or on a Monday then the following Tuesday, (3) if any holiday falls on a weekday then the preceding and the following days, and zero otherwise. Monday-Thursday: equals one if the trading day is Monday, Tuesday, Wednesday, or Thursday, and zero otherwise. February-December: equals one if the trading day is in one of these months, and zero otherwise. The tick size change dummy equals 1 for the period June 24, 1997 to January 28, 2001. The decimalization dummy equals 1 for the period January 29, 2001 till December 31, 2002. Estimation is done using the Ordinary Least Squares (OLS). All coefficients are multiplied by a factor of 100. Estimates marked **(*) are significant at the one (five) percent level or better.

	QSPR0	QSPR9
Intercept	19.218**	21.797**
Day of the week		
Monday	0.198**	-0.160**
Tuesday	-0.017	-0.133*
Wednesday	-0.060	-0.032
Thursday	-0.033	-0.018
Holiday	0.014	-0.002
Month		
February	0.170*	-0.172*
March	0.231**	-0.406**
April	0.281**	-0.285**
May	-0.031	-0.844**
June	-0.066	-0.993**
July	0.072	-0.967**
August	0.067	-1.023**
September	0.016	-1.292**
October	0.089	-0.722**
November	-0.196**	-1.147**
December	-0.605**	-1.042**
Tick size change dummy	-2.749**	-10.951**
Decimalization dummy	-6.598**	-13.879**
Trend, pre-tick size change		
Time	0.000**	-0.002**
$Time^2$	0.0000**	0.0000**
Trend, pre-decimalization		
Time	-0.0014*	0.0120**
$Time^2$	0.003**	0.0000**
Trend, post-decimalization		
Time	-0.0128**	-0.018**
$Time^2$	0.000**	0.000**

Table 3: Granger Causality Results

The table presents Granger causality results from a VAR with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, QSPR9, where VOL, RET, and QSPR denote volatility, return, and the quoted spread, respectively, and where the "0" and "9" suffixes respectively denote the deciles corresponding to the smallest and largest firms sorted by market capitalization. The VAR is estimated with two lags, includes a constant term, and uses 3782 observations. Cell *ij* shows chi-square statistics and p-values 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 zero when j is the dependent variable in the VAR. 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 to December 31, 2002. ** denotes significance at the 1% level and * denotes significance at the 5% level.

	VOL0	VOL9	RET0	RET9	QSPR0	QSPR9
VOL0		**38.808****	**38.119****	**10.213****	**32.659****	**19.282****
VOL9	2.123		**12.439****	5.537	**8.868***	**101.944****
RET0	**13.296****	**8.721***		2.891	4.505	2.540
RET9	**9.552****	**18.792****	**39.728****		1.035	**12.910****
QSPR0	**96.854****	0.953	0.314	4.959		**10.568****
QSPR9	**68.278****	**60.968****	4.859	2.176	**11.146****	

Table 4: VAR Results with Interaction Terms, for the Smallest Decile

The table presents results from VARs with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, QSPR9, where VOL, RET, and QSPR denote volatility, return, and the quoted spread, respectively, and where the "0" and "9" suffixes respectively denote the deciles corresponding to the smallest and largest firms sorted by market capitalization. In addition, one lag of the exogenous variables QRET09, QRET99, and QOIB99 are included in the equation for RET0, where QRET09= QSPR0*RET9, QRET99= QSPR9*RET9, 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 VAR is estimated with two lags, include a constant term, and uses 3782 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 to December 31, 2002. The Wald test reports the chi-square statistics for the null hypothesis that the coefficients of all exogenous variables are jointly zero. ** denotes significance at the 1% level and * denotes significance at the 5% level.

