Information Asymmetry about the Firm and the Permanent Price Impact of Trades: Is there a Connection?

Gideon Saar Lei Yu[1]

First Draft: July 2001
This Version: October 2001

[1]Both authors are from the Stern School of Business, New York University, 44 West Fourth Street, Suite 9-190, New York, NY 10012-1126. Saar can be reached at Tel: 212-998-0318, Fax: 212-995-4233, E-mail: gsaar@stern.nyu.edu, and Yu can be reached at Tel: 212-998-0321, Fax: 212-995-4233, E-mail: lyu@stern.nyu.edu. The authors thank I/B/E/S for providing the data on analysts' earnings forecasts, Frank Russell Company for data on the 1999 reconstitution of the Russell 1000 index, and Joel Hasbrouck and Soeren Hvidkjaer for helpful comments.

Abstract

Spread decomposition and variance decomposition methodologies have been developed and used in the literature to obtain measures of information asymmetry about firms. We examine the relation between these market microstructure measures and information asymmetry about the future cash flows of firms. First, to test whether differences in information asymmetry are sufficient to generate differences in the estimated measures, we examine a large cross-section of stocks employing various proxies for uncertainty about future cash flows or informativeness of prices. We find that the market microstructure measures do not consistently reflect uncertainty about future cash flows or relate to the informativeness of prices in a manner that is compatible with their use as proxies for information asymmetry. Second, to test whether changes in information asymmetry about the firm are necessary for the estimated measures to change, we conduct an event study of the Russell 1000 index reconstitution. We find that the information asymmetry measures change around the event despite the fact that Russell 1000 membership is based on market capitalization and therefore the event is not associated with any change in private information about the firms.

One of the heavily used paradigms in Finance is that of information asymmetry among investors (or between investors and managers) about the future cash flows of a firm. A problem with testing models that postulate information asymmetry is that the extent of information asymmetry among investors (or, for that matter, the amount of trading by informed investors) is usually unobservable. The market microstructure literature proposed a solution to this problem in the form of a number of econometric techniques for estimating the extent of information asymmetry (or the intensity of informed trading) using its implications for the patterns of prices and quotes. These econometric techniques were motivated by theoretical models that explain the price impacts of trades or the existence of the bid-ask spread by means of order processing costs, inventory costs, and information asymmetry among investors as to the future cash flows (or liquidation value) of the firm. While all these causes give rise to a spread, they differ slightly in their implications as to the adjustment patterns of prices and quotes. These differences enabled researchers to use the insights of the models to search for ways to empirically estimate a measure of price impact due to informed trading that can be used as a proxy for information asymmetry.

A number of papers set out to develop methodologies, each using a different set of assumptions, for estimating the "adverse selection" component of the spread or the "permanent" impact of the order flow innovation or the "trade-correlated component" of the random-walk variance of quote midpoint changes—all measures of the degree of information asymmetry among investors (see, for example, Glosten, 1987; Glosten and Harris, 1988; Hasbrouck, 1988; Stoll, 1989; George, Kaul, and Nimalendran, 1991; Hasbrouck, 1991a; Hasbrouck, 1991b; Madhavan and Smidt, 1991; Huang and Stoll, 1997; Madhavan, Richardson, and Roomans, 1997).[1] The theoretical models that motivated these econometric techniques have never been tested by independently identifying information asymmetry about the firm and

[1] Easley, Kiefer and O'Hara (1997a, 1997b) estimate a structural model that does not utilize price patterns to estimate the probability of informed trade from the daily number of buy and sell orders. While the current paper focuses on price-related methodologies, the conceptual issues raised here apply in general when trade data is used to characterize information asymmetry among investors.

then looking to see if it generates the predicted implications. Nonetheless, the idea that trading on private information about the future cash flows of a firm should have permanent impact on prices seemed intuitive enough to warrant using these techniques without attempting to test them empirically.

While trading on private information about the firm may very well cause permanent price impacts, it may not be the only reason for them. Saar (2000a, 2000b) provides a framework where all investors have the same information about the future cash flows of the firm, but trades can still have permanent price impacts. The driving force behind his model is uncertainty about the preferences and endowments of the investor population. When investors are risk averse, the preferences and endowments of investors join the future cash flows of the asset to determine equilibrium prices. The price impact of trades is created when prices adjust to reflect information that the market learns about characteristics of investors that have nothing to do with the firm. Saar (2000a) simulates the model and applies a spread decomposition procedure to the simulated data. Over 90% of the spread created in the model is attributed by the procedure to the "adverse selection" or permanent component despite the fact that all investors in the model have the same information about the firm.

Testing the investor uncertainty explanation for the permanent price impact of trades may not be straightforward, as quantifying the degree of uncertainty about the preferences and endowments of investors is perhaps even more difficult than identifying information asymmetry about the firm. However, if the econometric methodologies can pick up both information asymmetry about the firm and uncertainty about investors, interpretation of the estimates becomes difficult and using them to characterize the information environment may be ineffective.[2] The extensive usage of these methodologies to characterize the degree of information asymmetry about the firm necessitates examining the question whether the permanent price impact of trades estimated using these techniques really comes from private

[2] While uncertainty about the investor population can conceptually include uncertainty about investors' knowledge of the future cash flows, we use the term "investor uncertainty" do denote uncertainty about the preferences and endowments of investors. Differences among investors with respect to information about the firm are called "information asymmetry."

information about the prospects of firms. Our objective is to investigate this question.

We divide the investigation into "sufficient" and "necessary" parts. The "sufficient" part examines whether differences in information asymmetry about the firm's future cash flows are sufficient to produce differences in the information asymmetry measures estimated using these methodologies. We conduct this investigation by cross-sectionally relating these measures to estimates of uncertainty about the firms' future cash flows and price informativeness, employing various controls and econometric specifications. The "necessary" part examines whether changes in information asymmetry about the firm's future cash flows are necessary for the information asymmetry measures to change (holding constant the amount of uninformed trading). We analyze this issue using an event study where we look at changes in the information asymmetry measures around an event where we know there is no change in information about the firm: the Russell 1000 index reconstitution. Membership in the Russell 1000 is determined solely based on market capitalization at a single point in time (May 31), and therefore does not add information to what the market already knows.

We use two representative methods to estimate the information asymmetry measures. The first method, from Madhavan, Richardson, and Roomans (1997), provides an estimate of the permanent impact of the order flow innovation (MRR). The second method, from Hasbrouck (1991b), is a decomposition of the efficient price variance that provides us with two measures: (i) the trade-correlated efficient price variance component that serves as an "absolute" measure of information asymmetry (HASAB), and (ii) the ratio of the trade-correlated component to the total efficient price variance that serves as a "relative" measure of the amount of private information to the total amount of information (HASR).

For the cross-sectional part of the study, we need to quantify information asymmetry about future cash flows outside of these econometric techniques. We start by using uncertainty about future cash flows to represent information asymmetry since the greater the uncertainty, the more room (and incentives) there is for investors to acquire private information and trade on it profitably. Two of our proxies for uncertainty about future cash flows are based on dispersion in analysts' earnings forecasts, and the other two are based on past earnings variability. We find that only HASAB (the trade-correlated component of the efficient price variance) is related to both analysts' forecasts and past earnings proxies for uncertainty about future cash flows in the predicted direction. HASR is not related to any of the proxies for future cash flows uncertainty, and MRR provides conflicting results that are even more puzzling since it is negatively related to dispersion in analysts' forecasts.

We continue the "sufficient" part of the investigation with a set of regressions of current returns on future changes in earnings that we use to quantify how much of the information about future earnings is impounded into current prices. *Ceteris paribus*, the more trading on private information about the firm, the more informative should prices be with respect to future earnings (see, for example, Grossman and Stiglitz, 1980). We find that none of the information asymmetry measures relates to price informativeness in the predicted manner.

In the event study, we examine how the information asymmetry measures change from the last trading week in May (before the new ranking of the Russell 1000 is established) to the first trading week in July (when the new index takes effect). We find that MRR and HASAB decrease significantly, and show that the decrease cannot be attributed just to changes in normal trading around the event. We test the robustness of our findings with respect to alternative estimation intervals, price level effects, and possible changes in market-wide liquidity, and find that these do not change our conclusions.

The results we present have implications for the vast number of papers that use econometric procedures based on the price impact of trades to characterize the degree of information asymmetry about the firm. In addition to the aforementioned papers that establish the methodologies, other market microstructure papers use these measures to compare between markets (Neal, 1992; Affleck-Graves, Hedge, and Miller, 1994; Porter and Weaver, 1996), examine the importance of trade size (Lin, Sanger, and Booth, 1995; Heflin and Shaw, 2000b), characterize investors' intraday and daily trading patterns (Wei, 1992; Foster and Viswanathan, 1993), and look at the liquidity effects of introducing a futures contact

(Jegadeesh and Subrahmanyam, 1993).

These measures are also used outside of market microstructure to investigate information asymmetry in the contexts of corporate events (Brooks, 1994; Jennings, 1994; Barclay and Dunbar, 1996; Krinsky and Lee, 1996; Desai, Nimalendran, and Venkataraman, 1998), the debt-equity mix (Illesy and Shastri, 2000), corporate diversification (Fee and Thomas, 1999), disclosure quality (Heflin, Shaw, and Wild, 2000), the opaqueness of banking firms' assets (Flannery, Kwan, and Nimalendran, 2000), the size of the analyst following (Brennan and Subrahmanyam, 1995), and ownership structure (Heflin and Shaw, 2000a; Sarin, Shastri, and Shastri, 2000, Dennis and Weston, 2001; Dey and Radhakrishna, 2001).

Three papers in particular are related to our investigation: Neal and Wheatley (1998) and Clarke and Shastri (2000, 2001). Neal and Wheatley use econometric spread decomposition procedures to examine 17 closed-end funds. They postulate that since the net asset value of each fund is reported every week, there can be little information asymmetry about the current liquidation value, and therefore very little by means of an adverse selection component. Still, they find large adverse selection estimates for the closed-end funds in their sample. They conclude that either the estimates of the adverse selection components are unreliable, or that they reflect something other than information asymmetry about the current liquidation value. Clarke and Shastri (2001) examine a larger cross-section of closed-end funds and reach different conclusions. In particular, they argue that uncertainty about private benefits paid to blockholders lead to higher adverse selection costs, and document a positive relation between blockholdings and the adverse selection component. While the question Neal and Wheatley raise is in the spirit of our own motivation, the contrasting results using closed-end funds seem to suggest that a more fruitful approach could be a direct analysis of equities. Our focus on a large cross-section of equities and the investigation of the event study allow us to reach stronger conclusions on the relation between future cash flow information and both spread decomposition and variance decomposition measures.

Clarke and Shastri (2000) use six econometric methodologies to estimate information asymmetry measures for 320 firms. They look at how these measures relate to each other, and also at the correlations between the measures and certain firm characteristics. The correlations with firm characteristics are statistically significant only for net sales and a dummy for regulated firms. A correlation with analysts' forecast errors is statistically significant only for large firms. While our paper also uses analysts for some of the tests (though not forecast errors but rather dispersion in earnings' forecasts), the focus of their paper is different from ours. They take the estimated measures to represent information asymmetry, and examine whether all of them represent the same information, and whether they are correlated with variables such as the ratio of R&D expenses to operating expenses or sales growth. We, on the other hand, design the investigation to test whether the econometric measures indeed provide useful depictions of information asymmetry about future cash flows of the firms.

Our paper is also related to the literature that investigates what causes or explains price movements (or equivalently, total return variation). This literature usually focuses on a very different horizon from ours (e.g., monthly returns) and examines whether return variation is caused by shocks to expected future cash flows of the assets or by shocks to future expected returns (the "discount rate" effect). This distinction is similar in nature to the difference between attributing the permanent price impact of trades to news about future cash flows (e.g., Glosten and Milgrom, 1985; Kyle , 1985) or to uncertainty about the preferences and endowments of investors that affects the risk premium (Saar, 2000a, 2000b).[3] Fama (1990), for example, finds that the discount rate effect accounts for about 30% of the variance of annual real returns on a value-weighted portfolio of NYSE stocks. Campbell and Ammer

[3]In the classical CAPM, investors' preferences and endowments determine a risk premium that is a property of the market. In such a simple framework, order flow in any stock should provide information about the pricing of all stocks. However, Levy (1978), Merton (1987) and Shapiro (2001) develop models where investors do not invest in all assets in the economy. These papers formalize the "investor recognition hypothesis" whereby investors only trade stocks on which they know something. The expected return of each stock in these models depends on the characteristics of its specific clientele of investors. Such a framework allows information about investors to affect the value of a single stock or a subset of stocks without affecting all stocks in the market. The "investor recognition hypothesis" introduces yet another reason for uncertainty about the characteristics of the investor population—"awareness" of a stock. When awareness changes (without any private information about future cash flows), the aggregation of preferences and endowments of those investing in the stock changes as well. Therefore, learning about "awareness" of investors can generate permanent price impacts.

(1993) use a VAR decomposition of monthly stock excess returns and find that only 15% of the variance of returns is attributed to the variance of news about future dividends, while 70% is attributed to the variance of news about future excess returns.

