Available online at www.sciencedirect.com
ScienceDirect
J. Finan. Intermediation 16 (2007) 117–149
Journal of Financial Intermediation
www.elsevier.com/locate/jfi

Earnings announcements, private information, and strategic informed trading

Jin-Wan Cho *

Korea University Business School, 5 Ka 1, Anam-Dong, Sungbuk-Ku, Seoul 136-701, South Korea

Received 26 March 2004

Abstract

In this paper, we estimate and test a multi-period model of strategic informed trading developed by Foster and Viswanathan [Foster, F.-D., Viswanathan, S., 1996. Strategic trading when agents forecast the forecasts of others, J. Finance 51, 1437–1478]. We employ the GMM using intertemporal patterns of price, trading volume and market depth, leading up to the earnings announcements made by NYSE firms. We find that multiple informed traders with heterogeneous private signals trade prior to the announcements. In addition, by comparing the results from daily and intra-day estimations, we find that the number of informed traders increases while the intensity of liquidity trading decreases, and that the adverse selection problem becomes more pronounced as the announcements approach.

JEL classification: D82; C52; G14

Keywords: Market microstructure; Strategic informed trading; Earnings announcements; Informed trading

1. Introduction

This study evaluates the performance of stylized market microstructure models during periods leading up to earnings announcements. A typical market microstructure model assumes that informed traders exist, trading with a market maker in the presence of liquidity traders. As the informational content of a corporate event is initially known only to a small number of insid-

* Fax: +82 2 922 7220.
E-mail address: npvcho@korea.ac.kr.

1042-9573/$ – see front matter
doi:10.1016/j.jfi.2006.10.001

ers and/or investors, the period leading up to a corporate event is considered as an appropriate sample period for the testing of these models.

Several market microstructure studies attempt to analyze the influence on trading costs induced by possible informed trading during an event period. The majority of these studies attempt to characterize the observable measures of liquidity, such as spread or depth. For example, Krinsky and Lee (1996) estimate the components of spread surrounding earnings announcements. They report that the adverse selection component of spread increases during both pre- and post-event periods. Lee et al. (1993) assess both the spread and depth around earnings announcements, and determine that spreads widen and depths fall in anticipation of earnings announcements.[1] On a different track, Seppi (1992) constructs a model of block trading near earnings announcements, and finds that these trades reflect information contained in the upcoming earnings announcements. More recently, Chae (2005) investigates trading volume and price sensitivity around a number of types of corporate announcements. He finds mixed results as to trading volume, but finds that price sensitivity increases prior to all types of announcements. These papers collectively provide evidence that trades predicated on private information occur prior to the announcements, and also that liquidity providers reflect these anticipated adverse selection problems in their pricing function. In this paper, intertemporal patterns of price, trading volume and liquidity prior to earnings announcements are used to examine the ramifications of informationally motivated trades.

This paper differs from the above-mentioned studies in that structural model estimation is employed. The structural estimation of market microstructure models provides a number of advantages. Most importantly, it allows explicit estimation of the exogenous parameters of theoretical models. In addition, it helps provide insight into the fundamental sources of price-volume formation processes, and to measure the relative importance of these sources. The primary model used in our estimations is a model of speculative trading between a risk-neutral market maker and heterogeneously informed traders, as developed by Foster and Viswanathan (1996). The informed traders not only compete with one another for trading profits, but also learn about the other traders' information from past orders. Each of the traders is strategic, in that he possesses long-lived private information, and places orders to maximize his expected current *and* future profits. This model is a generalization of two simpler models: that of Holden and Subrahmanyam (1992), which features multiple traders but with homogeneous information, and the multi-period Kyle (1985) model, which involves a monopolistic informed trader. The Foster and Viswanathan (1996) model is the sole example among the Kyle-variant models which is able to explain the empirically stylized findings that market depth decreases as the time of information resolution approaches.[2] Testing the simpler models against the general model assists in learning the restrictiveness of the added assumptions, thereby providing insights as to the informationally motivated trades in an environment in which a certain degree of informed trading is likely to exist.

Although structural estimation clearly possesses some merits, some drawbacks emerge when the method is applied to market microstructure models. The principal drawback is the difficulty of estimation itself, which can be attributed to the inherent unobservability of state variables. For example, an econometrician can observe neither private information nor liquidity trading. In this paper, the unobservability problem is overcome by deriving moment conditions on the

[1] Morse and Ushman (1983) and Venkatech and Chiang (1986) investigate the changes in the spread near earnings and dividend announcements. Both studies document that the spread increases at least on the day of announcement.

[2] Back et al. (2000) developed a continuous time version of Foster and Viswanathan (1996). Even in the continuous time framework, the theoretical implications are essentially the same as in discrete time framework.

observable price volatility and trading volume as functions of the exogenous parameters of the model, and then applying the Generalized Method of Moments (GMM) to estimate the models. In doing so, we are able to estimate the model without measuring private information or liquidity trading.

The second drawback is that the Kyle-variant models are highly structured, and accordingly they are likely to be miss-specified. For example, these models do not consider the arrival of public information during trading periods. Instead, they assume that a one time public announcement at the terminal period reveals all the private information. These models further assume that no new pieces of public information, which the informed traders are not aware of at the beginning, are allowed to be revealed during the trading periods. In the real world, however, the private information can be revealed gradually via a series of public announcements. More often than not, new pieces of public information affect the market. Now, with respect to the timing of information acquisition, these models assume that the informed traders acquire their private signals simultaneously. Furthermore, the number of informed traders is assumed to remain unchanged during the trading periods. It is obvious that these assumptions are not realistic, and the actual data should reflect the effects arising from these realistic trading environments assumed away by the models. If we try to accommodate these realistic factors into the model, both the equilibrium order submitting strategy and pricing rule are affected in a complicated manner. Unlike the unobservability problem, the problem arising from the misspecification is extremely difficult to circumvent, and remains as one of the primary limitations of this study. We have, however, attempted to shed light on some of these issues within the limits of this study, via the estimation of the models using both intra-day and daily data.

Recently, several structural estimations of market microstructure models have been successfully undertaken. Among others, Easley et al. (1996) estimate the probability of informed trading by working with the number of buys and sells during a trading day. They find that the probability of informed trading is lower for high-volume stocks. Easley et al. (1997) further include the number of no-trades, within a comparable framework. In addition, Foster and Viswanathan (1995) develop a model in which private information is modeled as a latent variable, and test stylized empirical findings regarding the relationship between trading volume and price volatility. They impose a specific structure on the information process, and reject the model. Cho and Krishnan (2000) exploit the fact that the true value of a futures contract is known ex post, and then estimate Hellwig's noisy rational expectations equilibrium model using S&P 500 futures contracts. They find that, in the index futures market, the principal source of uncertainty is liquidity trading, rather than the diversity of private signals. Caballé and Krishnan (1995) estimate a model involving multiple securities and short-lived information. They establish moment restrictions on prices and order flows, and estimate the model using GMM and maximum-likelihood estimation. One important contribution of this paper is that the model being estimated and tested features multiple informed traders acting strategically using long-lived information. We use the data generated by an informational event, i.e., earnings announcements, and explicitly investigate the nature of the signals received by the informed traders prior to the announcements.