	Estimate	t-statistic	Estimate	t-statistic	Estimate	t-statistic
Endogenous variable: RET0						
RET9(-1)	0.088**	6.028	0.059	1.024	0.015	0.255
QRET09(-1)	---	---	-0.214	-0.689	-0.179	-0.578
QRET99(-1)	---	---	0.368*	2.226	0.313	1.892
QOIB99(-1)	---	---	---	---	0.047**	4.317
Wald Test						
Chi-square	---	---	4.994		23.727	
Probability	---	---	0.082		0.000	

Table 5: VAR Results With Interaction Terms, for all deciles

This table presents results from VARs with endogenous variables VOLN, VOL9, RETN, RET9, QSPRN, QSPR9, where N=1 through 8 refers to size deciles, and VOL, RET, and QSPR denote volatility, return, and the quoted spread, respectively. The deciles are numbered in order of increasing size, with the smallest and largest deciles being suffixed respectively by "0" and "9". In addition, one lag of the exogenous variables QRETN9, QRET99, and QOIB99 are included in the equation for RETN, where N=1 through 8, and QRETN9= QSPRN*RET9, QRET99= QSPR9*RET9, 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 the second and third columns RET is measured as the transaction price return. For the fourth and fifth columns, RET is the return based on closing quote midpoints on each day. All VARs are estimated with two lags, include a constant term, and use 3782 observations. 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. The sample spans the period January 4, 1988 to December 31, 2002. ** denotes significance at the 1% level and * denotes significance at the 5% level.

	Coefficient of QOIB99(-1) for transaction price returns		Coefficient of QOIB99(-1) for midpoint returns	
Size decile	Estimate	t-statistic	Estimate	t-statistic
0	0.047***	4.317	0.054***	3.024
1	0.030**	2.932	0.116***	8.196
2	0.042**	4.257	0.061***	3.119
3	0.045**	4.582	0.061***	3.028
4	0.049**	5.228	0.056***	2.837
5	0.041**	4.840	0.080***	4.232
6	0.034**	4.213	0.055***	3.127
7	0.022**	2.900	0.025	1.497
8	0.020**	3.245	0.047***	3.441

Table 6: VAR Results for the Smallest Decile, With Interaction Terms for Days Around Macroeconomic Announcements

The table presents results from VARs with endogenous variables VOL0, VOL9, RET0, RET9, QSPR0, QSPR9, where VOL, RET, and QSPR denote volatility, return, and the quoted spread, respectively, and where the “0” and “9” suffixes respectively denote the deciles corresponding to the smallest and largest firms sorted by market capitalization. The variables QRET09, QRET99, and QOIB99 are defined QRET09= QSPR0*RET9, QRET99= QSPR9*RET9, and QOIB99= QSPR9*OIB9, where 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. Interactions of these variables with dummies for days surrounding macroeconomic announcements are included as exogenous variables in the equation for RET0. The VAR is estimated with two lags, include a constant term, and uses 3782 observations. 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. RET is the decile return and VOL is the return volatility. The sample spans the period January 4, 1988 to December 31, 2002. The Wald test reports the chi-square statistics for the null hypothesis that the coefficients of the last two variables below are equal to each other. ** denotes significance at the 1% level and * denotes significance at the 5% level.

	Macro=1 for 2 days **including and after** **n**onfarm employment and CPI announcement days		Macro=1 for 2 days **before** nonfarm employment and CPI announcement days	
	Estimate	t-statistic	Estimate	t-statistic
Endogenous variable: RET0				
RET9(-1) *Macro	0.032	0.168	0.038	0.227
RET9(-1) *(1-Macro)	0.038	0.476	0.029	0.354
QRET09(-1)*Macro	0.333	0.368	-1.044	-1.186
QRET09(-1)*(1-Macro)	-0.314	-0.748	-0.036	-0.085
QRET99(-1) *Macro	-0.426	-0.696	0.893	1.769
QRET99(-1) *(1-Macro)	0.268	1.268	0.063	0.290
QOIB99(-1)*Macro (**D1**)	0.084**	3.301	0.061**	2.622
QOIB99(-1)*(1- Macro) (**D2**)	0.063**	4.246	0.068**	4.529
Wald Test (**D1=D2**)				
Chi-square	23.619		22.669	
Probability	0.000		0.000	