Our different focus is reflected in both the short horizon we examine—trade by trade—and in the fact that we do not try to quantify the contribution of either explanation (information asymmetry about future cash flows or investor uncertainty) to the permanent price impact for trades. We believe that, as the aforementioned papers found for the long horizon, both information asymmetry about future cash flows and uncertainty about investor characteristics that are unrelated to the firm contribute to price changes. The question we attempt to answer in this paper is whether estimates of the permanent price impact of the order flow produced by spread decomposition and random-walk variance decomposition procedures are useful for identifying changes or differences in the information asymmetry about future cash flows.

The rest of the paper proceeds as follows. Section 1 describes the information asymmetry measures we use and the design of the tests. Section 2 presents the cross-sectional analysis that relates the information asymmetry measures to proxies for uncertainty about future cash flows. Section 3 investigates the relation between the information asymmetry measures and price informativeness. Section 4 examines changes in the information asymmetry measures around the Russell 1000 index reconstitution. Section 5 concludes with a discussion of the results.

1 Information Asymmetry Measures

There are several methods for estimating the "permanent" impact of the order flow innovation or the "adverse selection" component of the spread using trade-indicator models.[4] As a representative of these techniques we take Madhavan, Richardson, and Roomans (1997). Let x_t denote an indicator variable taking the value of 1 if the transaction at time t is buyer

[4] See Huang and Stoll (1997) for a discussion of the relationships among these methods.

initiated and -1 if it is seller initiated, and let μ_t denote the post-trade expected value of a stock.[5] The revision of beliefs following a trade is the sum of the change in beliefs due to public information and change in beliefs due to the order flow innovation:

$$\mu_t = \mu_{t-1} + \theta \left(x_t - E[x_t | x_{t-1}] \right) + \epsilon_t \tag{1}$$

where θ is the permanent impact of the order flow innovation and is a measure of the degree of information asymmetry, and ϵ_t is the innovation in beliefs between times $t-1$ and t due to public information. Let p_t denote the transaction price at time t, and ϕ denote the market makers' cost per share of supplying liquidity (compensating them for order processing costs, inventory costs, and so on). The transaction price can then be expressed as:

$$p_t = \mu_t + \phi x_t + \xi_t \tag{2}$$

where ξ_t captures the effects of stochastic rounding errors induced by price discreteness or possibly time-varying returns.

Equations (1) and (2) can be used to obtain:

$$u_t = p_t - p_{t-1} - (\phi + \theta) x_t + (\phi + \rho\theta) x_{t-1} \tag{3}$$

where ρ is the first-order autocorrelation of the trade initiation variable. Then, the measure of information asymmetry θ, alongside ϕ, ρ, λ (the unconditional probability that a transaction occurs within the quoted spread), and a constant α can be estimated using GMM applied to the following moment conditions:

$$E \begin{pmatrix} x_t x_{t-1} - x_{t-1}^2 \rho \\ |x_t| - (1 - \lambda) \\ u_t - \alpha \\ (u_t - \alpha) x_t \\ (u_t - \alpha) x_{t-1} \end{pmatrix} = 0 \tag{4}$$

We use data from the TAQ database to estimate (4).[6] Following Madhavan *et al.*, the

[5] Madhavan *et al.* note that some trades (such as those with prices between the bid and the ask) can be viewed as both buyer and seller initiated, and set for those $x_t = 0$.

[6] We apply various filters to clean the data. We eliminate trades with irregular settlement or with prices that seem unreasonable relative to the previous trade or the prevailing quote. We also eliminate quotes with halt or fast trading conditions, and quotes where the bid is greater than the ask or that have unreasonable spreads.

classification into buys and sells is done as follows: (i) if a trade price is greater than or equal to the prevailing ask then $x_t = 1$, (ii) if the trade price is less than or equal to the bid then $x_t = -1$, and (iii) if the trade price falls between the bid and the ask then $x_t = 0$. We use the log transaction price for p_t, and refer to the estimate of θ as the MRR measure.[7]

The second methodology that we employ to estimate measures of information asymmetry based on the price changes induced by trades is the variance decomposition of Hasbrouck (1991b). In his specification, the quote midpoint q_t is the sum of two unobservable components:

$$q_t = m_t + s_t \tag{5}$$

where m_t is the efficient price (i.e., expected end-of-trading security value conditional on all time-t public information) and s_t is a residual discrepancy term that is assumed to incorporate inventory control, price discreteness, and other influences that cause the midquote to deviate from the efficient price. The efficient price evolves as a random walk:

$$m_t = m_{t-1} + w_t \tag{6}$$

where the innovation w_t reflects updates to the public information set including the information in the order flow.

The market's signal of private information is the current trade innovation defined as $x_t - E[x_t|\Phi_{t-1}]$, where Φ_{t-1} is the public information set prior to the trade. The impact of the trade innovation on the efficient price innovation is $E\left[w_t\,|x_t - E[x_t|\Phi_{t-1}]\right]$. Two measures of information asymmetry that Hasbrouck proposes are:

$$\sigma^2_{w,x} \equiv \mathrm{Var}\left(E\left[w_t\,|x_t - E[x_t|\Phi_{t-1}]\right]\right) \tag{7}$$

$$R^2_w \equiv \frac{\sigma^2_{w,x}}{\sigma^2_w} = \frac{\mathrm{Var}\left(E\left[w_t\,|x_t - E[x_t|\Phi_{t-1}]\right]\right)}{\mathrm{Var}(w_t)} \tag{8}$$

[7] Madhavan *et al.* estimate (4) using the transaction price, rather than its logarithmic transformation, and therefore the components they estimate are in dollar terms (i.e., components of the dollar spread). In the cross-sectional part of the paper (both in Section 2 and Section 3), We use log transaction prices to get components in percentage terms (i.e., components of the relative spread), since those seem to be better suited for comparisons across stocks. In the event study (Section 4), we estimate two versions of θ: one using dollar prices and the other using log prices. The results from both specifications are similar.

where $\sigma^2_{w,x}$ is the trade-correlated component of the random-walk variance of quote midpoint changes and is an absolute measure of information asymmetry, while R^2_w is a measure of the amount of private information relative to the total amount of information.

Let x^0_t be an indicator variable that takes the values $\{-1, 0, +1\}$, and define: $x^1_t = x_t$ (the signed volume), $x^2_t = \text{sign}(x_t)(x_t)^2$ (a signed, quadratic volume variable), and $\boldsymbol{x'_t} = \{x^0_t, x^1_t, x^2_t\}$. Let $r_t = \log q_t - \log q_{t-1}$. Following Hasbrouck, the absolute and relative information asymmetry measures are estimated using the vector autoregressive (VAR) model:

$$r_t = \sum_{i=1}^{5} a_i r_{t-i} + \sum_{i=0}^{5} \boldsymbol{b'_i x_{t-i}} + v_{1,t} \tag{9}$$

$$\boldsymbol{x_t} = \sum_{i=1}^{5} c_i r_{t-i} + \sum_{i=1}^{5} \boldsymbol{D_i x_{t-i}} + \boldsymbol{v_{2,t}} \tag{10}$$

where $\boldsymbol{b_i}$ is a 3×1 vector of coefficients, $\boldsymbol{D_i}$ is a 3×3 matrix of coefficients, and $\boldsymbol{v_{2,t}}$ is a 3×1 vector of error terms.

We use data from the TAQ database to estimate the VAR model in (9) and (10). Following Hasbrouck, the trade classification into buys and sells is done as follows: (i) if a trade price is greater than the prevailing midquote then $x^0_t = 1$, (ii) if the trade price is less than the prevailing midquote then $x^0_t = -1$, and (iii) if the trade price is at the prevailing midquote then $x^0_t = 0$.[8] We estimate the VAR system using OLS. A finite VAR generally possesses a vector moving average (VMA) representation of an infinite order. The random walk variance and the trade-correlated component are calculated from the coefficients of the VMA representation truncated after 100 lags.[9] Throughout the paper, we will refer to the estimate of the absolute measure, expressed in the form of standard deviation per hour, as the HASAB measure, and to the estimate of the relative measure as the HASR measure.

[8]Hasbrouck considers two types of event: (i) a trade, or (ii) a quote change without a trade in the previous 5 seconds. Each event has an associated r_t and $\boldsymbol{x_t}$. The return associated with a trade event is the difference between the prevailing log midquote at the time of the trade and the log midquote subsequent to the trade. If there is a change of quote in the 5 seconds after a trade, it is considered to be brought about by the trade and the return associated with that trade is the change in log midquotes. If there is no quote change in the five seconds after the trade, the prevailing midquote continues to be in effect, and the return associated with the trade is zero. Whenever a quote changes without a trade in the previous 5 seconds, there is a quote change event, and it is considered a change in prices due to public information. In this case, the return is the change in log midquotes and $\boldsymbol{x_t}$=$\mathbf{0}$.

[9]See Hasbrouck (1991b) for a detailed exposition of the methodology.

2 Cross-Sectional Analysis

2.1 Sample

The initial sample for the cross-sectional investigation is constructed by including all common stocks with information in the CRSP database for the entire year of 1999 that had more than one analyst covering them in the I/B/E/S database. The reason for requiring more than one analyst is to allow the calculation of the standard deviation of analysts' earnings forecasts. This screen leaves 3,144 stocks. The sample is further restricted to firms with information in TAQ, COMPUSTAT, and Value Line Investment Survey in order to allow the computation of the information asymmetry measures and the construction of other variables that will be explained in greater detail below. This requirement eliminates 268 firms from the sample, leaving a final sample of 2876 firms.

Table 1 presents summary statistics for the sample. The average daily market capitalization (AvgCap) over the sample period (January 4 through December 31, 1999) for firms in our sample ranges from \$6.05M to \$446.16B, testifying to its heterogeneous nature. The sample also spans a range of trading activity and price levels. The most active firm has on average 29,809.45 daily trades (AvgTrd), while the average firm has 442.67 daily trades, and the least actively traded firm in the sample has 3.03 trades per day. Average daily closing prices (AvgPrc) range from \$0.24 to \$550.05. The sample is also diverse with respect to analyst following and institutional holdings. NumEst is the average of 12 monthly observations (January through December, 1999) of the number of analysts with end-of-fiscal-year earnings forecast for the current year in I/B/E/S. While the mean of NumEst is 8.81, it ranges from a minimum of 2 to a maximum of 40.18. The percentage institutional holdings from Value Line Investment Survey (InstHol) ranges from 0.23% to 99.59%.

Table 2 presents summary statistics for the MRR, HASAB, and HASR measures of information asymmetry. The mean MRR measure, or the permanent component of the relative spread, is 0.07368 percent. The mean daily closing prices (AvgPrc) for the stocks in our sample is \$26.68, which means that the permanent component of the dollar spread for the average stock is about 2 cents.[10] The estimates of MRR for seven stocks are negative. While there is nothing in the estimation to constrain the information asymmetry measure from being negative, it is unclear how to interpret these estimates. The negative measures do not affect our results as our findings are practically identical with and without these seven stocks.

The mean of the HASAB measure in our sample is 0.391; it decreases with AvgCap and increases with StdRet.[11] HASAB also seems to increase, although not monotonically, with MedTurn (the median daily turnover of the stock). This is perhaps driven by the way the HASAB measure, expressed in standard deviation per hour, is calculated: it is the square root of the estimated variance per event multiplied by the number of events over the sample period and divided by the number of hours. Since many of the events are trades, it is not surprising that the measure increases with MedTurn.[12]

2.2 Methodology

The first test we run examines whether cross-sectional differences in MRR, HASAB, and HASR reflect differences in information asymmetry about future cash flows. Since we do not have a way to directly measure information asymmetry (which is one of the reasons econometric methodologies that use trade data became so popular in the first place), we focus instead on uncertainty about future cash flows. The logic behind the test is that the greater the uncertainty about future cash flow, the more room (and incentives) there is for investors to acquire private information and trade on it profitably. A theoretical treatment

[10] Table 2 in Madhavan *et al.* (1997) shows a mean measure of about 3 cents. It seems reasonable that our estimates are lower considering that Madhavan *et al.* use 1990 data while our sample period is 1999.

[11] The VMA could not converge for eight stocks in our sample: ADOG, FOBB, NRG, OAK, OMNI, RIT, SHLR, and YANB. They are eliminated from the sample for the tests involving the HASAB and HASR measures

[12] The mean of the absolute measure that we report (0.391) is very similar to the one reported by Hasbrouck (0.412). The mean of our relative measure is 16.17%, which is about half of the mean reported by Hasbrouck. While his sample consisted of NYSE/AMEX stocks in the third quarter of 1989, our sample includes both NYSE/AMEX and Nasdaq stocks in 1999. For the 1413 Nasdaq stocks in our sample, the estimate of HASR is 10.19%, while for the NYSE/AMEX stocks it is 21.90%, which is closer to the magnitude reported by Hasbrouck. Our results indicate that the magnitude of the trade-correlated component remained in 1999 as it was in 1989, while the proportion of the variation due to trades decreased relative to the overall variability of the efficient price.

of the information environment of analysts is provided by Barron, Kim, Lim, and Stevens (1998). They show that an increase in private information about the firm, *ceteris paribus*, will increase the dispersion of analysts' forecasts.