Accordingly, this paper attempts to answer several questions, including the following:

(i) Are there informationally motivated trades close to the time of earnings announcements?
(ii) Are there multiple informed traders?
(iii) If so, do they trade on homogeneous information?
(iv) If the information is heterogeneous, to what degree are the signals correlated?

(v) How much of the trading is conducted by informed traders, liquidity traders, and the market maker? and
(vi) To what degree do informed traders and liquidity traders contribute to price volatility?.

The major findings of this study are (i) multiple informed traders with heterogeneous signals appear to trade close to the earnings announcements, (ii) as the time of the announcements approaches, the number of informed traders increases, the adverse selection problem intensifies, and the liquidity trading decreases, (iii) the structure of the information, i.e. whether it is homogeneous or heterogeneous, is of greater importance than simply the number of informed traders, in explaining the trading activities close to the time of the announcements. Finally, (iv) the contribution of informed trading to trading volume, as well as price volatility, increases as the time of the announcements approaches.

The subsequent sections of this paper are organized as follows. Section 2 introduces the models and their moment restrictions for GMM estimation. Section 3 describes the data and the price and trading volume normalization process. Section 4 discusses the estimation and testing strategies. Section 5 presents the empirical findings, and Section 6 concludes with a summary.

2. Models

In this section, we introduce three models used for estimations: Foster and Viswanathan (1996) (hereafter referred to as FV), Holden and Subrahmanyam (1992) (hereafter, HS) and Kyle (1985). FV is the most general model among those derived from Kyle's (1985) multi-period model.[3] In FV, multiple informed traders trade using heterogeneous information. The heterogeneity of the information is modeled by assuming that the informed traders receive not identical, but correlated signals regarding the terminal value of the asset. The informed traders are strategic in that they compete with each other for profits that they expect to make not only from current trading, but also from future trading. In addition, the informed traders learn about the other traders' private signals by observing the history of net orders. This model serves as the unrestricted model for this paper.

The other two models, HS and Kyle, are tested against FV. These models are restricted models, in that they are nested in FV. In restricting the correlation characterizing the heterogeneity of private information to one, FV is reduced to a model involving homogeneously informed multiple traders. Models along the lines of HS and Foster and Viswanathan (1993b) fall into this category. If it is further assumed that only one informed trader exists, FV is reduced to Kyle. The following provides a brief description of the models and their implications on moment restrictions on price volatility and trading volume.

2.1. Description of the models

In FV, three types of risk-neutral traders exist; a market maker, M informed traders, and a number of liquidity traders. There are T trading periods, characterized by the following event time-line. Both the market maker and the informed traders initially share common priors regarding the distribution of the liquidating value of the asset. Before trading starts, i.e. at time 0,

[3] Cho (2004), HS, Foster and Viswanathan (1993a), and FV developed Kyle-variant models comparable to the present model. Cho (2004) assumes that private information evolves during trading periods, while the others assume multiple informed traders.

information concerning the mean of the distribution is revealed to the informed traders. Each of the informed traders receives a private signal that is correlated to, but may well be different from those of the others. Formally, let the liquidating value be $\tilde{S}$. The ith informed trader receives a signal s_{i0} at the start of trading. Further, assume that $\tilde{S} = p_0 + \sum_{i=1}^{M} s_{i0}$, where p_0 is the common prior about the mean of the terminal value of the asset. The signal vector for all informed traders, $(s_{10}, s_{20}, \ldots, s_{M0})$, is drawn from a multivariate normal distribution. The mean of each signal is zero. Both the variance of each signal and the covariance between any two signals are identical across informed traders. The variance and covariance of signals are denoted as Λ_0 and Ω_0, respectively. In the context of our empirical analysis, this signal can be viewed as private information regarding the upcoming earnings announcement.

Once trading starts, four events occur during each trading period t, for $t = 1, \ldots, T$.[4] First, each of the informed traders updates valuation of the asset by observing the orders of the previous period. In doing so, the informed traders learn about other traders' information, which was partially revealed in past orders. Second, the informed traders submit orders which represent changes in their demand. They act strategically in that when submitting their orders, they attempt to maximize the expected sum of *current and future* profits.[5] Third, the liquidity traders arrive at the market and submit their orders. The market maker observes only the net order, which is the sum of the orders from the informed traders and the liquidity traders. Finally, the market maker sets the price so that it is equal to the expected liquidating value of the asset, and then the market clears at this price.

The exogenous parameters of the model estimated are the number of the informed traders, M, the initial informativeness of the price, Σ_0,[6] the parameter that characterizes the correlation between individual signals, χ, and the amount of liquidity trading, σ_u.[7] Given the values of these exogenous parameters, all of the endogenous parameters, including the parameters characterizing liquidity in the market, λ, and trading aggressiveness, β, are specified completely. Proposition 1[8] in FV shows the recursive linear Markov equilibrium that represents the endogenous parameters as functions of the exogenous parameters. More detailed discussion on FV is provided in Appendix A.

In FV, in which the informed traders receive correlated signals, it is necessary to define two second moments of the distribution of private signals: the conditional covariance between two private signals, Ω_t, and the conditional variance of each of the private signals, Λ_t. One of the

[4] In the empirical analysis, we use two specifications. First, in the intra-day analysis, we use price changes and trading volumes leading up to earnings announcements in 30-minute intervals. In this set-up, each trading period can be interpreted as a 30-minute interval. Second, in the daily analysis, we use daily price changes and trading volumes leading up to the announcements. In this specification, therefore, each period is a trading day.

[5] An informed trader with short-lived information, as in Admati and Pfleiderer (1988), Hughson (1990) and Caballé and Krishnan (1995) is myopic in that only current expected profits are maximized. This means that an informed trader perceives no benefit in limiting the resolution of information for future profits. The informed trader in the present model may trade less aggressively than would a myopic informed trader, particularly early in trading periods, in order to obtain larger *overall* profits.

[6] In FV, Σ_0 is the conditional variance of the *sum* of the private signals. In the present paper, however, Σ_0 is defined as the conditional variance of the *average* of the private signals. In other words, it is required to multiply Σ_0 in the present paper by M^2, in order to obtain Σ_0 in FV. This presents no major problems, as both the sum and the average are sufficient statistics for the liquidating value of the asset. This re-scaling will allow aligning of the meaning of Σ_0 across all of models estimated.

[7] In the original FV model, there is another exogenous parameter: the initial prior. In the empirical component, instead of estimating the parameter, the prevailing price is substituted at the beginning of each trading period for the initial prior.

[8] In Appendix A, this proposition is renamed Proposition A.1.

exogenous parameters, χ, is defined as the difference between Λ_t and Ω_t, i.e. $\chi \equiv \Lambda_t - \Omega_t$,[9] and is used to measure the closeness between FV and HS. It is important to note that as FV approaches HS, the correlation between the signals approaches one. As a result, χ approaches 0. Therefore, a small χ indicates that the underlying model is close to HS. At the extreme, if χ becomes 0, the FV model is reduced to HS. This fact is used in the testing of FV against HS. Similarly, the number of informed traders, M, reflects the closeness between HS and Kyle. As HS approaches Kyle, M approaches 1. Therefore, in assigning a value of 1 to M, HS can be tested against Kyle.

2.2. Moment restrictions on price volatility and trading volume

In this sub-section, we present a series of propositions that summarize the moment restrictions on price volatility and trading volume implied by each of the three aforementioned models. These conditional moments function as the orthogonality conditions in the GMM procedure. First, the following proposition presents the conditional means of price volatility, trading volume and the inverse market depth parameter, λ, implied in FV.