Table 7: VARs Including Nasdaq Stocks

This table presents the results from a VAR with endogenous variables VOLN, VOL9, RETN, RET9, QSPRN, QSPR9 where VOL, RET, and QSPR denote volatility, return, and the quoted spread, respectively, and the suffixes "N" and "9" respectively correspond to Nasdaq stocks and the largest firm decile of NYSE stocks sorted by market capitalization. The VAR uses 3782 observations. Panel A of the table presents causality results from the VAR. Cell *ij* shows chi-square statistics and p-values 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 zero when *j* is the dependent variable in the VAR. In Panel B one lag of the exogenous variables QRETN9, QRET99, and QOIB99 are included in the equation for RETN, and QRETN9= QSPRN*RET9, QRET99= QSPR9*RET9, 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 January 4, 1988 to December 31, 2002. The Wald test reports the chi-square statistics for the null hypothesis that the coefficients of all exogenous variables are jointly zero. ** denotes significance at the 1% level and * denotes significance at the 5% level.

Panel A: Granger Causality Results

	VOLN	VOL9	RETN	RET9	QSPRN	QSPR9
VOLN		**18.951****	3.296	2.407	**11.291***	4.513
VOL9	4.720		7.719	**8.504***	0.310	**96.061****
RETN	**19.960****	6.282		2.863	5.109	0.912
RET9	2.079	**13.014****	0.321		7.173	**11.824****
QSPRN	**23.513****	**10.091***	5.416	2.248		5.479
QSPR9	**37.964****	**55.411****	3.833	0.499	**18.562****	

Panel B: VAR Results With Interaction Terms

	Estimate	t-statistic
Endogenous variable: RETN		
RET9(-1)	-0.066	-1.326
QRETN9(-1)	0.162	0.984
QRET99(-1)	-0.074	-0.217
QOIB99(-1)	0.077**	2.799
Wald Test		
Chi-square	10.815	
Probability	0.013	

Figure 1

Quoted bid-ask spreads for NYSE stocks, 1988 to 2002. QSPR9 is the spread for the largest size decile, and QSPR0 that for the smallest one.