We use two definitions of dispersion in analysts' forecasts of future earnings to represent uncertainty about future cash flows: the standard deviation of analysts' forecasts (StdEst) and the coefficient of variation (CoefEst), which a-priori has the appeal of a normalized measure that may perform better in a cross-sectional analysis.[13] Monthly observations for the standard deviation and the mean of analysts' earnings forecasts for the upcoming end-of-fiscal-year are taken from I/B/E/S. StdEst is the simple average of the monthly standard deviation observations for all months in 1999, and CoefEst is the ratio of StdEst to the absolute value of the average monthly mean observations. The monthly observations of the standard deviation of the forecasts in I/B/E/S (as well as the number of estimates) seem to depend on the amount of time to the end-of-fiscal-year. However, since we use twelve consecutive months of data, the average should have similar properties across firms with different end-of-fiscal-year dates.

We also use two proxies for uncertainty about future cash flows based on past earnings information. The first is the earning predictability measure provided by Value Line Investment Survey (VLPred). This variable, available for 2081 stocks in our sample, is derived from the standard deviation of percentage changes in quarterly earnings over an eight-year period, with special adjustments for negative observations and observations around zero. VLPred takes the values 1 to 100, where higher values are associated with less variability of past earnings and hence higher predictability of future earnings. We also compute the standard deviation of past annual earnings from COMPUSTAT using the five years prior to 1999 (StdEPS).[14] The idea behind both measures is that higher past earnings variability

[13]We also used the difference between the high and the low forecasts as an additional definition of uncertainty about future cash flow. The results using this definition were very similar to the results using the standard deviation of earnings' forecasts, and are therefore omitted for brevity.

[14]We repeated the analysis with the standard deviation of annual earnings computed using five years that include 1999. The results were very similar to those obtained using StdEPS.

makes it more likely that future earnings cannot be predicted with great precision (allowing more room for private information production and profitable informed trading).

We then estimate regressions with the information asymmetry measures as dependent variables and the proxies for uncertainty about future cash flows as independent variables. This econometric specification attests to an underlying assumption that we make when using the analysts' forecasts and past earnings proxies in the cross-sectional test. Intuitively, variables that explicitly incorporate information about the firm's earnings (or their expectations) are a-priori more tightly connected to the firm's future cash flows than measures that do not use such direct information (like the estimated price impact measures), and therefore should be the independent variables in the regressions.

We estimate these regressions using OLS with White's Heteroskedasticity-consistent standard errors. A potential econometric problem with this specification for the two proxies that use analysts' forecasts is that the dispersion in forecasts can be endogenous.[15] In other words, the amount of time analysts spend on producing their estimates (which may affect the dispersion), the number of estimates or the frequency of their update may be influenced by how much profit can be made from this information. The profit opportunity, in turn, depends on the permanent price impact of trades—the same permanent price impact that the econometric methodologies estimate and we use as a dependent variable. This rationale suggests that the dispersion of analysts' forecasts may be correlated with the error term, causing OLS to be inconsistent. To examine this potential problem, we also estimate the cross-sectional relation using two-stage-least-squares (2SLS) and conduct Durbin-Wu-Hausman tests.[16]

[15] This is not a problem with the proxies that use actual earnings data since it is hard to imagine how the permanent price impact of trades can affect the firms' earnings per share.

[16] See Wu (1973), Hausman (1978), and Davidson and MacKinnon (1989) for a discussion of this test. Using data in COMPUSTAT and CRSP, we construct six instruments for the dispersion in analysts' forecasts. The first instrument is the log standard deviation of EPS for the five years prior to 1999 (available for 2845 firms in our sample). The second instrument is the one year lagged EPS growth (percentage change in EPS from 1997 to 1998), using the absolute value of EPS in the denominator to maintain consistent signs when EPS is negative (available for 2822 firms in our sample). The motivation for this instrument is that growth firms are associated with more uncertainty about future earnings. The third instrument is the average absolute value of quarterly sales growth in 1999 from COMPUSTAT (available for 2774 firms in our sample). This is a measure of the magnitude of change in the firm's operations that presumably is correlated with uncertainty about its sales and earnings potential. The fourth instrument is the number of days with listing records

The choice of control variables to add to the regressions is guided by the proxies we use for future cash flows uncertainty. More dispersion among analysts may reflect the nature of the industry or the number of analysts that typically follow firms in a specific industry. Therefore, we add to the regressions a set of 2-digit SIC dummies to control for industry effects. To relate our measures of uncertainty about future cash flows to the information asymmetry measures we need to control for public information, and we (partially) do it by including the standard deviation of daily returns in 1999 from CRSP as a control variable. Similarly, there are more public information events (e.g., news articles) about larger firms. So, we add a second control for public information flow by including AvgCap, the average market capitalization in 1999 from CRSP (in log form).

Two other controls are motivated by the market microstructure literature. First, the impact of informed trading on prices diminishes as the amount of liquidity trading or normal trading increases (see, for example, Kyle, 1985; Easley and O'Hara, 1987). Hence, we want to hold constant the amount of liquidity trading when we compare across firms. As a proxy for normal trading, we use the median daily turnover (daily number of shares traded over the number of shares outstanding) in 1999 from CRSP (MedTurn). The median is less sensitive to informational shocks that generate a lot of trading, and therefore is a better description of "normal" volume.[17] Lastly, we use dummies for the three primary markets (NYSE, AMEX, and Nasdaq) to control for differences in the manner specialists or market makers set prices in these markets. The motivation behind this control is that on the NYSE, for example, the price continuity requirement affects the way in which specialists change prices, and this may be picked up by the spread decomposition and variance decomposition procedures (for

in CRSP, which is a measure of the maturity of the firm. Presumably, it is easier to estimate the future earnings of more mature firms. The fifth instrument is the log number of SIC codes that are associated with the businesses of the firm (available for 2751 firms in our sample). Firms comprised of more businesses may be more difficult for analysts to evaluate, and this should increase the dispersion of the forecasts. The sixth instrument is the dividend payout ratio for 1999 (available for 2685 firms in our sample). The idea behind this instrument is that the more money is paid out, the less money is spent of new projects and therefore the less uncertainty there is about earnings.

[17]We use median turnover in log form, and add a small constant (0.000001) to the raw data before making the transformation to be able to accommodate stocks with zero median turnover.

a discussion of this point see Hasbrouck, 1991a). Table 3 presents summary statistics for the analysts' forecasts and past earnings proxies, as well as correlations among the (log transformed) proxies and control variables that are used in the regressions.

2.3 Results

Panel A of Table 4 presents the results of the cross-sectional regressions with the permanent price impact of the order flow innovation, MRR, as the dependent variable. Of the two proxies using analysts' forecasts, only StdEst is statistically significant, and it is negative. This implies that information asymmetry is negatively related to uncertainty about future cash flows, which does not seem reasonable. As for the two proxies using past earnings, VLPred is negative and (weakly) statistically significant. Here, the direction is consistent with our expectation: higher predictability is associated with lower information asymmetry. It is interesting to note that when StdEst and VLPred are together in the regression, StdEst losses statistical significance and VLPred is strengthened.

Panel B of Table 4 presents the results with the trade-correlated component of the efficient price variance, HASAB, as the dependent variable. Here, StdEst is not significant, but CoefEst is positive and statistically significant. The positive relation between HASAB and uncertainty about future cash flows is in line with its role as an information asymmetry measure. VLPred is (weakly) statistically significant and negative, again consistent with our expectation. Unlike the case with MRR, VLPred losses statistical significance when it is in the regression with the analysts' forecasts measure.[18]

Panel C of Table 4 presents the results using the amount of private information relative to the total amount of information, HASR, as the dependent variable. None of the proxies for uncertainty about future cash flows is statistically significant when in the regression alone. When CoefEst and StdEPS are together in the regression, CoefEst is negative but only marginally significant. This result does not agree with our intuition about the relation

[18]The correlation between StdEst and VLPred is -0.406, and between CoefEst and VLPred is -0.571. These do not seem high enough for multicollinearity to be the sole explanation for the loss of significance of StdEst in the MRR regression and VLPred in the HASAB regression.

between measures of information asymmetry and uncertainty about future cash flows, but seems too weak to be meaningful.[19]

The Durbin-Wu-Hausman exogeneity tests and the coefficients of the two-stage-least-squares (2SLS) estimation are presented in Table 5. For all three measures, the exogeneity tests are insignificant (with very low chi-squared statistics). This gives us confidence that the results of the OLS estimations provide an accurate picture of the relationships among the variables.[20]

To summarize the results of the cross-sectional analysis: the only information asymmetry measure that relates to both types of proxies (analysts' forecasts and past earnings) in the predicted direction is HASAB. MRR provides conflicting results that are difficult to interpret, and HASR is unrelated to any of the proxies.

3 FERC and FINC Analysis

3.1 Methodology

We use the same sample to perform a different kind of cross-sectional analysis with two proxies for the informativeness of prices about future earnings that were developed in the accounting literature (see, Kothari, 1992; Kothari and Sloan, 1992; Collins, Kothari, Shanken, and Sloan, 1994; Durnev, Morck, Yeung, and Zarowin, 2001). Let $r_t = (P_t - P_{t-1})/P_{t-1}$ be the current quarterly return and $r_{t+\tau}$ be the return τ quarters ahead. Similarly, let ΔE_t be the current earnings-per-share change and $\Delta E_{t+\tau}$ be the earnings change τ quarters ahead, where all earnings changes are scaled by the price at the beginning of the current quarter, P_{t-1}. To measure stock price informativeness, we regress current stock returns on current

[19] Since HASR is a proportion, we also estimated the regressions with a logit transformation of HASR as the dependent variable. The results were qualitatively similar to those reported in Panel C and are therefore not presented here.

[20] We also conducted Durbin-Wu-Hausman tests with subsets of the six instruments to see how sensitive are the results to the exact specification of the instruments. Similar results were found for all subsets except in one case. In the regression of MRR on CoefEst, the test was significant when only two instruments were included. Since other instruments were highly significant in the first stage regression (where CoefEst is regressed on the instruments), they were gradually added to the specification. As additional instruments were added, however, the 2SLS coefficient of CoefEst switched sign and the test became insignificant.

and future earnings changes:

$$r_t = a + b_0 \Delta E_t + \sum_\tau b_\tau \Delta E_{t+\tau} + \sum_\tau c_\tau r_{t+\tau} + u_t \tag{11}$$

Future stock returns are included among the regressors to control for the measurement error problem described by Collins *et al.* (1994).

The two proxies for price informativeness are:

$$\text{FERC} \equiv \sum_\tau b_\tau \tag{12}$$

$$\text{FINC} \equiv R^2_{a+b_0\Delta E_t + \sum_\tau b_\tau \Delta E_{t+\tau} + \sum_\tau c_\tau r_{t+\tau} + u_t} - R^2_{a+b_0\Delta E_t + u_t} \tag{13}$$

FERC (Future Earnings Response Coefficients) is defined as the sum of the coefficients on the future earnings changes, and FINC (Future earnings INCremental) is the incremental explanatory power of the future earnings changes (given that current earnings are already in the model). Both FERC and FINC are designed to detect the extent to which current stock prices capitalize information about future earnings.

The general intuition behind using FERC and FINC to test the information asymmetry measures is that the more trading on private information, holding constant the extent of public information, the more informative are prices with respect to future earnings (for a theoretical justification see, for example, Grossman and Stiglitz, 1980). The need to control for public information arises since if we take the extreme case where all future cash flow information is revealed publicly, then both FERC and FINC will be at their maximum but there will be no information asymmetry. However, once we control for public information, the relation between trading on private information and FERC and FINC should be monotonically increasing. This is true in particular since MRR and HASAB (or HASR) are not just measures of information asymmetry, but rather measures of trading based on private information that moves prices. Therefore, these measures directly relate to information incorporation into prices and thus should be more tightly related to FERC and FINC.

Since we are interested in evaluating how much information about future earnings is incorporated into prices during our sample period, we use returns for the four calendar quarters in 1999. Our earnings data is taken from COMPUSTAT, and we have earnings up to and including the second quarter of 2000. To be able to use all four quarterly return observations, we set $\tau = 2$ and estimate the information that is impounded into current prices about earnings two quarters in the future. Since each firm has only four observation, we cannot estimate FERC and FINC for individual firms. We group stocks with similar information asymmetry measures and estimate for each group a common price informativeness proxy. The motivation behind this procedure is that each group of stocks that have the same level of informed trading (holding public information constant) should have similar FERC and FINC estimates, and hence can be estimated together.

For example, to estimate FERC and FINC for the MRR measure, we sort all stocks according to MRR and form 287 groups of ten stocks each.[21] Each such group has approximately 40 firm-quarter observations that can be used for the estimation of FERC and FINC. We estimate Equation (11) using OLS with White's heteroskedasticity-consistent standard errors. We then use the mean of the MRR coefficients of stocks in each group ($\mathrm{MRR_G}$) to represent the level of information asymmetry for the group. $\mathrm{FERC_{MRR}}$ and $\mathrm{FINC_{MRR}}$ are computed according to (12) and (13). We perform separate estimations of Equation (11) for groups formed by sorting on HASAB and HASR. We then use correlations to examine the relationship between the price informativeness proxies and the group measures of information asymmetry.