Proposition 1 *(FV: General Model). Given the history of prices, the conditional means of price volatility and trading volume for $t = 1, 2, \ldots, T$ are*

$$E\left(\Delta p_t^2 \mid p_{t-1}, \ldots, p_1\right) = M\lambda_t^{FV^2}\beta_t^{FV^2}\Lambda_{t-1} + M(M-1)\lambda_t^{FV^2}\beta_t^{FV^2}\Omega_{t-1} + \lambda_t^{FV^2}\sigma_u^2, \quad (1)$$

$$E(TV_t \mid p_{t-1}, \ldots, p_1) = \sqrt{\frac{1}{2\pi}}\left\{M\sqrt{\beta_t^{FV^2}\Lambda_{t-1}} + \sigma_u + \sqrt{M\beta_t^{FV^2}\Lambda_{t-1} + M(M-1)\beta_t^{FV^2}\Omega_{t-1} + \sigma_u^2}\right\}, \quad (2)$$

and

$$E\left(\lambda_t^{FV} \mid p_{t-1}, \ldots, p_1\right) = \lambda_t^{FV} = \frac{\beta_t^{FV}\theta}{\sigma_u^2}\Sigma_t^{FV}, \quad (3)$$

where θ is a constant defined in Eq. (A.24) in Appendix A.

The proof for (1) and (2) is given in Appendix B. The expected price changes given in (1) have three terms. These measure, respectively, the amount of price change attributable to (i) the degree of residual ‘privateness’ of each of the informed trader’s private information, (ii) the degree of learning by the informed traders regarding other informed traders’ information, and (iii) liquidity trading.

In order to determine the conditional expectation of trading volume in (2), it is first required to derive an appropriate definition of trading volume. As in Admati and Pfleiderer (1988), there are three types of contributors to trading volume: M informed traders, liquidity traders, and the market maker. The trading volume is defined as one-half of the sum of the absolute values of orders submitted or received by the above market participants. For example, suppose that two informed traders want to buy 5 and 3 shares, and that the liquidity traders want to sell 6 shares. Then, the market maker observes the net order to buy 2 shares. The market maker then sets an equilibrium price and absorbs the order. In this example, the total trading volume will be 8 shares, which is precisely one half of $(5+3)+6+2$. The case in which both the informed traders and the

[9] One of the properties of FV is that χ is constant for all t.

liquidity traders submit orders in the same direction can be conducted in a similar fashion. In (2), the first term represents the trading volume generated by the informed traders, the second term, by the liquidity traders, and the last term, by the market maker. One of the primary advantages of structural model estimation is that, once estimation is achieved, the relative importance of the sources that determine price innovation and trading volume can be explicitly specified. This is presented in subsequent sections.

In (3), we present the expected value of the inverse market depth parameter, λ_t^{FV}. Since this is not a stochastic variable, its expected value should be the parameter itself. The exact form is from (A.15) in Appendix B. In this equation, θ is the regression coefficient of the true value on the average signal. Under the HS specification, this parameter becomes M. As illustrated later, depending on the values of the exogenous parameters, the market depth can either increase or decrease as time approaches the terminal trading period. Even though the FV model allows more general intertemporal patterns for λ_t^{FV}, the essential characteristics implied in Kyle or HS specifications are well persevered. For example, λ_t^{FV} is positively related to β_t^{FV} and Σ_t^{FV}, while negatively related to σ_u. Therefore, just as in Kyle or HS, the market depth increases as liquidity trading intensifies, informed traders trade less aggressively, or the degree of informational asymmetry becomes weaker.

Recall that by restricting some parameter values, either HS or Kyle can be obtained. In the following propositions, we derive the conditional means of price volatility and trading volume implied by HS and Kyle.

Proposition 2 *(HS: Homogeneous Private Information). If $\chi = 0$, i.e. the informed traders receive homogeneous private signals, then*

$$E\left(\Delta p_t^2 \mid p_{t-1}, \dots, p_1\right) = M^2 \lambda_t^{HS^2} \beta_t^{HS^2} \Sigma_{t-1} + \lambda_t^{HS^2} \sigma_u^2, \tag{4}$$

$$E(TV_t \mid p_{t-1}, \dots, p_1) = \sqrt{\frac{1}{2\pi}} \left\{ M\sqrt{\beta_t^{HS^2} \Sigma_{t-1}} + \sigma_u + \sqrt{M^2 \beta_t^{HS^2} \Sigma_{t-1} + \sigma_u^2} \right\}, \tag{5}$$

and

$$E\left(\lambda_t^{HS} \mid p_{t-1}, \dots, p_1\right) = \lambda_t^{HS} = \frac{\beta_t^{HS} M}{\sigma_u^2} \Sigma_t^{HS} \tag{6}$$

for $t = 1, 2, \dots, T$.

Proposition 3 *(Kyle: Monopolistic Informed Trader). If there is only one informed trader receiving private information, then*

$$E\left(\Delta p_t^2 \mid p_{t-1}, \dots, p_1\right) = \lambda_t^{Kyle^2} \beta_t^{Kyle^2} \Sigma_{t-1} + \lambda_t^{Kyle^2} \sigma_u^2, \tag{7}$$

$$E(TV_t \mid p_{t-1}, \dots, p_1) = \sqrt{\frac{1}{2\pi}} \left\{ \sqrt{\beta_t^{Kyle^2} \Sigma_{t-1}} + \sigma_u + \sqrt{\beta_t^{Kyle^2} \Sigma_{t-1} + \sigma_u^2} \right\}, \tag{8}$$

and

$$E\left(\lambda_t^{Kyle} \mid p_{t-1}, \dots, p_1\right) = \lambda_t^{Kyle} = \frac{\beta_t^{Kyle}}{\sigma_u^2} \Sigma_t^{Kyle} \tag{9}$$

for $t = 1, 2, \dots, T$.

2.3. Intertemporal patterns of price volatility, trading volume, and the inverse market depth parameter

In the estimation of the models, we use three observable series as moment conditions: price volatility, trading volume and the inverse of market depth parameter, λ. A commonality shared by these three models is that each of the informed traders acts strategically, yielding specific patterns in price volatility, trading volume, and market depth parameter. These patterns depend on the level of correlation in the informed traders' private signals, as well as on the number of competing informed traders. In this subsection, we compare the intertemporal patterns of price volatility, trading volume, and market depth implied by these models. In order to conduct numerical evaluation, the values of the exogenous parameters must first be established, and then the endogenous parameters are solved. For FV, we set the values of the exogenous parameters, (M, Σ_0, σ_u), to (5, 1, 1), while changing the χ value from 0.2, 1 to 2. The various values of χ imply that the correlation between the individual signals is 0.44, 0, and -0.11, respectively. We then solve for the endogenous parameters, λ_t, β_t, Λ_t, and Ω_t. For HS and Kyle, χ is set to 0. When χ is 0, it can be demonstrated that the endogenous parameters, Λ_t and Ω_t, are both equal to Σ_t. For Kyle, M is further set to 1.[10] Since we estimate the model over six trading periods in the empirical analysis, we will also examine the numerical patterns over six trading periods.[11]

With the values of the exogenous parameters given above, we solve for the endogenous parameters by using the recursive linear Markov equilibrium given in Eqs. (A.13) to (A.29) in Proposition A.1 of Appendix A. The key to solving the system is the variance of the signals, Λ. Using the values of the exogenous parameters, the initial variance, Λ_0, can be evaluated by using the relationship, $\Lambda_0 = (\Sigma_0 + M(M-1)\chi)/M^2$.[12] Since Proposition A.1 can be solved only via backward induction, the precise Λ_T that corresponds to Λ_0 must be determined. Exploiting the uniqueness of the equilibrium, first, a search is made numerically for an internally consistent Λ_T. Once Λ_T is obtained, the entire set of endogenous parameters is solved by backward induction. The series of inverse market depth parameter, λ_t, is obtained via this process.