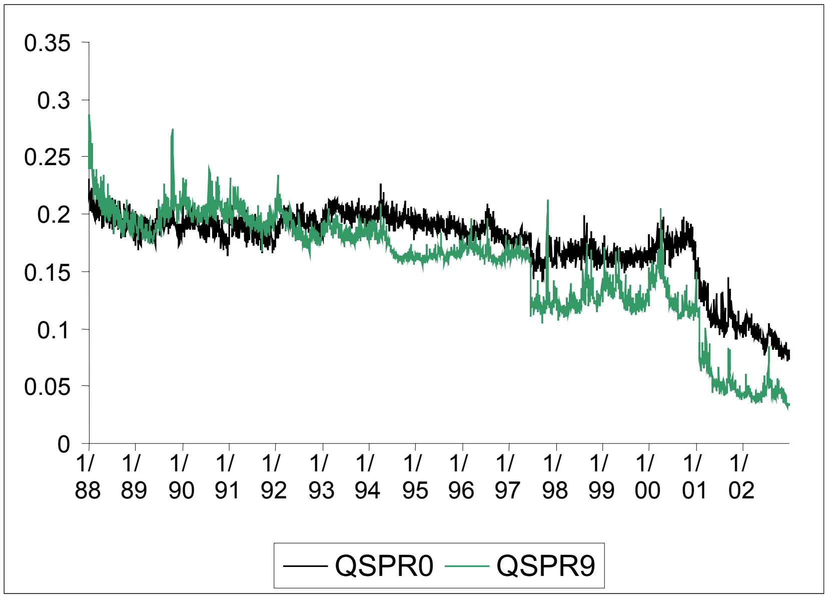
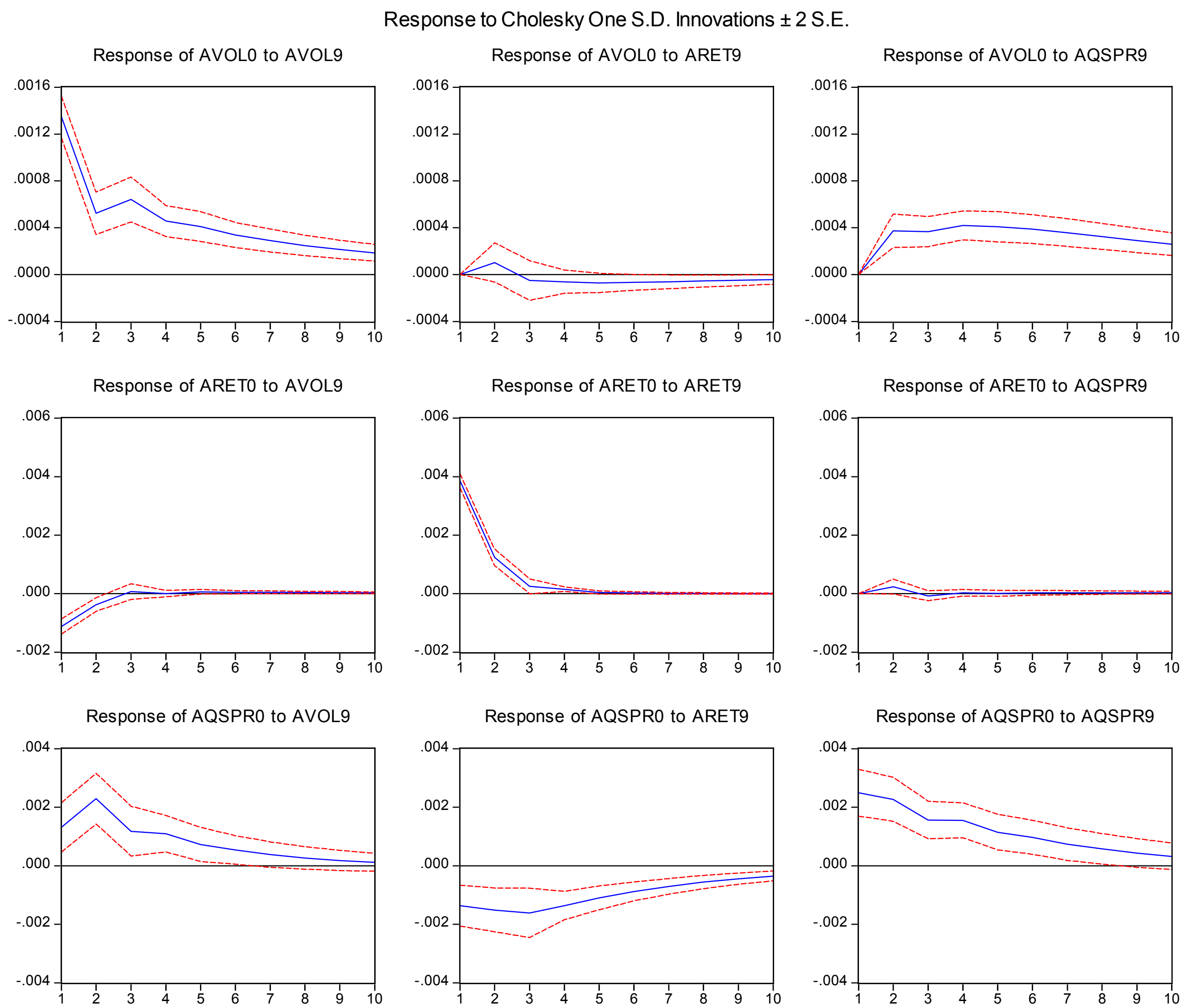


Figure 2. Impulse Response Functions

The figure presents impulse response functions from the VARs with endogenous variables representing volatility (VOL), returns (RET) and quoted bid-ask spreads (QSPR). The ordering is VOL0, VOL9, RET0, RET9, QSPR0, QSPR9, with the smallest and largest deciles being respectively denoted by the suffixes "0" and "9".