Ideally, one should control for cross-sectional differences in public information and "noise" trading when looking at the relation between the information asymmetry measures and the proxies for price informativeness. However, since we group stocks according to their information asymmetry measure, we cannot simply use a regression and plug-in control variables. The mean of the market capitalization (or the volatility of returns) of stocks in a group has no real economic content. To exercise some control, we divide the stocks into four categories. First, we sort the 2869 stocks according to their market capitalization.

[21] We dropped the seven stocks for which the MRR measure is negative. Inclusion of these stocks in the groups does not materially affect the results.

We then separate them into two categories, Low AvgCap and High AvgCap, with equal number of stocks. Each category is then sorted according to MedTurn and divided into two equal categories, Low MedTurn and High MedTurn. This procedure produces four control categories: High AvgCap and High MedTurn, High AvgCap and Low MedTurn, Low AvgCap and High MedTurn, and Low AvgCap and Low MedTurn. Within each category, we sort according to the information asymmetry measure, create 72 groups of ten stocks each, and estimate FERC and FINC for the groups. We then compute the correlations separately for each category. Such a structure provides (partial) control for public information and noise trading. We use a similar procedure to control for public information and liquidity trading with StdRet and MedTurn, respectively.

3.2 Results

The results of the FERC and FINC analysis are provided in Table 6. The predicted direction is that more informed trading is associated with a higher degree of price informativeness, which means larger magnitudes of FERC and FINC. The results of the unconditional analysis, where we look at the correlations for the entire sample, seem to go the other way. MRR has a negative correlation with FERC (-0.1003) that is statistically significant; HASAB is negatively correlated with both FERC (-0.1419) and FINC (-0.1647), and both are statistically different from zero; and HASR is not significantly correlated with either FERC or FINC.

The conditional analysis, where we sort the stocks into four categories according to two conditioning variables, should lessen the influence of differences across stocks in the degree of public information or the amount of liquidity trading. Panel B of Table 6 shows the conditional analysis for the MRR measure. When the controls are average market capitalization (AvgCap) and median turnover (MedTurn), MRR is negatively correlated with FERC, but significantly so only for high MedTurn stocks. It is negatively correlated with FINC both for the small, less actively-traded stocks, and for the large, actively-traded stocks. MRR in the two other categories is positively correlated with FINC, but these correlations are not statistically different from zero. When the controls are the standard deviation of daily returns (StdRet) and MedTurn, none of the correlations is statistically different from zero.

Panel C of Table 6 shows the conditional analysis for the HASAB measure. When the controls are AvgCap and MedTurn, only one category (small, actively-traded stocks) has correlations with FERC and FINC that are significantly different from zero, and these correlations are negative. When the controls are StdRet and MedTurn, the only correlation that is statically different from zero (and negative) is in the High StdRet and Low MedTurn category. Panel D of Table 6 presents the conditional analysis for the HASR measure. There are both positive and negative correlations of HASR with FERC and FINC, but none is significantly different from zero.

The bottom line of the FERC and FINC analysis is that despite our expectation that more trading by investors who are informed about future cash flows would increase the informativeness of prices with respect to future earnings, all information asymmetry measures seem to be negatively correlated with the proxies for price informativeness.[22]

4 Event Study: Russell 1000 Reconstitution

In this section we investigate whether changes in information about the firm are necessary for the information asymmetry measures to change, holding constant the amount of liquidity trading. In principle, only one example where the information asymmetry measures change without corresponding changes in information about the firm is required to show that the "necessary" link does not exist. The Russell 1000 index reconstitution provides a suitable

[22] We have used one additional proxy for public information in the conditional analysis of HASAB. The VAR decomposition of the random-walk variance also produces a non-trade-correlated component that can be viewed as an estimate of public information. There is no mechanical relation between the public information component and HASAB (though both components make up the efficient price variance). We sorted the stocks first according to the public information component and then according to MedTurn to get four control categories similar to those that are presented in Panel C where AvgCap or StdRet are used to control for public information. HASAB had both positive and negative correlations with FERC and FINC in the control categories, but none was statistically different from zero.

environment for constructing such a test.[23]

The Russell 1000 index is comprised of the top 1000 firms based on market capitalization. Membership in the Russell 1000 index is determined strictly by market capitalization rather than by subjective opinion or committee decision.[24] While the market capitalization of firms fluctuates with prices, the index is changed only once a year in a "reconstitution" process. All firms are ranked based on their market capitalization at the end of the last trading day in May. The top 1000 firms constitute the new Russell 1000 index, the next 2000 firms the Russell 2000 index, and the firms in both indexes are combined to make the Russell 3000 index. During June, Frank Russell Company issues a list of firms to be added and deleted from the indexes. The new indexes become effective July 1.

The Russell indexes are popular benchmarks of stock market performance in the United States. In 2001, an estimated 117 billion dollars in investment were indexed to the Russell 1000, 2000, and 3000 indexes.[25] Therefore, the annual reconstitution event can generate buying and selling pressures on stocks that are added to or deleted from the indexes. On the other hand, since membership in the Russell 1000 index is solely based on the market capitalization of a firm, which is publicly known, additions to and deletions from the index convey no new information about the firm. Thus controlling for the amount of noise trading, the measures should not change if they arise solely from private information about the future cash flows of the firms. If these measures are affected by other factors, like changes in the characteristics of investors who trade them, we would detect changes around the reconstitution event.

It is clear that the Russell reconstitution event is fairly transparent, and it is possible to

[23]For other papers that examine changes to equity indexes (in particular the S&P 500 index) see Harris and Gurel (1986), Shleifer (1986), Jain (1987), Dhillon and Johnson (1991), and Lynch and Mendenhall (1997).

[24]Excluded from all Russell indexes are foreign firms, limited partnerships, limited liability companies, royalty trusts, closed-ended investment management companies, ADRs, preferred stocks, pink-slipped companies, OTC bulletin board companies, and warrants and rights. The indexes also exclude stocks trading at below $1.

[25]This estimate is taken from a report by Investment Technology Group, Inc. on the 2001 Russell index reconstitution.

predict prior to the last day in May which stocks are more likely to be added or deleted. In fact, firms like Investment Technology Group, Inc. offer the service of predicting the changes in order to assist fund managers who wish to begin rebalancing their portfolios in advance of the announced index changes. Such behavior will bias the results against finding any change between the estimated information asymmetry measures just before the end of May and immediately after July 1. Thus any effect that we detect is likely to be an underestimate of the real change in the measures.

4.1 Sample and Methodology

The sample is constructed from a list, obtained from the Frank Russell Company, of 157 stocks that were added to the Russell 1000 index in 1999 and 103 stocks that were deleted from the index.[26] We restrict the sample to common stocks (for compatibility with the cross-sectional tests) with data from April 1 through mid-July. In order to make sure that no major corporate event occurs at the same time that can confound the results, we use the CRSP database and Dow Jones Interactive to identify firms that were involved in mergers around the time of the Russell reconstitution. We eliminate from the sample firms about which we find news reports discussing the mergers during May, June, and July 1999. The final sample consists of 130 additions and 87 deletions.

Table 7 presents summary statistics on the firms used in the event study. There seems to be a slight increase in mean price between May 28 (the last trading day in May) and July 1 (when the new composition of the index takes effect). Average turnover increases slightly from the last week of May to the first week of July for both samples, but the cross-sectional mean of the average number of trades decreases for the addition sample. Dollar spreads and percentage spreads decrease for both additions and deletions.

Our event study methodology is an adaptation of standard specifications.[27] Let $\phi_{i,\tau}$

[26] The preliminary list of the changes was published by the Frank Russell Company on June 11 and the final list was published on July 7 (though the new composition of the index took effect on July 1). To construct the sample, we eliminate 12 stocks that were not present in both the preliminary and final lists.

[27] See, for example, Schipper and Thompson (1983) and Thompson (1985).

denote the information asymmetry measure for stock i estimated over time interval $\tau \in \{\text{pre}, \text{post}\}$ (i.e., either the pre-event interval or the post-event interval). The simplest specification assumes that the information asymmetry measure can be described as the sum of a stock-specific unconditional mean (μ_i), an event effect (α), and an error term ($\epsilon_{i,\tau}$):

$$\phi_{i,\tau} = \mu_i + \alpha\delta_{i,\tau} + \epsilon_{i,\tau} \tag{14}$$

where $\delta_{i,\tau}$ is an indicator variable that takes the value zero in the pre-event interval and one in the post-event interval, and $\epsilon_{i,\tau}$ is assumed i.i.d and normally distributed with mean zero and variance $\sigma_i^2/2$.

The theoretical models that examine the price impact of trades in the context of information asymmetry also demonstrate how the price impact of trades can be affected by changes to normal or liquidity trading (i.e., volume due to uninformed traders that is unrelated to information events). A formulation that takes normal trading into account can be written as:

$$\phi_{i,\tau} = \mu_i + \alpha\delta_{i,\tau} + \beta V_{i,\tau} + \epsilon_{i,\tau} \tag{15}$$

where $V_{i,\tau}$ is a proxy for the amount of normal volume.

By examining differences between the pre- and post-event intervals, we can eliminate the firm-specific mean. Equation 14 can be written in terms of differences as:

$$\Delta\phi_i \equiv \phi_{i,\text{post}} - \phi_{i,\text{pre}} = \alpha + \epsilon_i \tag{16}$$

where ϵ_i is normally distributed with mean zero and variance σ_i^2. Similarly, the formulation in terms of differences of Equation 15 is:

$$\Delta\phi_i = \alpha + \beta\,\Delta V_i + \epsilon_i \tag{17}$$

where $\Delta V_i \equiv V_{i,\text{post}} - V_{i,\text{pre}}$. Testing that the Russell 1000 reconstitution event affects the information asymmetry measures is therefore equivalent to testing whether the intercept (α) of either Equation 16 or Equation 17 is statistically different from zero.

If the event affects the level of prices but does not affect the information asymmetry environment, it may be preferable to estimate the MRR measure using changes in dollar prices rather than changes to the logarithmic transformation of prices. For consistency with the previous sections, we estimate the same information asymmetry measures, but also present the results of MRR$, which is the permanent price impact of the order flow innovation developed in Madhavan *et al.* (1997) in dollar terms.[28] We use one week as the length of the interval over which the information asymmetry measures are estimated. Since the ranking of firms for the reconstitution is based on the closing prices on May 28, we use the week ending May 28 as the pre-event interval. Similarly, since the effective date of the changes is July 1, we use five trading days beginning July 1 as the post-event interval.[29]

For start, we assume that the error terms are identically distributed across stocks and Equation 16 is estimated as a simple t-test. We also use the non-parametric Wilcoxon signed rank test to perform the analysis under less restrictive assumptions. Then, Equation 17 is estimated using OLS with White's heteroskedasticity-consistent standard errors.

We use two proxies for the amount of normal volume: the average daily turnover and the average daily number of trades in the interval. We fully acknowledge that these may not be perfect controls. The analysis of both the addition and deletion samples, however, provides us with an internal check on the effects of liquidity trading. This, because most stocks in our sample that were added to the Russell 1000 index (109 out of 130) were deleted from the Russell 2000 index, and all stocks that were deleted from the Russell 1000 index were added to the Russell 2000. Say we believe that a move from the Russell 1000 to the Russell 2000 index increases the amount of liquidity trading (perhaps more passive funds

[28]While we conduct a cross-sectional analysis of $\Delta\phi_i$, Equation 16 and Equation 17 show how the firm-specific mean effect is eliminated. Cross-sectional differences can therefore be accommodated by the firm-specific error term ϵ_i.

[29]Unlike the pre-event interval, the five days beginning July 1 do not include all days of the week (due to Independence Day). Since there may be a day-of-the-week effect in the estimated information asymmetry measures (see, for example, Foster and Viswanathan, 1993), we also conducted the analysis with the first five days in July that comprise a full set of the days of the week (dates 1, 2, 6, 7, and 12). The results obtained were very similar to the results using the first five trading days in July, and are therefore omitted for brevity of exposition.

follow the Russell 2000 index). If the proxy for normal volume is not good enough, some of the increase in liquidity trading will be captured by the intercept, making it negative. At the same time, stocks that move from the Russell 2000 to the Russell 1000 should experience a decrease in liquidity trading. For them, insufficient control for liquidity trading will result in a positive intercept. So, opposite results for the addition and deletion samples is consistent with insufficient adjustment for the effects of liquidity trading. If the intercepts of the two samples have the same sign, the results are most likely not due to lack of control for changes in normal trading.

4.2 Results

Panel A of Table 8 presents the results of the event study for the addition sample. The t-tests indicate that the decrease in the mean measure between the pre- and post-event intervals is statistically different from zero for MRR$, MRR, and HASAB. The non-parametric Wilcoxon signed rank test is also highly significant. Normalizing by the pre-event magnitude of the measures, the median percentage changes in MRR$, MRR, and HASAB are -29.67%, -31.08%, and -21.86%, respectively. The fourth column of Panel A shows the intercept from the regression with AvgTurn as a proxy for normal volume and the sixth column shows the intercept from the regression with AvgTrd as a proxy for normal volume. The intercepts in the regressions with MRR$, MRR, and HASAB are all negative and statistically significant, while the volume variables are not significant in these regressions. The situation is reversed in the regressions with HASR: the intercept is not statistically significant, but the volume variables are significant and negative.