From the values of the endogenous parameters, the simulated series for trading volume, price volatility and market depth are produced by utilizing the propositions from the previous section. Figures 1 through 3 plot simulated trading volume, price volatility, and inverse of market depth, respectively.

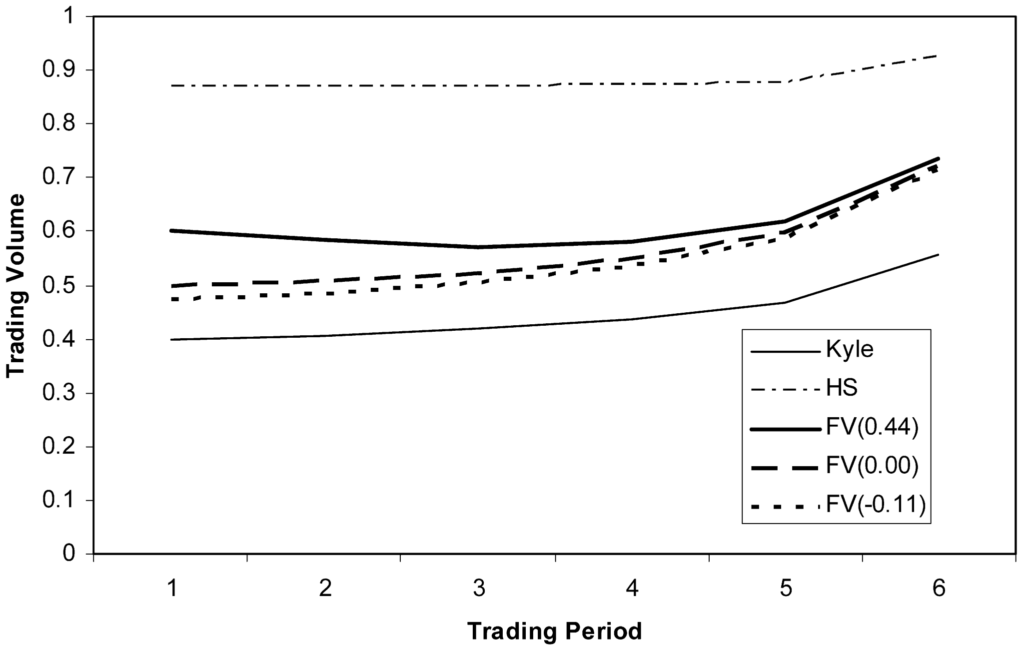
First, in Kyle, both price volatility and trading volume rise moderately as the time of the announcements approaches. However, the inverse market depth parameter, λ, decreases moderately, thereby implying that liquidity improves as the time of the announcements approaches. This reflects the controlled revelation of information by the monopolistic informed trader. The informed trader trades less aggressively during the early stages of the trading periods, in order to maximize the overall expected trading profits. For HS, however, price and λ change most significantly during the first round of trading, whereas trading volume remains relatively flat throughout the trading periods, with a slight increase occurring as the time of the announcement draws nearer.

---

10 In this numerical exercise, the value of Σ_0 that renders the endogenous parameters from HS consistent with those from FV is 5. When using this parameter value, however, the price volatility for period 1 becomes too high (at over 20) to make the patterns for the other cases recognizable. Therefore, this value has been normalized by the number of informed traders. This change would not result in any material changes in the basic shape of the intertemporal patterns. In fact, the expected trading volume will still be the same, even with Σ_0 set to 5.

11 The basic shape of the patterns remains the same, even with more trading periods.

12 This relationship is given in Equation (8) of Foster and Viswanathan (1996).

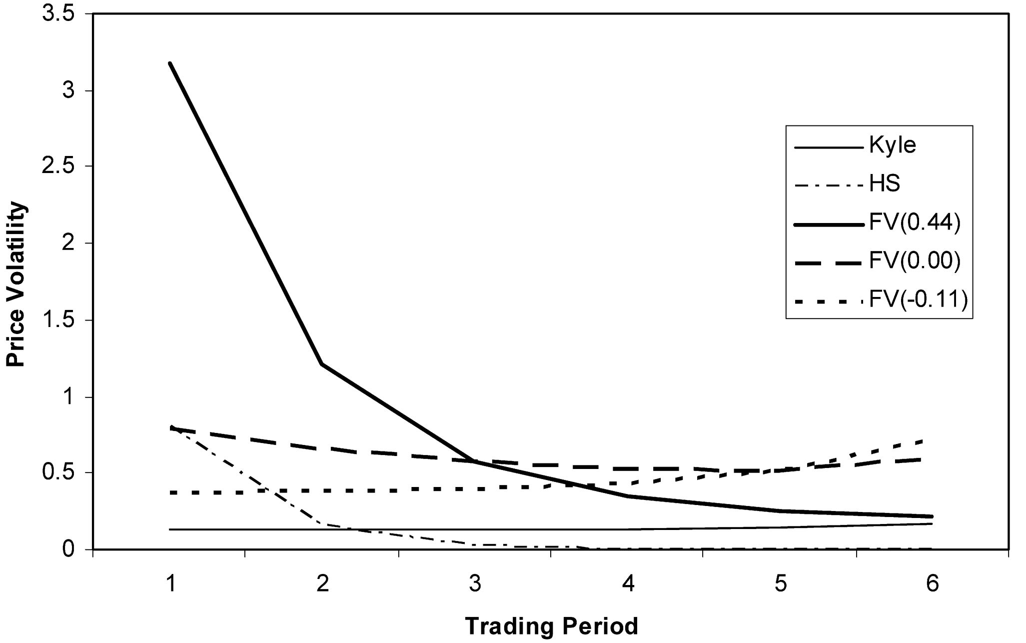


In this figure, we compare the intertemporal patterns of trading volume implied by Kyle, HS and FV with different levels of correlation between private signals. In order to conduct a numerical evaluation of the models, it is first required to solve for the endogenous parameters for each of the trading periods. For FV, the endogenous parameters are λ_t, β_t, Λ_t, and Ω_t. When solving for the endogenous variables, we set the values of the exogenous parameters, (M, Σ_0, σ_u) to $(5, 1, 1)$ respectively, while changing the value of χ from 0.2, 1 and to 2. The various values of χ imply that the correlations between individual signals are 0.44, 0, and -0.11, respectively. For HS and Kyle, χ is set to 0. For Kyle, we set M to 1.

Fig. 1. Simulated trading volume.