Panel A. Response of Decile 0 to Decile 9



Panel B. Response of Decile 9 to Decile 0

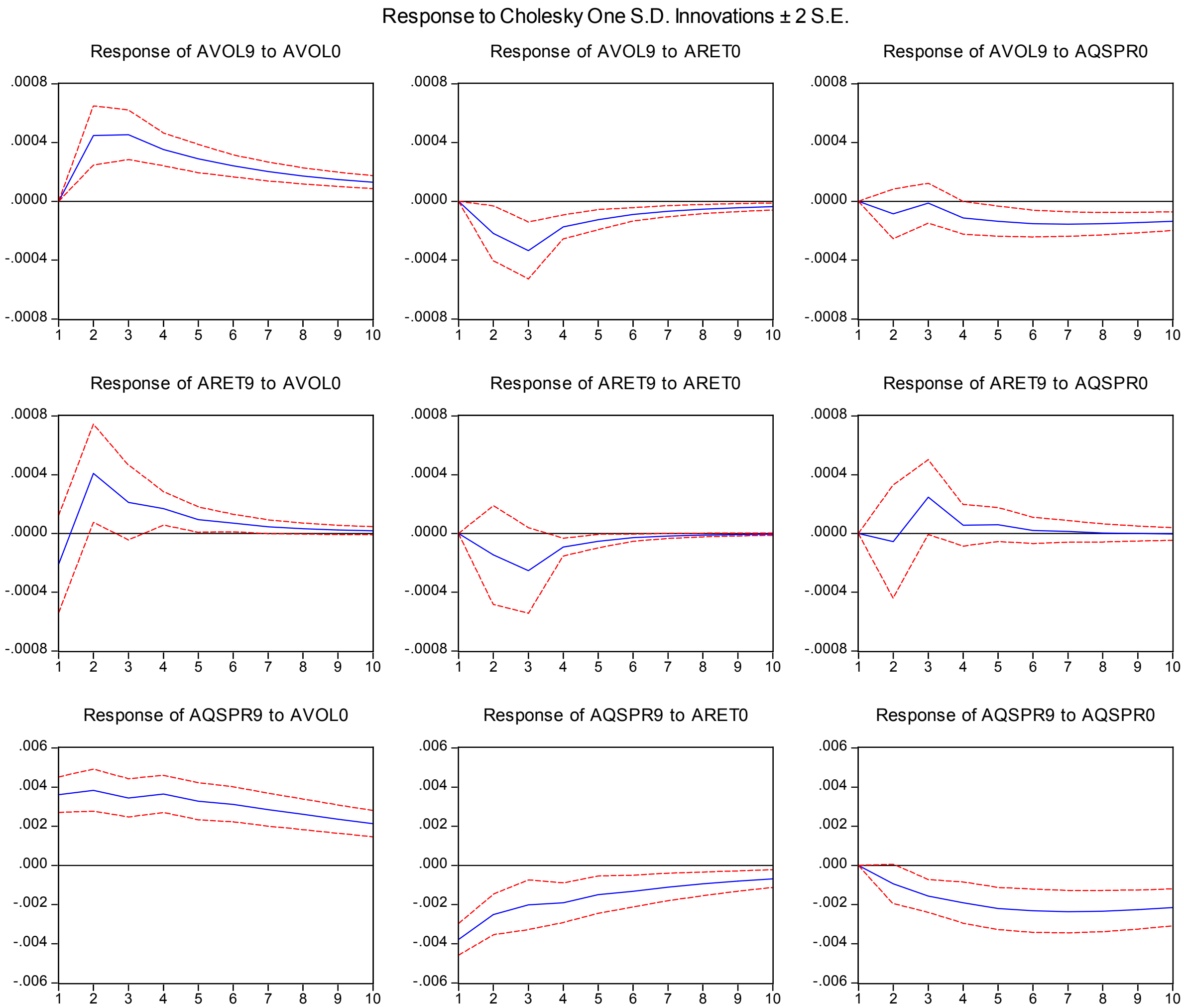


Figure 2, continued

Panel B. Response of Decile 0 to Decile 9