The results for the deletion sample are presented in Panel B of Table 8. The mean difference between the pre- and post-event measures is negative, and the t-test confirms that it is statistically different from zero for MRR, and HASAB. The Wilcoxon signed rank test points to the same conclusion. The non-parametric test also confirms that the MRR$ information asymmetry measure decreased, though the t-test is not statistically significant.

A closer inspection reveals that the insignificant t-test is the result of one positive outlier that is more than 7 standard deviations from the mean. Eliminating this outlier results in a very significant t-test. Normalizing by the pre-event magnitude of the measures, the median percentage changes in MRR\$, MRR, and HASAB are -22.99%, -29.44%, and -7.76%, respectively. The intercepts of the regressions with the proxies for normal volume tell the same story: significant negative intercepts for MRR and HASAB (and also for MRR\$ without the outlier).

The results from both the addition and deletion samples point to the same conclusion: the Russell 1000 reconstitution event is associated with a decrease in the information asymmetry measures (with the exception of HASR). This cannot be due to a change in private information about the future cash flows of the firms since the only criterion used for the reconstitution—market capitalization on May 28—is public information. The finding of a decrease in the measures for both the addition and deletion samples also rules out an explanation based on changes in liquidity trading. As the discussion in Section 4.1 points out, even if the proxies for normal volume are not perfect, a change in normal trading would produce opposite results in the addition and deletion samples.

The decrease in the measures is consistent with an explanation that postulates changes in uncertainty about the investor population (in the spirit of Saar 2000a, 2000b). Uncertainty about whether a stock will end up in the Russell 1000 index or the Russell 2000 index during the last week of May creates uncertainty with respect to the composition of the investors (e.g., index funds) who trade the stock. While some fund managers could be trading ahead of time based on a certain assessment as to which stocks will switch indexes, others could be trading based on a different assessment or waiting until the new composition is made known. This increased uncertainty for stocks that may be added to or deleted from the index increases the price impact of trades and is picked up by the information asymmetry measures. In the post-event interval, there is less uncertainty about the investor population since the changes to the index are announced and therefore the permanent price impact of trades decreases.

We conducted several tests to evaluate the robustness of our results. First, we used interval lengths other than five days. Two-day intervals displayed a similar significant decrease in the information asymmetry measures MRR$, MRR, and HASAB.[30] We also used one-month intervals and obtained results similar to those presented in Table 8 (i.e., a statistically significant decrease in MRR$, MRR, and HASAB).[31]

Since MRR$ and MRR behave in a similar fashion, it seems unlikely that the results are due to possible changes in the price level that may affect the estimated measures due to discreteness of the price gird. Nonetheless, we added the change in AvgPrc as a regressor to the specification in Equation 17 to account for changes in the price level, and found that it did not alter our results. We also used a dummy variable to examine whether the 21 stocks that were added to the Russell 1000 index but were not previously in the Russell 2000 index behave differently from the rest of the stocks. The coefficient of the dummy variable was not significant and similar results were obtained. For MRR$ and MRR, we repeated the tests using only positive estimates and only estimates that were statistically different from zero. This was done to eliminate the influence of stocks that did not have sufficient amount of trading to produce reliable estimates of the information asymmetry measures. The results were similar to those reported in Table 8.

Lastly, we constructed a matched sample of stocks based on both market capitalization and average turnover in April 1999.[32] The motivation behind constructing the matched sample was to investigate whether the Russell 1000 reconstitution event coincided with a market-wide change in liquidity that could have affected the estimates of the information asymmetry measures of all stocks. Unlike the results for the addition and deletion samples,

[30]We used May 27 and 28 for the pre-event interval and July 1 and 2 for the post-event interval. Both intervals are comprised of the days Thursday and Friday, and hence the results are not due to a day-of-the-week effect.

[31]The only difference was with respect to the HASR measure, which was significantly negative in the deletion sample and significantly positive in the addition sample.

[32]The universe of stocks that was used for the matching consisted of all common stocks in the CRSP database with information from April 1 to July 12 that were not added to or deleted from the Russell 1000 index.

there was no overwhelming significant decrease in the information asymmetry measures in the matched sample. The only statistically significant change was a decrease in HASAB in the addition-matched sample. The decrease in the matched sample, however, was smaller in magnitude than the decrease in the addition sample, and a paired t-test (as well as a Wilcoxon signed rank test) showed that the difference between them was statistically significant.

5 Conclusion

The objective of this paper is to examine whether spread decomposition and variance decomposition methodologies developed in the market microstructure literature to produce measures of information asymmetry about the future cash flows of firms in fact do so. We separate this task into two parts. In the first, we investigate whether cross-sectional differences in information asymmetry about future cash flows are sufficient to produce differences in the estimated measures. In the second part, we use the Russell 1000 index reconstitution to look at whether changes in information asymmetry are necessary to produce changes in the estimated measures.

The cross-sectional analysis is important since if we conclude that these measures are related in the appropriate direction to our proxies for future cash flow uncertainty and price informativeness, then these measures can be useful for detecting changes in the information asymmetry environment. Even if this is the case, however, the measures are vulnerable to the claim that other things in the economic environment can cause them to change, and therefore changes in these measures would always be open to alternative interpretations. This is the motivation for the second part of the analysis. It is very difficult to prove that nothing but information asymmetry about future cash flows affects the estimated measures since it requires to consider every possible scenario. To disprove this claim, however, requires only one example. We use the Russell 1000 index reconstitution because it provides a clean experiment of a change to the composition of the investor population without any new information about the firm.

Already the results of the cross-sectional and price informativeness analyses, however, cast doubt on the reliability of the measures as depictions of information asymmetry about future cash flows. The MRR measure does not seem to relate to uncertainty about future cash flows in a consistent fashion. It is negatively related to both the dispersion of analysts' forecasts and to price informativeness, exactly the opposite of what we would expect. It is negatively related to a measure of the predictability of earnings based on past data, as it should be, but the mixed results indicate that this measure probably cannot be used reliably to represent information asymmetry about future cash flows.

While the HASAB and HASR measures are both based on the variance decomposition methodology, they turn out to have different properties. In particular, it does not seem as if the HASR measure represents the phenomena we associate with information asymmetry about future cash flows. It may be that the amount of private information relative to all information is not what affects the market, but rather the amount of private information per se. HASAB shows promise in the cross-sectional analysis in that it has the predicted relations with two proxies for uncertainty about future cash flows, one using analysts' forecasts and another using past earnings variability. The picture changes when we consider the price informativeness analysis, where more informed trading according to HASAB is associated with less information in prices about future earnings. This contradictory result weakens our belief in the usefulness of the measure, but the HASAB measure still seems to perform better than MRR.

One could ask why researchers do not use proxies for uncertainty about future cash flows instead of the spread or variance components if indeed these proxies are more strongly associated with information asymmetry about future cash flows. We believe that the abundance of trading data and the fact that the market microstructure measures can be estimated for very short periods or for portions of the day greatly contributed to the popularity of these measures. Past earnings data or even analysts' forecasts cannot be measured at fine intervals, nor do they allow the flexibility of intraday data in establishing the information environment for portions of the day or hours around a corporate event. By looking at the relation between the market microstructure measures and proxies for future cash flows uncertainty or price informativeness (when those can be obtained), we are able to examine whether the former reflect information asymmetry about future cash flows and can serve as a convenient substitute for the latter.

The design of the event study was guided by our conjecture that other things besides future cash flow information indeed affect the information asymmetry measures. This conjecture is supported by our findings: the MRR (both in dollar and percentage terms) and the HASAB measures decrease significantly around the reconstitution event, and this cannot be attributed to mere changes in uninformed or liquidity trading.

Callahan, Lee, and Yohn (1997) claim that the information asymmetry risk faced by a dealer, presumably the one reflected in the permanent price impact of trades, is not a measure of the long-term fundamental risk of investing in a particular firm. Market makers are concerned with order imbalances over very short horizons (often less than an hour), and these need not correspond to the creation of new information about the firm or the trading by investors with special knowledge about the firm. It may be that the kind of information that is picked up by spread decomposition or variance decomposition procedures is not as "permanent" as we would like to think and can arise from many reasons that give rise to imbalances in the order flow. Even if information about the firm is one cause for these order flow imbalances, its effects may be miniscule relative to other factor that influence minute-by-minute trading.

This may also explain the better performance of HASAB (as opposed to MRR) in the cross-sectional analysis. If information about the firm indeed affects prices in a more permanent fashion than other factors, then the better we are at identifying permanent price changes, the stronger will be the connection between the measures and future cash flow information. The methodology that generates HASAB uses five lags of price changes to identify permanent effects. The econometric model that generates the MRR measure essentially uses only one lag. As such, the MRR measure may incorporate more of the other factors that affect trading, reducing its effectiveness in identifying information about the firm. Further increasing the number of lags in the variance decomposition methodology, while preventing its use for thinly traded stock, may increase its effectiveness. This is especially true in future work that will use prices quoted in decimals, as the reduced tick size dramatically increased in the number of quote changes.

Our findings raise doubts as to the strength of the conclusions one can draw when using the measures of information asymmetry produced by spread decomposition and variance decomposition techniques. While some of our results are consistent with the traditional use of these measures, the overall weight of the evidence points to a conclusion that measures based on the permanent price impact of trades may not be suitable for describing information asymmetry about the firm.

References

Affleck-Graves, J., S. P. Hedge, and R. E. Miller (1994): "Trading Mechanisms and the Components of the Bid-Ask Spread," *Journal of Finance*, 49(4), 1471–1488.

Barclay, M. J., and C. G. Dunbar (1996): "Private Information and the Costs of Trading around Quarterly Earnings Announcements," *Financial Analysts Journal*, November-December, 75–84.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens (1998): "Using Analysts' Forecasts to Measure Properties of Analysts' Information Environment," *Accounting Review*, 73(4), 421–433.

Brennan, M. J., and A. Subrahmanyam (1995): "Investment Analysis and Price Formation in Securities Markets," *Journal of Financial Economics*, 38, 361–381.

Brooks, R. M. (1994): "Bid-Ask Spread Components Around Anticipated Announcements," *Journal of Financial Research*, 42(3), 375–386.

Callahan, C. M., C. M. Lee, and T. Lombardi Yohn (1997): "Accounting Information and Bid-Ask Spreads," *Accounting Horizons*, 11(4), 50–60.

Campbell, J. Y., and J. Ammer (1993): "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns," *Journal of Finance*, 48(1), 3–37.

Clarke, J., and K. Shastri (2000): "On Information Asymmetry Metrics," working paper, Katz Graduate School of Business, University of Pittsburgh.

——— (2001): "Adverse Selection Costs and Closed-End Funds," working paper, Katz Graduate School of Business, University of Pittsburgh.

Collins, D. W., S. P. Kothari, J. Shanken, and R. G. Sloan (1994): "Lack of Timeliness and Noise as Explanations for the Low Contemporaneous Return-Earnings Association," *Journal of Accounting and Economics*, 18, 289–324.

Davidson, R., and J. G. MacKinnon (1989): "Testing for Consistency Using Artificial Regressions," *Econometric Theory*, 5, 363–384.

——— (1993): *Estimation and Inference in Econometrics*. Oxford University Press, New York.

Dennis, P. J., and J. P. Weston (2001): "Who's Informed? An Analysis of Stock Ownership and Informed Trading," working paper, McIntire School of Commerce, University of Virginia.

Desai, A. S., M. Nimalendran, and S. Venkataraman (1998): "Changes in Trading Activity Following Stock Splits and their Effect on Volatility and the Adverse-Information Fomponent of the Bid-Ask Spread," *Journal of Financial Research*, 21(2), 159–183.

Dey, M. K., and B. Radhakrishna (2001): "Institutional Trading, Trading Volume, and Spread," working paper, Carlson School of Management, University of Minnesota.

Dhillon, U., and H. Johnson (1991): "Changes in the Standard and Poor's 500 List," *Journal of Business*, 64(1), 75–85.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin (2001): "Does Greater Firm-specific Variation Mean More or Less Informed Stock Pricing?," working paper, Stern School of Business, New York University.

Easley, D., N. M. Kiefer, and M. O'Hara (1997a): "The Information Content of the Trading Process," *Journal of Empirical Finance*, 4, 159–186.

——— (1997b): "One Day in the Life of a Very Common Stock," *Review of Financial Studies*, 10(3), 805–835.

Easley, D., and M. O'Hara (1987): "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, 19, 69–90.

Fama, E. F. (1990): "Stock Returns, Expected Returns, and Real Activity," *Journal of Finance*, 45(4), 1089–1108.

Fee, C. E., and S. Thomas (1999): "Corporate Diversification, Asymmetric Information, and Firm Value: Evidence from Stock Market Trading Characteristics," working paper, Eli Broad College of Business, Michigan State University.