In contrast to original model Holden and Subrahmanyam's (1992), the trading volume for HS in this paper does not reflect the well-known 'rat race' pattern. In other words, trading volume does not reach the peak in the first round of trading. This arises from the fact that the definitions of trading volume used in HS and this paper are different. In the original HS, what is plotted is the expected *orders* submitted by *the informed traders* conditional on *their signal*, i.e. $E(M \times \Delta x_{it} \mid v)$, where Δx_{it} represents the orders submitted by the ith informed trader in period t and v, the signal acquired by the informed traders. Therefore, the orders are likely to be either buy or sell, depending on the sign of the private signal. Their numerical illustration is for the case when the private signal is positive. In this case, recognizing that the other informed traders would have received the same signal, the informed traders then simultaneously turn in buy orders. As a result, the "rat race" pattern of buy orders is exhibited.

However, in this paper, we plot the expected *trading volume* conditional on the *past prices*, i.e. $E(|\Sigma \Delta x_{it}| + |\Delta y_t| + |u_t| \mid p_1 \dots p_{t-1})$, where Δy_t are the net orders received by the market maker, u_t, the orders submitted by the liquidity traders. The statistical properties of trading volume defined in this way are quite different from that defined in the original HS. The key to the difference is the conditioning variable. In order to have a 'rat race', the informed traders have to turn in a rush of buy (or sell) orders up front given p_0 at the first round of trading. This however is not possible in this paper, since the informed traders do not rush on public information. In fact,



In this figure, we compare the intertemporal patterns of price volatility implied by Kyle, HS, and FV with different levels of correlation among private signals. In order to conduct a numerical evaluation of the models, it is first required to solve for the endogenous parameters for each trading period. For FV, the endogenous parameters are λ_t, β_t, Λ_t, and Ω_t. When solving for the endogenous variables, we set the values of the exogenous parameters, (M, Σ_0, σ_u) to (5, 1, 1) respectively, while changing the value of χ from 0.2, 1 and to 2. The various values of χ imply that the correlations between individual signals are 0.44, 0, and -0.11, respectively. For HS and Kyle, χ is set to 0. For Kyle, we set M to 1.

Fig. 2. Simulated price volatility.

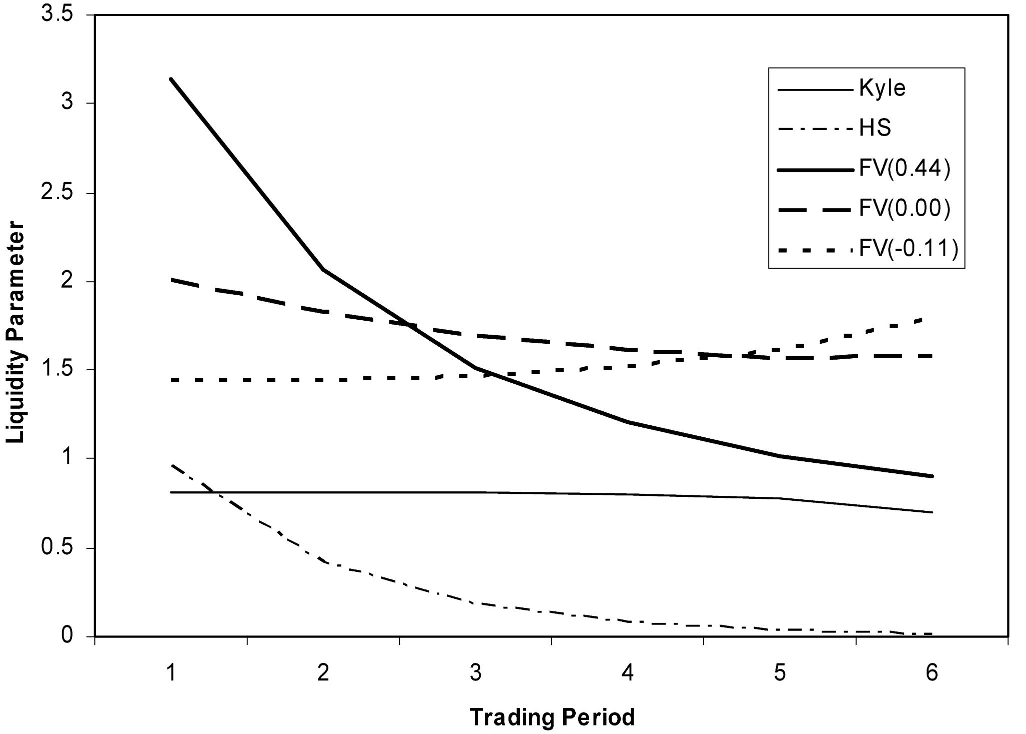
even in HS, if the conditioning variable were the past prices, the expected trading volume defined in their specification would yield a 0 expected trading volume.[13] As a result, the expected trading volume for the first period is not as great as that observed by Holden and Subrahmanyam.

The pattern of price volatility, however, reflects active trading by the informed traders during the first round of trading similar in the original HS model.[14] As presented in Fig. 2, the majority of the price changes are expected to occur within the first round. This implies that the price at the beginning of the second round already reflects the majority of the information contained in the private signal.

In the Kyle and HS models, the shapes of the intertemporal patterns are robust to changes in the exogenous parameter values. In the FV model, however, these patterns change with changes in the correlation among private signals. As the correlation changes from positive to negative, the intertemporal pattern of trading volume changes from a U-shape to a monotone increase. The changes in the patterns of price volatility and λ are even more conspicuous. With positively

[13] Note that $E(M \times \Delta X_{it} \mid p_{t-1}) = E(M\beta_t(v - p_{t-1}) \mid p_{t-1}) = 0$, as the next period's signed net order flow conditional on public information is equally likely to be a net buy or a net sell.

[14] The price volatility in FV (0.44) also reveals a disproportionate amount of price discovery in the first round of trading. Note that this is the case in which the private signals are positively correlated, and also that the information structure of FV is somewhat similar to that of HS.



In this figure, we compare the intertemporal patterns of the inverse market depth parameter, λ implied by Kyle, HS, and FV, with different correlation levels among private signals. In order to conduct a numerical evaluation of the models, it is first required to solve for the endogenous parameters for each of the trading periods. For FV, the endogenous parameters are λ_t, β_t, Λ_t, and Ω_t. When solving for the endogenous variables, we set the values of the exogenous parameters, (M, Σ_0, σ_u) to (5, 1, 1) respectively, while changing the value of χ from 0.2, 1 and to 2. The various values of χ imply that the correlations between individual signals are 0.44, 0, and −0.11, respectively. For HS and Kyle, χ is set to 0. For Kyle, we set M to 1.

Fig. 3. Simulated inverse market depth parameter, λ.

correlated signals, price volatility and λ both decrease monotonically over time, whereas with negatively correlated signals, price volatility and λ exhibit a monotonic increase. The latter case, in which λ increases as the announcement time approaches, can be generated neither by Kyle nor by HS. Therefore, only FV is able to explain the empirically documented decrease in market depth occurring near the time of announcement.[15]

3. Data

The sample moments of price volatility and trading volume are constructed from the prices and trading volumes leading up to quarterly earnings announcements made by NYSE firms during the months of July and October of 1998, and January and April of 1999. The announcement data are obtained from the PR Newswire, in which each announcement is time-stamped to the nearest minute. We run two sets of estimation: intra-day and daily. In estimating the model over

[15] See Krinsky and Lee (1996) and Chae (2005) for example.

different time intervals, we are able to observe whether the estimated number of traders and the correlation among signals differ for the two time intervals. In the process of explaining the results, the information acquisition process and the model itself can be better understood. The data regarding daily series are acquired from the CRSP database, whereas those regarding intra-day transaction prices and trading volumes are obtained from the TAQ database. In total, 2685 earnings announcements are in the original PR Newswire database. The final number of announcements employed in this paper is 2491 for the intra-day analysis, and 2518 for the daily analysis. Missing price or volume data is attributable to the omission of the announcements.