Flannery, M. J., S. H. Kwan, and M. Nimalendran (2000): "Market Evidence on the Opaqueness of Banking Firms' Assets," working paper, Graduate School of Business Administration, University of Florida.

Foster, F. D., and S. Viswanathan (1993): "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models," *Journal of Finance*, 48(1), 187–211.

George, T. J., G. Kaul, and M. Nimalendran (1991): "Estimation of the Bid-Ask Spread and its Components: A New Approach," *Review of Financial Studies*, 4(4), 623–656.

Glosten, L. R. (1987): "Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices," *Journal of Finance*, 42, 1293–1307.

Glosten, L. R., and L. E. Harris (1988): "Estimating the Components of the Bid-Ask Spread," *Journal of Financial Economics*, 21, 123–142.

Glosten, L. R., and P. Milgrom (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

Grossman, S. J., and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70(3), 393–408.

Harris, L., and E. Gurel (1986): "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures," *Journal of Finance*, 41(4), 815–829.

Hasbrouck, J. (1988): "Trades, Quotes, Inventories, and Information," *Journal of Financial Economics*, 22, 229–252.

——— (1991a): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46(1), 179–207.

——— (1991b): "The Summary Informativeness of Stock Trades: An Econometric Analysis," *Review of Financial Studies*, 4(3), 571–595.

Hausman, J. A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46, 1251–1272.

Heflin, F., and K. W. Shaw (2000a): "Blockholder Ownership and Market Liquidity," *Journal of Financial and Quantitative Analysis*, 35(4), 621–633.

——— (2000b): "Trade Size and the Adverse Selection Component of the Spread: Which Trades are "Big"?," working paper, Krannert Graduate School of Management, Purdue University.

Heflin, F., K. W. Shaw, and J. J. Wild (2000): "Disclosure Quality and Market Liquidity," working paper, Krannert Graduate School of Management, Purdue University.

Huang, R. D., and H. R. Stoll (1997): "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10(4), 995–1034.

Illessy, J. K., and K. Shastri (2000): "The Debt-Equity Mix in Preferred Stock and Adverse Selection Costs: An Empirical Investigation," working paper, University of Pittsburgh.

Jagadeesh, N., and A. Subrahmanyam (1993): "Liquidity Effects of the Introduction of the S&P500 Index Futures Contract on the Underlying Stocks," *Journal of Business*, 66(2), 171–187.

Jain, P. C. (1987): "The Effect on Stock Price of Inclusion in or Exclusion from the S&P 500," *Financial Analyst Journal*, pp. 58–65.

Jennings, R. (1994): "Intraday Changes in Target Firms' Share Price and Bid-Ask Quotes Around Takeover Announcements," *Journal of Financial Research*, 17(2), 255–270.

Kothari, S. P. (1992): "Price-Earnings Regressions in the Presence of Prices Leading Earnings," *Journal of Accounting and Economics*, 15, 173–202.

Kothari, S. P., and R. G. Sloan (1992): "Information in Prices about Future Earnings: Implications for Earnings Reseponse Coefficients," *Journal of Accounting and Economics*, 15, 143–171.

Krinsky, I., and J. Lee (1996): "Earnings Announcements and the Components of the Bid-Ask Spread," *Journal of Finance*, 51(4), 1523–1535.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335.

Levy, H. (1978): "Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio," *American Economic Review*, 68(4), 643–658.

Lin, J.-C., G. C. Sanger, and G. G. Booth (1995): "Trade Size and Components of the Bid-Ask Spread," *Review of Financial Studies*, 8(4), 1153–1183.

Lynch, A., and R. R. Mendenhall (1997): "New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index," *Journal of Business*, 70(3), 351–383.

Madhavan, A., M. Richardson, and M. Roomans (1997): "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks," *Review of Financial Studies*, 10(4), 1035–1064.

Madhavan, A., and S. Smidt (1991): "A Bayesian Model of Intraday Specialist Pricing," *Journal of Financial Economics*, 30, 99–134.

Merton, R. C. (1987): "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42(3), 483–510.

Neal, R. (1992): "A Comparison of Transaction Costs between Competitive Market Maker and Specialist Market Structures," *Journal of Business*, 65, 317–334.

Neal, R., and S. M. Wheatley (1998): "Adverse Selection and Bid-Ask Spreads: Evidence from Closed-End Funds," *Journal of Financial Markets*, 1, 121–149.

Porter, D. C., and D. G. Weaver (1996): "Estimating Bid-Ask Spread Components: Specialist versus Multiple Market Maker Systems," *Review of Quantitative Finance and Accounting*, 6, 167–180.

Saar, G. (2000a): "Demand Uncertainty and the Information Content of the Order Flow," working paper, Stern School of Business, New York University.

——— (2000b): "Prices and Spreads in Sequential Markets with Information Imperfections," working paper, Stern School of Business, New York University.

Sarin, A., K. A. Shastri, and K. Shastri (2000): "Ownership Structure and Stock Market Liquidity," working paper, Katz Graduate School of Business, University of Pittsburgh.

Schipper, K., and R. Thompson (1983): "The Impact of Merger-Related Regulations on the Shareholders of Acquiring Firms," *Journal of Accounting Research*, 21(1), 184–221.

Shapiro, A. (2001): "The Investor Recognition Hypothesis in a Dynamic General Equilibrium: Theory and Evidence," Forthcoming in the *Review of Financial Studies.*

Shleifer, A. (1986): "Do Demand Curves for Stocks Slope Down?," *Journal of Finance*, 41(3), 579–590.

Stoll, H. R. (1989): "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests," *Journal of Finance*, 44(1), 115–134.

Thompson, R. (1985): "Conditioning the Return-Generating Process on Firm-Specific Events: A Discussion of Event Study Methods," *Journal of Financial and Quantitative Analysis*, 20, 151–168.

Wei, P.-H. (1992): "Intraday Variations in Trading Activity, Price Variability, and the Bid-Ask Spread," *Journal of Financial Research*, 15(3), 265–285.

Wu, D.-M. (1973): "Alternative Tests of Independence Between Stochastic Regressors and Disturbances," *Econometrica*, 41(4), 733–750.

Table 1
Summary Statistics for the Cross-sectional Sample

The cross-sectional sample is constructed by including all common stocks with information in the CRSP database between January 4 and December 31, 1999 (the sample period) that have at least one analyst covering them in the I/B/E/S database. The sample is further restricted to stocks with information in TAQ, COMPUSTAT, and Value Line Investment Survey. The following variables are calculated for each stock over the sample period using data in CRSP: AvgCap is the average daily market capitalization (the number of shares outstanding multiplied by the daily closing price), StdRet is the standard deviation of daily returns, MedTurn is the median daily turnover (the number of shares traded divided by the number of shares outstanding), AvgVol is the average daily number of shares traded, and AvgPrc is the average daily closing price of the stock. From the TAQ database, AvgTrd is calculated as the average daily number of trades, and %Sprd is the average percentage spread (the bid-ask spread divided by the midquote). NumEst is the average of 12 monthly observations in 1999 of the number of analysts with end-of-fiscal-year earnings forecasts for the current year in I/B/E/S. InstHol is the percentage of shares held by institutions from Value Line Investment Survey. The table presents summary statistics for the entire sample, as well as means for four equal groups sorted by AvgCap, StdRet, and MedTurn, respectively.

		AvgCap (in million $)	**StdRet** (in %)	**MedTurn** (in %)	**AvgVol** (in 100s)	**AvgPrc** (in $)	**AvgTrd**	**%Sprd** (in %)	**NumEst**	**InstHol** (in %)
Entire	Mean	4,414.29	3.62	0.50	5,128.72	26.68	442.67	1.25	8.81	50.38
Sample	Median	563.45	3.30	0.31	1,617.20	20.93	117.56	0.89	6.52	51.12
	Std. Dev.	19,183.64	1.70	0.65	13,875.91	24.50	1,584.67	1.17	6.77	22.72
	Min.	6.05	0.50	0.00	11.00	0.24	3.03	0.07	2.00	0.23
	Max.	446,161.27	25.73	15.64	280,833.61	550.05	29,809.45	19.95	40.18	99.59
	Obs.	2,876	2,876	2,876	2,876	2,876	2,876	2,876	2,876	2,876
Means	Q1 (Low)	106.38	4.47	0.37	831.57	10.18	76.21	2.62	3.87	36.60
For	Q2	354.09	3.90	0.51	1,833.17	18.55	163.85	1.27	5.78	48.10
AvgCap	Q3	1,047.68	3.25	0.54	3,007.62	29.25	248.40	0.74	8.59	55.68
Groups	Q4 (High)	16,149.00	2.86	0.59	14,842.49	48.73	1,282.20	0.38	16.99	61.16
Means	Q1 (Low)	7,428.85	1.90	0.21	3,661.77	36.65	223.66	0.68	10.37	51.58
for	Q2	7,212.71	2.80	0.31	5,972.29	28.43	419.18	0.99	9.97	56.30
StdRet	Q3	2,124.60	3.90	0.53	5,715.42	21.72	489.40	1.41	8.42	52.55
Groups	Q4 (High)	890.99	5.89	0.96	5,165.39	19.92	638.43	1.93	6.48	41.11
Means	Q1 (Low)	3,697.73	2.84	0.11	1,247.20	21.77	82.03	1.80	5.91	37.07
for	Q2	5,639.16	3.00	0.24	3,781.44	26.29	222.07	1.18	9.51	52.71
MedTurn	Q3	4,832.07	3.67	0.42	5,163.21	27.42	326.38	1.13	9.87	57.68
Groups	Q4 (High)	3,488.18	4.97	1.23	10,323.00	31.24	1,140.19	0.91	9.96	54.07

Table 2
Summary Statistics for the Information Asymmetry Measures

Three measures of information asymmetry, MRR, HASAB, and HASR, are estimated for each of the stocks in the cross-sectional sample using intraday transactions data (trades and quotes) between January 4 and December 31, 1999 from the TAQ database. MRR is the permanent price impact of the order flow innovation developed in Madhavan, Richardson, and Roomans (1997). HASAB is the trade-correlated component of the efficient price variance (in standard deviation per hour form), and HASR is the trade-correlated component relative to the total efficient price variance, both introduced in Hasbrouck (1991b). The table presents summary statistics for the entire sample, as well as means for four equal groups sorted by AvgCap (average daily market capitalization), StdRet (the standard deviation of daily returns), and MedTurn (the median daily turnover), respectively.

		MRR	**HASAB**	**HASR**
Entire	Mean	0.0737	0.3907	0.1617
Sample	Median	0.0513	0.3685	0.1490
	Std. Dev.	0.0768	0.1970	0.0938
	Min.	-0.0102	0.0483	0.0008
	Max.	1.0569	2.2993	0.8836
	Obs.	2,876	2,868	2,868
Means	Q1 (Low)	0.1168	0.4717	0.1515
for	Q2	0.0813	0.3993	0.1541
AvgCap	Q3	0.0629	0.3488	0.1639
Groups	Q4 (High)	0.0336	0.3437	0.1770
Means	Q1 (Low)	0.0736	0.2788	0.1997
for	Q2	0.0819	0.3693	0.1907
StdRet	Q3	0.0748	0.4330	0.1512
Groups	Q4 (High)	0.0645	0.4817	0.1051
Means	Q1 (Low)	0.1314	0.3683	0.1949
for	Q2	0.0766	0.3867	0.1836
MedTurn	Q3	0.0590	0.4138	0.1643
Groups	Q4 (High)	0.0277	0.3939	0.1040

Table 3
Summary Statistics for Future Cash Flows Uncertainty Proxies

We use as proxies for uncertainty about future cash flows two definitions of dispersion of analysts' earnings forecasts: log of the standard deviation of analysts' forecasts (StdEst) and log of the coefficient of variation of analysts' forecasts (CoefEst). I/B/E/S provides monthly observations of the standard deviation and mean of the analysts' forecasts for the current end-of-fiscal-year. StdEst is the simple average of the monthly standard deviation observations for all months in 1999, and CoefEst is the ratio of the average standard deviation to the average absolute mean forecast. We also use two proxies for uncertainty about future cash flows based on past earnings: VLPred is an earning predictability measure provided by Value Line Investment Survey derived from the standard deviation of percentage changes in quarterly earnings over an eight-year period, and StdEPS is log of the standard deviation of annual earnings for the five years prior to 1999. Panel A presents summary statistics for the entire sample (before taking the log transformations), as well as for four equal groups sorted by AvgCap (average daily market capitalization), StdRet (the standard deviation of daily returns), and MedTurn (the median daily turnover), respectively. Panel B presents pairwise correlations and *p*-values of the four future cash flows uncertainty proxies and the three control variables used in the cross-sectional regressions (where LAvgCap and LMedTurn are the log transformations of AvgCap and MedTurn, respectively). The two-sided *p*-values for the asymptotic test of zero correlations are shown in parentheses. "**" indicates significance at the 1% level and "*" indicates significance at the 5% level.

Panel A: **Summary statistics of untransformed variables**

		StdEst	**CoefEst**	**VLPred**	**StdEPS**
Entire	Mean	0.0882	0.3842	47.7827	0.9176
Sample	Median	0.0490	0.0410	45.0000	0.5600
	Std. Dev.	0.1265	7.0859	28.3613	1.2713
	Min.	0.0008	0.0008	5.0000	0.0100
	Max.	2.5400	363.0000	100.0000	22.0300
	Obs.	2,876	2,876	2,072	2,845
Means	Q1 (Low)	0.0948	1.0437	33.4708	0.6904
for	Q2	0.0862	0.1973	40.5537	0.8333
AvgCap	Q3	0.0790	0.1709	51.7693	0.8990
Groups	Q4 (High)	0.0929	0.1249	59.6099	1.2466
Means	Q1 (Low)	0.0809	0.0504	62.5261	0.9303
for	Q2	0.0724	0.0818	51.1898	0.9590
StdRet	Q3	0.0861	0.7715	38.9826	0.8787
Groups	Q4 (High)	0.1134	0.6331	27.0140	0.9022
Means	Q1 (Low)	0.0746	0.1347	54.4925	0.7259
for	Q2	0.0769	0.1538	53.6217	0.8425
MedTurn	Q3	0.0987	0.9287	45.4728	1.0325
Groups	Q4 (High)	0.1027	0.3196	35.3902	1.0675

Panel B: **Unconditional correlations of transformed variables used in the regressions**

				Control Variables			
	StdEst (p-value)	**CoefEst** (p-value)	**VLPred** (p-value)	**StdEPS** (p-value)	**LAvgCap** (p-value)	**StdRet** (p-value)	**LMedTurn** (p-value)
StdEst	_	0.6748**	-0.4057**	0.3733**	-0.0672**	0.1149**	0.0489*
		(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0000)	(0.0119)
CoefEst		_	-0.5713**	0.1524**	-0.3123**	0.4500**	0.1386**
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
VLPred			_	-0.3557**	0.3813**	-0.4090**	-0.2190**
				(0.0000)	(0.0000)	(0.0000)	(0.0000)
StdEPS				_	0.2125**	-0.0266	0.1337**
					(0.0000)	(0.2050)	(0.0000)
LAvgCap					_	-0.3794**	0.1722**
						(0.0000)	(0.0000)
StdRet						_	0.4667**
							(0.0000)
LMedTurn							_

Table 4
Cross-Sectional OLS Estimation Results

This table provides the results of cross-sectional regressions where the information asymmetry measures are the dependent variables and the proxies for uncertainty about future cash flows (alongside controls) are the independent variables. We use as proxies for uncertainty about future cash flows two definitions of dispersion of analysts' earnings forecasts: log of the standard deviation of analysts' forecasts (StdEst) and log of the coefficient of variation of analysts' forecasts (CoefEst). I/B/E/S provides monthly observations of the standard deviation and mean of the analysts' forecasts for the current end-of-fiscal-year. StdEst is the simple average of the monthly standard deviation observations for all months in 1999, and CoefEst is the ratio of the average standard deviation to the absolute value of the average monthly mean observations. We also use two proxies for uncertainty about future cash flows based on past earnings: VLPred is an earning predictability measure provided by Value Line Investment Survey that is derived from the standard deviation of percentage changes in quarterly earnings over an eight-year period, and StdEPS is log of the standard deviation of annual earnings for the five years prior to 1999. The following set of control variables is used in each regression: the logarithmic transformation of the average daily market capitalization (LAvgCap), the standard deviation of daily returns (StdRet), and the logarithmic transformation of the median daily turnover (LMedTurn). The regressions also include (but they are not shown in the table) a set of two-digit SIC dummies to control for industry effects, and dummies for the three primary markets (NYSE, AMEX, and Nasdaq) to control for institutional characteristics of the markets. Panel A presents the results for the MRR measure, the permanent price impact of the order flow innovation, estimated using the methodology in Madhavan, Richardson, and Roomans (1997). Panel B presents the results for the HASAB measure, the trade-correlated component of the efficient price variance (in standard deviation per hour form), estimated using the methodology in Hasbrouck (1991b). Panel C presents the results for HASR, the ratio of the trade-correlated component to the total efficient price variance, also from Hasbrouck (1991b). All estimations are done using OLS, and we report p-values in parenthesis calculated using White’s heteroskedasticity-consistent standard errors. "**" indicates significance at the 1% level and "*" indicates significance at the 5% level.

Panel A: **MRR (dependent variable)**

				Control variables				
StdEst (p-value)	**CoefEst** (p-value)	**VLPred** (p-value)	**StdEPS** (p-value)	**LAvgCap** (p-value)	**StdRet** (p-value)	**LMedTurn** (p-value)	**AdjR^2** (in %)	**Obs.**
-0.00340**				-0.0209**	0.6214**	-0.0334**	61.14	2867
(0.0014)				(0.0000)	(0.0000)	(0.0000)		
	-0.00112			-0.0208**	0.6255**	-0.0338**	61.02	2867
	(0.1551)			(0.0000)	(0.0000)	(0.0000)		
		-0.00008*		-0.0187**	0.4460**	-0.0336**	65.43	2067
		(0.0379)		(0.0000)	(0.0000)	(0.0000)		
			0.00022	-0.0206**	0.5654**	-0.0333**	60.86	2836
			(0.8191)	(0.0000)	(0.0000)	(0.0000)		
-0.00184		-0.00010*		-0.0188**	0.4605**	-0.0335**	65.46	2067
(0.1416)		(0.0118)		(0.0000)	(0.0006)	(0.0000)		
	-0.00153	-0.00011**		-0.0189**	0.4952**	-0.0337**	65.47	2067
	(0.1263)	(0.0077)		(0.0000)	(0.0006)	(0.0000)		
-0.00386**			0.00147	-0.0211**	0.5969**	-0.0330**	61.02	2836
(0.0009)			(0.1660)	(0.0000)	(0.0000)	(0.0000)		
	-0.00123		0.00050	-0.0209**	0.6033**	-0.0333**	60.88	2836
	(0.1250)		(0.6191)	(0.0000)	(0.0000)	(0.0000)		

Panel B: **HASAB (dependent variable)**

				Control variables				
StdEst (p-value)	**CoefEst** (p-value)	**VLPred** (p-value)	**StdEPS** (p-value)	**LAvgCap** (p-value)	**StdRet** (p-value)	**LMedTurn** (p-value)	**AdjR^2** (in %)	**Obs.**
-0.00012				-0.0280**	6.0764**	-0.0060	38.13	2859
(0.9735)				(0.0000)	(0.0000)	(0.2491)		
	0.01325**			-0.0254**	5.6598**	-0.0066	38.67	2859
	(0.0000)			(0.0000)	(0.0000)	(0.2014)		
		-0.00031*		-0.0228**	6.4685**	-0.0014	42.39	2062
		(0.0395)		(0.0000)	(0.0000)	(0.8252)		
			0.00100	-0.0278**	6.0960**	-0.0058	38.29	2828
			(0.7764)	(0.0000)	(0.0000)	(0.2845)		
-0.00434		-0.00036*		-0.0230**	6.5028**	-0.0012	42.39	2062
(0.2928)		(0.0286)		(0.0000)	(0.0000)	(0.8502)		
	0.01313**	-8.22e-05		-0.0216**	6.0456**	-5.34e-5	42.83	2062
	(0.0006)	(0.5943)		(0.0000)	(0.0000)	(0.9930)		
-0.00117			0.00137	-0.0280**	6.1056**	-0.0254	38.28	2828
(0.7684)			(0.7207)	(0.0000)	(0.0000)	(0.4709)		
	0.01322**		-0.00192	-0.0251**	5.6894**	-0.0059	38.80	2828
	(0.0000)		(0.5992)	(0.0000)	(0.0000)	(0.2762)		

Panel C: **HASR (dependent variable)**

				Control variables				
StdEst (p-value)	**CoefEst** (p-value)	**VLPred** (p-value)	**StdEPS** (p-value)	**LAvgCap** (p-value)	**StdRet** (p-value)	**LMedTurn** (p-value)	**AdjR^2** (in %)	**Obs.**
-0.00106				-0.0127**	-0.8555**	-0.0067**	45.69	2859
(0.4882)				(0.0000)	(0.0000)	(0.0046)		
	-0.00191			-0.0130**	-0.8054**	-0.0068**	45.72	2859
	(0.0877)			(0.0000)	(0.0000)	(0.0043)		
		0.00001		-0.0128**	-0.9170**	-0.0045	42.02	2062
		(0.8512)		(0.0000)	(0.0000)	(0.1198)		
			0.00102	-0.0126**	-0.8660**	-0.0069**	45.53	2828
			(0.4936)	(0.0000)	(0.0000)	(0.0042)		
0.00188		0.00003		-0.0127**	-0.9318**	-0.0045	42.02	2062
(0.3211)		(0.6422)		(0.0000)	(0.0000)	(0.1137)		
	-0.00180		-0.00002	-0.0130**	-0.8592**	-0.0046	42.03	2062
	(0.2405)		(0.7962)	(0.0000)	(0.0000)	(0.1070)		
-0.00177			0.00160	-0.0128**	-0.8516**	-0.0068**	45.54	2828
(0.2842)			(0.3227)	(0.0000)	(0.0000)	(0.0048)		
	-0.00228*		0.00152	-0.0131**	-0.7959**	-0.0069**	45.58	2828
	(0.0452)		(0.3143)	(0.0000)	(0.0000)	(0.0042)		

Table 5
Durbin-Wu-Hausman Tests and Cross-Sectional 2SLS Estimation

This table provides the results of the Durbin-Wu-Hausman tests as well as cross-sectional two-stage-least-squares (2SLS) regressions where the information asymmetry measures are the dependent variables and the analysts proxies for uncertainty about future cash flows (alongside controls) are the independent variables. The following instruments are used for analysts' earnings forecasts dispersion: (i) the log standard deviation of EPS for the five years prior to 1999 from COMPUSTAT, (ii) the one-year lagged growth of EPS, (iii) the average absolute value of quarterly sales growth in 1999 from COMPUSTAT, (iv) the number of days with listing records in CRSP, (v) the log of the number of SIC codes associated with the firm, and (vi) the dividend payout ratio for 1999 from COMPUSTAT. We use as proxies for uncertainty about future cash flows two definitions of dispersion of analysts' earnings forecasts: log of the standard deviation of analysts forecasts (StdEst) and log of the coefficient of variation of analysts forecasts (CoefEst). I/B/E/S provides monthly observations of the standard deviation and mean of the analysts' forecasts for the current end-of-fiscal-year. StdEst is the simple average of the monthly standard deviation observations for all months in 1999, and CoefEst is the ratio of the average standard deviation to the absolute value of the average monthly mean observations. The following set of control variables constructed over the sample period is used in each regression: the logarithmic transformation of the average daily market capitalization (LAvgCap), the standard deviation of daily returns (StdRet), and the logarithmic transformation of the median daily turnover (LMedTurn). The regressions also include (but they are not shown in the table) a set of two-digit SIC dummies to control for industry effects, and dummies for the three primary markets (NYSE, AMEX, and Nasdaq) to control for institutional characteristics of the markets. The table presents the regression results for the MRR measure, the permanent price impact of the order flow innovation, estimated using the methodology in Madhavan, Richardson, and Roomans (1997), the HASAB measure, the trade-correlated component of the efficient price variance (in standard deviation per hour form), estimated using the methodology in Hasbrouck (1991b), and HASR, the ratio of the trade-correlated component to the total efficient price variance, also from Hasbrouck (1991b). "**" indicates significance at the 1% level and "*" indicates significance at the 5% level.