In the theoretical model, the terminal trading period occurs when private information is revealed to market participants. In other words, this is the time at which the price reflects the 'true' value of the asset. The timing of the actual announcement, however, cannot be regarded as the time at which the content of the earnings announcement is fully reflected in the price. The market would probably require some time to assess the nature of the announcement. Accordingly, in this paper, it is assumed that at least 30 minutes are required for the price to correctly reflect the informational content of an announcement. As a result, if an announcement was made after 3:30 pm EST, we take, for the intra-day estimation, the next trading day's price at 9:30 a.m.,[16] and for the daily estimation, we take the next day's closing price as the terminal price. Once the time of the terminal price is determined, we obtain 5 prices, trading volumes, and market depths, leading up to the terminal price. For the intra-day estimation, this series is taken over 30-minute intervals, while for the daily estimation, we use the daily series.

It is important to note that the moment restrictions provided in the preceding section are expressed in terms of prices rather than as returns. This gives rise to the necessity for the normalization of the price and volume series, as each firm has a different stock price level. This normalization was accomplished as follows. For the sake of simplicity, let 'period 6' represent the terminal period. Then, the value of a '*variable*' of t periods prior to the earnings announcement is denoted as the period $6 - t$ '*variable*'. For example, the stock price 5 periods prior to the announcement would be designated 'period 1 stock price.' Now, in order to normalize the prices, we first set the period 0 price to 100.[17] We then take the price series from period 1, and generate a set of normalized stock prices. Once all the normalized prices have been calculated, the price volatilities are calculated by taking the square of the difference between adjacent normalized prices.

Once prices are normalized, trading volume should also be normalized. This is necessary since the original trading volume is measured with respect to the original prices, i.e. prices that are not normalized. Therefore, using the original trading volumes without normalization can result in problems. For example, suppose that actual period 0 price and trading volume were \$50 and 100 shares, respectively, and period 1 price and trading volume, were \$55 and 300 shares, respectively. Then, the normalized prices for the first two periods would be 100 and 110 respectively. In this case, the original trading volumes 100 and 300 may not be used, since this would mean that the dollar trading volumes are \$10,000 and \$33,000 respectively, when in fact the actual dollar trading volume is \$5000 and \$16,500. To circumvent this problem, the trading volume is normalized in the following way. First, the period 0 stock price is divided by the period 0 trading volume. In this example, this number will turn out to be 0.5. Second,

[16] The reason that opening prices are omitted is two-fold. First, they are not determined by the specialists. Second, they have been shown to behave differently from prices at other times.

[17] One of the benefits of setting the initial price at 100 is that estimated parameters related to price changes can be interpreted as percentages of prices.

the normalized trading volumes are obtained by multiplying this number by the actual trading volumes. For the above example, the period 0 and period 1 normalized trading volumes would be 0.5 (100) = 50, and 0.5 (300) = 150 shares. Note that, coupled with the normalized prices, this procedure preserves the dollar trading volumes ($5000 and $16,500 respectively), and the fact that period 1 trading volume is three times larger than period 0 trading volume.

After we obtain the normalized price and volume series, we evaluate the series of market depth by dividing the absolute price change by the trading volume, i.e. $InvD_t \equiv |\Delta p_t / TV_t|$. A larger *InvD* implies that the market is less liquid, so that the price must move more significantly in order to accommodate one unit of trading volume. Therefore, this series can serve as a natural sample moment for the inverse market depth parameter, λ.

To further analyze these data, we group the whole sample into three sub-samples, according to two criteria. First, we create sub-samples by sorting on firm size. Since large firms tend to attract heavy coverage by the analysts, it is conjectured that the adverse selection problem is less severe for these firms. For small firms, however, it is conjectured that the opposite would be the case. We use the market capitalization as of the end of the month preceding the announcement as the basis of sorting. For example, for firms announcing in July 1998, we take the firm size at the end of June 1998. The resulting sub-samples are referred to as 'Large Firms', 'Medium Firms', and 'Small Firms'.

Second, we group the announcements by absolute price changes prior to the trading periods. For the daily series, we utilize a one-day window, while for the intra-day series, we employ a one hour window. In the notation, this corresponds to $|p_0 - p_{-1}|$ for the daily series, and $|p_0 - p_{-2}|$ for the intra-day series. The resulting sub-samples are referred to as 'Large Move', 'Medium Move', and 'Small Move'. The announcements included in the 'Large Move' sample can then be regarded as announcements made in the midst of great uncertainty, whereas the announcements included in the 'Small Move' sample are those made in a less uncertain environment. Tables 1 and 2 show the summary statistics of the data for the intra-day and daily estimations.

First of all, for both intra-day and daily series, price volatility (ΔP^2) and trading volume (*TV*) are highest for the terminal period across all samples. The inverse market depth (*InvD*), however, reveals somewhat different patterns for intra-day and daily series. In particular, for the majority of the intra-day samples, *InvD* exhibits a U-shape as the time of the announcement approaches. For the daily series, however, the patterns vary depending on the samples. In casually investigating these patterns, it can be conjectured that only FV is a reasonable candidate in explaining the trading environment prior to the earnings announcement.[18] This is due to two reasons. First, since HS predicts that the price volatility is highest at the beginning, it may not serve as the reasonable model in that the actual price volatility is highest at the end. Furthermore, Kyle may not be rejected by HS since the sharp decrease in the price volatility only occurs with more than one informed traders with perfectly correlated signals. Second, HS and Kyle predict that liquidity improves towards the end of the trading rounds, implying *InvD* is decreasing. The actual *InvD*, however, does not reveal a decreasing pattern for the majority of the samples. Since only FV can generate increasing, decreasing, U-shaped, or reverse U-shaped patterns of *InvD*, the patterns of *InvD* also suggest that only FV is a reasonable model to apply.

[18] As was discussed in the introduction, the true data generating process may well not originate from the narrow specifications of Kyle-variant models. 'Identifying' and testing for the true model is beyond the scope of this paper. Instead, this paper attempts to estimate and test the informational structure inherent to the Kyle-variant models.