			Control variables			Durbin-Wu-Hausman statistic	
Dependent variables	**StdEst** (*p*-value)	**CoefEst** (*p*-value)	**LAvgCap** (*p*-value)	**StdRet** (*p*-value)	**LMedTurn** (*p*-value)	(*p*-value)	**Obs.**
MRR	-0.0051		-0.0214**	0.6606**	-0.0316**	0.3770	2454
	(0.0861)		(0.0000)	(0.0000)	(0.0000)	(0.5392)	
		-0.0020	-0.0212**	0.6821**	-0.0321**	0.0599	2454
		(0.5901)	(0.0000)	(0.0005)	(0.0000)	(0.8066)	
HASAB	0.0007		-0.0258**	6.1250**	-0.0028	0.0076	2447
	(0.9484)		(0.0000)	(0.0000)	(0.6321)	(0.9304)	
		0.0117	-0.0236**	5.7541**	-0.0030	0.0132	2447
		(0.3739)	(0.0000)	(0.0000)	(0.5810)	(0.9086)	
HASR	0.0006		-0.0126**	-0.7809**	-0.0060*	0.1400	2447
	(0.9031)		(0.0000)	(0.0000)	(0.0139)	(0.7084)	
		0.0022	-0.0123**	-0.8460**	-0.0060*	0.5510	2447
		(0.6986)	(0.0000)	(0.0003)	(0.0110)	(0.4579)	

Table 6
FERC and FINC Analysis

This table provides estimates of correlations between the information asymmetry measures and two proxies for price informativeness about future earnings, FERC and FINC. To construct these proxies, we run the following two regressions:

$$r_t = a + b_0 \Delta E_t + \sum_{\tau=1}^{2} b_\tau \Delta E_{t+\tau} + \sum_{\tau=1}^{2} c_\tau r_{t+\tau} + u_t$$

$$r_t = a + b_0 \Delta E_t + v_t$$

where r_t is the quarterly return at time t, $r_{t+\tau}$ is the return τ quarters ahead, ΔE_t is the change in EPS at time t, and $\Delta E_{t+\tau}$ is the change in EPS τ quarters ahead. All earnings changes are scaled by the price at time *t-1*. We construct the returns for four quarters in 1999 using CRSP, and take EPS data from COMPUSTAT for 1999 and the first two quarters of 2000 (we set τ=2). The definitions of the price informativeness proxies are:

$$\text{FERC} \equiv \sum_{\tau} b_\tau \qquad \text{FINC} \equiv R^2_{r=a+b_0\Delta E_t+\sum_{\tau=1}^{2} b_\tau \Delta E_{t+\tau}+\sum_{\tau=1}^{2} c_\tau r_{t+\tau}+u_t} - R^2_{r_t=a+b_0\Delta E_t+v_t}$$

Since each firm has only four observations, we group stocks with similar information asymmetry measures and estimate for each group a common price informativeness proxy. The information asymmetry measure for the group is the average of the measures of the stocks in the group. For the MRR measure, we sort the sample firms according to MRR, form 287 groups of 10 stocks each, estimate the two equations above for each group separately, and calculate FERC and FINC for the group. A similar procedure is applied to the HASAB and HASR measures. Panel A presents the Pearson correlations between the groups' information asymmetry measures (where a group measure is the mean of the measures of the individual stocks in the group) and both FERC and FINC. Panels B, C, and D present the correlations and *p*-values of the three information asymmetry measures with FERC and FINC controlling for the amount of public information and liquidity trading by dividing the stocks into four categories. For the MRR measure, we sort the stocks according to AvgCap, and separate into two categories (Low and High) with equal number of stocks. Each category is sorted according to MedTurn and divided into two equal categories. We then sort the stocks within each category according to MRR, and form groups of 10 stocks each. We estimate FERC and FINC for each group and compute the correlations separately for each category. The process is repeated with control variables StdRet and MedTurn. A similar procedure is applied to the HASAB and HASR measures. In all panels, the two-sided *p*-values for the asymptotic test of zero correlations are shown in parentheses. "**" indicates significance at the 1% level and "*" indicates significance at the 5% level.

Panel A: **Unconditional correlations**

	FERC (p-value)	**FINC** (p-value)	**Obs.**
MRR	-0.1003*	-0.0369	287
	(0.0441)	(0.4644)	
HASAB	-0.1419*	-0.1647**	287
	(0.0126)	(0.0033)	
HASR	-0.0608	-0.0679	287
	(0.3967)	(0.2009)	

Panel B: **MRR Conditional Correlations**

Control Categories	**FERC** (p-value)	**FINC** (p-value)	**Obs.**	**Control Categories**	**FERC** (p-value)	**FINC** (p-value)	**Obs.**
Low AvgCap	-0.0502	-0.2411	72	Low StdRet	-0.0725	0.0492	72
Low MedTurn	(0.5880)	(0.0536)		Low MedTurn	(0.4598)	(0.6291)	
Low AvgCap	-0.1419	0.1349	72	Low StdRet	-0.0505	0.1040	72
High MedTurn	(0.0849)	(0.2030)		High MedTurn	(0.6861)	(0.3294)	
High AvgCap	-0.1083	0.0995	72	High StdRet	-0.0288	0.0065	72
Low MedTurn	(0.4240)	(0.3496)		Low MedTurn	(0.7059)	(0.9258)	
High AvgCap	-0.1819	-0.1978*	72	High StdRet	-0.0070	0.0123	72
High MedTurn	(0.0722)	(0.0270)		High MedTurn	(0.9189)	(0.8797)	

Panel C: **HASAB Conditional correlations**

Control Categories	**FERC** (p-value)	**FINC** (p-value)	**Obs.**	**Control Categories**	**FERC** (p-value)	**FINC** (p-value)	**Obs.**
Low AvgCap	-0.1658	-0.0560	71	Low StdRet	0.0154	-0.0004	71
Low MedTurn	(0.1153)	(0.8066)		Low MedTurn	(0.8830)	(0.9964)	
Low AvgCap	-0.1840*	-0.2599**	72	Low StdRet	-0.2159	-0.1206	72
High MedTurn	(0.0299)	(0.0096)		High MedTurn	(0.1875)	(0.3291)	
High AvgCap	-0.1933	-0.0819	72	High StdRet	-0.1529	-0.0085	72
Low MedTurn	(0.1688)	(0.5028)		Low MedTurn	(0.0648)	(0.9471)	
High AvgCap	0.1789	0.0565	72	High StdRet	-0.0381	-0.1047	72
High MedTurn	(0.1246)	(0.5075)		High MedTurn	(0.7735)	(0.3227)	

Panel D: **HASR Conditional correlations**

Control Categories	**FERC** (p-value)	**FINC** (p-value)	**Obs.**	**Control Categories**	**FERC** (p-value)	**FINC** (p-value)	**Obs.**
Low AvgCap	0.0766	-0.1423	71	Low StdRet	-0.1046	-0.0011	71
Low MedTurn	(0.6243)	(0.1209)		Low MedTurn	(0.4612)	(0.9921)	
Low AvgCap	-0.0827	-0.0178	72	Low StdRet	-0.1657	-0.1123	72
High MedTurn	(0.5330)	(0.8732)		High MedTurn	(0.3883)	(0.3533)	
High AvgCap	-0.1580	0.0902	72	High StdRet	0.0492	-0.1736	72
Low MedTurn	(0.4148)	(0.2783)		Low MedTurn	(0.6667)	(0.0817)	
High AvgCap	0.1538	-0.0858	72	High StdRet	-0.0660	0.0143	72
High MedTurn	(0.1564)	(0.4109)		High MedTurn	(0.5183)	(0.8929)	

Table 7
Summary Statistics for the Event-study Sample

The basic universe for the event-study sample consists of all stocks that were added to or deleted from the Russell 1000 index in 1999. We restrict the sample to common stocks with trading data between April and mid-July, and eliminate firms about which we find news reports discussing mergers in the months around the Russell reconstitution dates. The following variables are calculated for each stock using data in CRSP: CapMay28 is the market capitalization on May 28 (the number of shares outstanding multiplied by the closing price), PrcMay28 and PrcJuly1 are the closing prices on these two dates, BegTurn is the average daily turnover (the number of shares traded divided by the number of shares outstanding) for the pre-event interval (May 24–28), and EndTurn is the average daily turnover for the post-event interval (July 1–8). The following variables are calculated for each stock using data in TAQ: BegTrd and EndTrd are the average daily number of trades in the pre- and post-events intervals, Beg$Sprd and End$Sprd are the average dollar spreads in the pre- and post-event intervals, and Beg%Sprd and End%Sprd are the average percentage spreads (dollars spreads divided by the mid-quotes) in the pre- and post-event intervals. Panel A presents the statistics for the 130 stocks in the addition sample, and Panel B presents the statistics for the 87 stocks in the deletion sample.

Panel A: **Addition Sample**

	CapMay28 (in million $)	**PrcMay28** (in $)	**PrcJuly1** (in $)	**BegTurn** (in %)	**EndTurn** (in %)	**BegTrd**	**EndTrd**	**Beg$Sprd** (in $)	**End$Sprd** (in $)	**Beg%Sprd** (in %)	**End%Sprd** (in %)
Mean	2,485.28	50.46	51.45	1.76	1.82	1,631.86	1,622.83	0.29	0.24	0.65	0.54
Median	1,786.01	46.50	46.56	0.79	0.81	404.20	422.60	0.24	0.20	0.58	0.49
Std. Dev.	2,499.59	27.15	28.80	2.33	3.11	3,753.72	3,231.21	0.17	0.12	0.33	0.26
Min.	197.75	8.50	8.13	0.03	0.07	18.40	25.20	0.10	0.08	0.13	0.12
Max.	21,402.83	177.19	159.69	15.38	25.58	26,396.80	18,667.60	1.19	0.69	2.64	1.72
Obs.	130	130	130	130	130	130	130	130	130	130	130

Panel B: **Deletion Sample**

	CapMay28 (in million $)	**PrcMay28** (in $)	**PrcJuly1** (in $)	**BegTurn** (in %)	**EndTurn** (in %)	**BegTrd**	**EndTrd**	**Beg$Sprd** (in $)	**End$Sprd** (in $)	**Beg%Sprd** (in %)	**End%Sprd** (in %)
Mean	972.23	21.25	22.27	0.60	0.64	253.11	274.76	0.24	0.18	1.08	0.95
Median	1,066.62	18.88	19.19	0.46	0.49	144.00	156.80	0.17	0.15	0.84	0.80
Std. Dev.	305.24	19.55	20.25	0.57	0.55	314.93	366.13	0.63	0.28	0.58	0.53
Min.	192.69	3.25	1.88	0.02	0.07	7.20	14.00	0.06	0.06	0.47	0.24
Max.	1,346.02	181.00	185.00	3.91	2.94	1,796.40	2,693.80	6.04	2.72	3.59	4.04
Obs.	87	87	87	87	87	87	87	87	87	87	87

Table 8
Event Study Analysis

This table presents the analysis of changes in the information asymmetry measures around the Russell 1000 reconstitution in 1999. The pre-event interval is May 24–28 and the post-event interval is July 1–8 (each contains 5 trading days). ΔMRR$ is the change in the MRR$ measure (similar to MRR but estimated with price differences rather than log price differences) between the post- and pre-event intervals. Similarly, ΔMRR is the change in the MRR measure, ΔHASAB is the change in the HASAB measure, and ΔHASR is the change in the HASR measure. The second column reports the mean of the change in measures and the *p*-value (in parentheses) of a t-test against the hypothesis of zero change. The third column reports the median of the change in measures and the *p*-value (in parentheses) of a Wlicoxon signed rank test against the hypothesis of zero change. Regressions 1 and 2 (columns four through seven) both have the following form:

$$\Delta\phi_i = \alpha + \beta\Delta V_i + \varepsilon_i$$

where $\Delta\phi_i$ is the change in the information asymmetry measures and ΔV_i is the change in a proxy for normal volume. In Regression 1, the proxy for normal volume is the average daily turnover over the interval (i.e., ΔV_i=EndTurn-BegTurn). In Regression 2, the proxy for normal volume is the average daily number of trades over the interval (i.e., ΔV_i=EndTrd-BegTrd). The estimation of the regressions is done using OLS, and *p*-values (in parentheses) are calculated using White's heteroskedasticity-consistent standard errors. Panel A presents the results for the 130 stocks in the addition sample, and Panel B presents the results for the 87 stocks in the deletion sample. "**" indicates significance at the 1% level and "*" indicates significance at the 5% level (both against a two-sided alternative).

Panel A: **Addition Sample**

Dependent Variable	Mean (*p*-value of t-test)	Median (*p*-value of Wilcoxon test)	Regression 1		Regression 2	
			α (*p*-value)	**β(Turn)** (*p*-value)	**α** (*p*-value)	**β(Trd)** (*p*-value)
ΔMRR$	-0.005421**	-0.001407**	-0.005403**	-0.000317	-0.005424**	0.000000
	(0.0007)	(0.0001)	(0.0008)	(0.1910)	(0.0008)	(0.2563)
ΔMRR	-0.000118**	-0.000040**	-0.000117**	-0.000007	-0.000118**	0.000000
	(0.0004)	(0.0000)	(0.0007)	(0.1409)	(0.0007)	(0.2871)
ΔHASAB	-0.126197**	-0.091680**	-0.126586**	0.006924	-0.126204**	-0.000001
	(0.0000)	(0.0000)	(0.0000)	(0.3412)	(0.0000)	(0.9194)
ΔHASR	-0.000449	0.006005	-0.000344	-0.001858*	-0.000479	-0.000003**
	(0.9436)	(0.6141)	(0.9568)	(0.0348)	(0.9398)	(0.0008)

Panel B: **Deletion Sample**

Dependent Variable	Mean (*p*-value of t-test)	Median (*p*-value of Wilcoxon test)	Regression 1		Regression 2	
			α (*p*-value)	**β(Turn)** (*p*-value)	**α** (*p*-value)	**β(Trd)** (*p*-value)
ΔMRR$	-0.002081	-0.001672**	-0.001990	-0.001921	-0.002069	-0.000001
	(0.3322)	(0.0008)	(0.3556)	(0.1741)	(0.3384)	(0.6608)
ΔMRR	-0.000222**	-0.000089**	-0.000218**	-0.000081	-0.000221**	-0.000000
	(0.0027)	(0.0001)	(0.0033)	(0.0740)	(0.0030)	(0.3088)
ΔHASAB	-0.047807*	-0.035549*	-0.050112*	0.049070	-0.050869*	0.000141
	(0.0276)	(0.0227)	(0.0150)	(0.4430)	(0.0169)	(0.0577)
ΔHASR	-0.022325	-0.020980	-0.021578	-0.015907	-0.021994	-0.000015
	(0.0708)	(0.1670)	(0.0765)	(0.4362)	(0.0758)	(0.2030)