Table 1
Summary statistics and sample moments: intra-day series

Trading periods	(M2) Large firms [830]				Medium firms [831]				Small firms [830]			
	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ
1	100.07	0.42	22.28	0.15	100.02	0.55	2.17	1.20	100.01	0.79	0.49	6.66
2	100.07	0.47	22.78	0.15	100.04	0.50	1.97	1.51	100.02	0.95	0.66	6.02
3	100.08	0.49	25.43	0.15	100.04	0.65	2.95	0.99	100.05	1.32	0.65	4.85
4	100.11	0.58	30.44	0.13	100.12	0.66	3.31	1.51	100.01	1.09	0.77	7.84
5	100.16	3.42	48.45	0.13	100.06	4.48	4.34	1.07	100.17	5.63	1.53	8.77
6	100.26	5.55	65.37	0.70	100.36	10.94	6.83	1.17	100.66	10.93	1.87	7.52

Trading periods	Large move [830]				Medium move [831]				Small move [830]				Whole sample [2491]			
	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ
1	100.08	1.97	13.15	7.22	100.02	0.08	9.32	0.72	100.00	0.00	2.60	0.16	100.03	0.68	8.36	2.68
2	100.13	0.94	12.28	3.41	100.00	0.36	10.23	0.63	99.99	0.64	2.97	1.97	100.04	0.64	8.50	2.56
3	100.14	0.99	13.42	1.68	100.00	0.34	12.29	0.57	100.01	1.13	3.57	2.80	100.05	0.82	9.76	2.01
4	100.17	1.05	14.77	2.72	100.06	0.52	14.69	0.52	100.01	0.78	5.24	3.45	100.08	0.78	11.57	3.18
5	100.16	6.38	22.75	3.16	100.11	4.06	23.99	0.57	100.11	2.99	8.38	4.47	100.13	4.48	18.38	3.33
6	100.38	10.72	29.81	3.03	100.27	6.37	32.66	2.48	100.63	10.25	12.99	6.81	100.43	9.11	25.16	3.14

Notes. This table provides summary statistics of the normalized price (P), price volatility (ΔP^2), trading volume (TV) and inverse of market depth series (*InvD*). Each of the periods corresponds to a 30-minute trading interval. ‘Period 6’ contains the earnings announcement, and represents the maturity of the theoretical model. The prices and trading volumes are normalized as described in Section 2. The numbers in brackets, [ ], are the number of observations in the corresponding sample. Each of the numbers demonstrates the cross-sectional averages across announcements. The trading volume is in 1000 normalized shares.

Table 2
Summary statistics and sample moments: daily series

Trading periods	(M2) Large firms [839]				Medium firms [840]				Small firms [839]			
	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ
1	100.09	8.55	567.00	0.017	100.38	10.86	50.98	0.142	100.36	13.98	9.59	1.824
2	100.32	7.89	549.12	0.020	100.72	11.29	51.44	0.132	100.93	15.17	9.94	2.237
3	100.63	9.88	584.75	0.038	101.29	12.85	52.75	0.107	101.27	14.67	10.28	2.631
4	100.79	7.45	543.83	0.028	101.57	12.20	48.01	0.149	101.66	15.32	9.68	1.520
5	100.97	11.70	617.97	0.018	102.15	16.98	56.88	0.129	102.13	16.82	10.71	2.345
6	100.86	19.09	851.23	0.017	102.18	30.88	84.68	0.116	102.64	34.83	15.94	1.695

Trading periods	Large move [839]				Medium move [840]				Small move [839]				Whole sample [2518]			
	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ	P	ΔP^2	TV	λ
1	100.84	31.55	259.94	0.174	99.91	2.14	141.14	0.103	99.98	0.19	89.10	0.080	100.27	11.13	209.25	0.661
2	101.72	21.76	241.69	0.392	100.22	7.44	146.53	0.070	100.12	5.16	88.36	0.089	100.66	11.45	203.56	0.796
3	102.44	21.26	243.34	0.091	100.63	8.77	156.72	0.107	100.31	6.99	113.25	0.117	101.07	12.47	215.99	0.925
4	102.83	20.09	216.95	0.071	100.86	8.01	148.22	0.101	100.61	6.72	104.84	0.071	101.34	11.66	200.56	0.569
5	103.72	24.46	243.16	0.033	101.19	11.36	172.35	0.080	100.83	9.18	118.37	0.088	101.75	15.16	228.59	0.831
6	104.23	36.17	304.49	0.091	101.38	17.04	213.81	0.102	100.83	15.98	148.36	0.100	101.89	28.27	317.37	0.609

Notes. This table provides summary statistics of the normalized price (P), price volatility (ΔP^2), trading volume (TV) and inverse of market depth series (λ). Each period corresponds to a one day trading interval. 'Period 6' contains the earnings announcement, and represents the maturity of the theoretical model. The prices and trading volumes are normalized as described in Section 2. The numbers in brackets, [ ], are the number of observations in the corresponding sample. Each number demonstrates the cross-sectional averages across announcements. The trading volume is in 1000 normalized shares.

Second, when the sub-samples sorted by the firm size are compared, various clear patterns emerge. For both intra-day and daily series, *TV* is largest for the 'Large Firms' sample, second for the 'Medium Firms' sample, and smallest for the 'Small Firms' sample. The order is reversed for ΔP^2 and *InvD*. These are somewhat expected, since larger firms are believed to attract heavier trading volume making the market deeper. With a deeper market, price volatility and the sensitivity of price to orders becomes smaller. In summary, it can be expected that the adverse selection problem would be the smallest in the 'Large Firms' sample. This is analyzed in greater detail by estimating the models in the following section.

4. Estimation and test methodology

We estimate the models via the Generalized Method of Moments (GMM). The FV, HS, and Kyle models are represented by their implications on expected price volatility and trading volume. Although the models' state variables, i.e. private signals and liquidity trading, are all unobservable, these models can be estimated, since the equilibria are expressed in observable price volatility and trading volume.

There are three variables upon which the moment restrictions are constructed: price volatility, trading volume, and the inverse of market depth parameter, λ. The theoretical moments on price volatility, trading volume and inverse market depth are provided in Eqs. (1) through to (9) It is straightforward to obtain the sample moments of price volatility and trading volume. The moment restrictions for λ are derived from the observation that this parameter is typically measured by dividing price change by trading volume. We estimate the models for six trading periods leading up to the earnings announcements. This implies that 18 moment restrictions are obtained: 6 on the price volatility, 6 on trading volume, and the other 6 on λ. More specifically, the moment restrictions are as follows:

$$\begin{pmatrix} E(\Delta p_1^2 \mid p_0) \\ \vdots \\ E(\Delta p_6^2 \mid p_5, \ldots, p_0) \\ E(TV_1 \mid p_0) \\ \vdots \\ E(TV_6 \mid p_5, \ldots, p_0) \\ E(\Delta p_1/TV_1 \mid p_0) \\ \vdots \\ E(\Delta p_6/TV_6 \mid p_5, \ldots, p_0) \end{pmatrix}$$

$$
= \begin{pmatrix}
M\lambda_1^{FV^2}\beta_1^{FV^2}\Lambda_0 + M(M-1)\lambda_1^{FV^2}\beta_1^{FV^2}\Omega_0 + \lambda_1^{FV^2}\sigma_u^2 \\
\vdots \\
M\lambda_6^{FV^2}\beta_6^{FV^2}\Lambda_5 + M(M-1)\lambda_6^{FV^2}\beta_6^{FV^2}\Omega_5 + \lambda_6^{FV^2}\sigma_u^2 \\
\sqrt{\frac{1}{2\pi}}\{M\sqrt{\beta_1^{FV^2}\Lambda_0} + \sigma_u + \sqrt{M\beta_1^{FV^2}\Lambda_0 + M(M-1)\beta_1^{FV^2}\Omega_0 + \sigma_u^2}\} \\
\vdots \\
\sqrt{\frac{1}{2\pi}}\{M\sqrt{\beta_6^{FV^2}\Lambda_5} + \sigma_u + \sqrt{M\beta_6^{FV^2}\Lambda_5 + M(M-1)\beta_6^{FV^2}\Omega_5 + \sigma_u^2}\} \\
\lambda_1^{FV} \\
\vdots \\
\lambda_6^{FV}
\end{pmatrix} \quad \text{for FV,} \tag{10}
$$

$$
\begin{pmatrix}
E(\Delta p_1^2 \mid p_0) \\
\vdots \\
E(\Delta p_6^2 \mid p_5, \ldots, p_0) \\
E(TV_1 \mid p_0) \\
\vdots \\
E(TV_6 \mid p_5, \ldots, p_0) \\
E(\Delta p_1 / TV_1 \mid p_0) \\
\vdots \\
E(\Delta p_6 / TV_6 \mid p_5, \ldots, p_0)
\end{pmatrix}
$$

$$
= \begin{pmatrix}
M^2\lambda_1^{HS^2}\beta_1^{HS^2}\Sigma_0 + \lambda_1^{HS^2}\sigma_u^2 \\
\vdots \\
M^2\lambda_6^{HS^2}\beta_6^{HS^2}\Sigma_5 + \lambda_6^{HS^2}\sigma_u^2 \\
\sqrt{\frac{1}{2\pi}}\{M\sqrt{\beta_1^{HS^2}\Sigma_0} + \sigma_u + \sqrt{M^2\beta_1^{HS^2}\Sigma_0 + \sigma_u^2}\} \\
\vdots \\
\sqrt{\frac{1}{2\pi}}\{M\sqrt{\beta_6^{HS^2}\Sigma_5} + \sigma_u + \sqrt{M^2\beta_6^{HS^2}\Sigma_5 + \sigma_u^2}\} \\
\lambda_1^{HS} \\
\vdots \\
\lambda_6^{HS}
\end{pmatrix} \quad \text{for HS,} \tag{11}
$$

and

$$\begin{pmatrix} E(\Delta p_1^2 \mid p_0) \\ \vdots \\ E(\Delta p_6^2 \mid p_5, \ldots, p_0) \\ E(TV_1 \mid p_0) \\ \vdots \\ E(TV_6 \mid p_5, \ldots, p_0) \\ E(\Delta p_1 / TV_1 \mid p_0) \\ \vdots \\ E(\Delta p_6 / TV_6 \mid p_5, \ldots, p_0) \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1^{Kyle^2} \beta_1^{Kyle^2} \Sigma_0 + \lambda_1^{Kyle^2} \sigma_u^2 \\ \vdots \\ \lambda_6^2 \beta_6^{Kyle^2} \Sigma_5 + \lambda_6^{Kyle^2} \sigma_u^2 \\ \sqrt{\frac{1}{2\pi}} \{ \sqrt{\beta_1^{Kyle^2} \Sigma_0} + \sigma_u + \sqrt{\beta_1^{Kyle^2} \Sigma_0 + \sigma_u^2} \} \\ \vdots \\ \sqrt{\frac{1}{2\pi}} \{ \sqrt{\beta_6^{Kyle^2} \Sigma_5} + \sigma_u + \sqrt{\beta_6^{Kyle^2} \Sigma_5 + \sigma_u^2} \} \\ \lambda_1^{Kyle} \\ \vdots \\ \lambda_6^{Kyle} \end{pmatrix} \quad \text{for Kyle.} \tag{12}$$

In order to apply the GMM, we first obtain the sample moments via the evaluation of the cross-sectional averages of normalized price volatility, normalized trading volume, and inverse market depth, for each of the 6 periods. Then, using Eqs. (10)–(12) as theoretical moments, we estimate the parameters, by minimizing the distance between theoretical moments and sample moments. In doing so, the distance is measured with regard to a weighting matrix. The standard two-stage GMM methodology is then conducted, in which the first stage involves the use of an identity matrix as the weighting matrix, and the second stage uses the optimal consistent weighting matrix, which incorporates the Newey and West (1987) corrections for heteroskedasticity and serial correlation.

A number of issues must be addressed. First, it is important to note that the data used in this work encompasses four quarters' worth of earnings announcements by different companies. In addition, it is assumed that the individual announcements in the data set contain similar information structures, thus making a cross-sectional estimation feasible. Perhaps the best way to estimate a model such as this is to estimate the model using the time-series of earnings announcements by individual companies on the first pass, and then take the averages of the estimates on the second pass. However, in practice, this is an impossible approach, because considering the paucity of earnings announcements issued by a single company, it is impossible to estimate the model on the first pass. In order to ameliorate the difficulty associated with pure cross-sectional analysis, we use earnings announcements issued over different quarters, and then estimate the model in a sub-group, for which we assume that the companies in the sub-groups have similar information structures.

Second, the reason the number of trading periods is limited to six is, principally because the optimal weighting matrix explained above must be estimated as well as the exogenous parameters. Since the size of the weighting matrix is $(3T \times 1)$, the more is added to the number of trading periods, the more elements in the variance-covariance matrix will have to be estimated. This, then, clearly would attenuate the power of the test.

Third, the exogenous parameters include the number of informed traders, M, the covariance between any two signals, Ω_0, the difference between the initial variance and the covariance of the signals, χ,[19] and the amount of liquidity trading, σ_u^2.

Finally, although the number of informed traders is an integer, in order to facilitate estimation, it must be interpreted as if it were a real number. Since the estimate in the empirical section can be interpreted as the cross-sectional average of informed traders at the firm level, this change does not affect economic intuition regarding the interpretation of the equilibrium provided in Proposition 1.[20]

In order to test the restricted models, HS and Kyle, against the general model, FV, we evaluate the Distance Metric statistics (*DM*), as suggested by Newey and McFadden (1994).[21] If $\widehat{Y}$ represents the GMM estimator for the general model, and $\overline{Y}$ is the GMM estimator for the restricted model, the *DM* statistic can be calculated as follows:

$$DM = 2n[Q(\overline{Y}) - Q(\widehat{Y})],$$

where n is the number of observations, and the $Q(\cdot)$'s are the minimized values of the criterion function, evaluated at the respective GMM estimates. As HS assumes a perfect correlation among private signals, HS can be tested with the null hypothesis, such that $\chi = 0$. Under the null hypothesis, the *DM* statistic converges to a chi-square with one degree of freedom. In addition, Kyle assumes that the number of informed traders is 1. In Kyle, therefore, the pertinent null hypothesis is $\chi = 0$ and $M = 1$, and under the null hypothesis, the *DM* statistic converges to a chi-square with two degrees of freedom.

5. Empirical results

There are three parts to this section. We first estimate the parameters of the FV model. In presenting the results, we start by briefly discussing the results from the daily series, and then those from the intra-day series. We focus on the results from the intra-day series, since this time frame is better suited for the market microstructure models. However, we present the results from the daily series first since the daily series starts earlier than the intra-day series in calendar time. This assists in explaining the intertemporal changes in the parameter estimates. Second part shows the results from testing the restrictions of HS and Kyle against the FV model. Last, we use parameter estimates of the FV model to measure the contributions of the informed traders, market maker and noise traders to trading volume and price volatility. The second and last parts are done using the intra-day series.

[19] The difference between the conditional variance and covariance of the private signals is a constant. Therefore, by estimating any two of Λ_0, Ω_0 and χ, it is possible to identify all three parameters. In the following section, we estimate Ω_0 and χ. As χ is positive, a constrained minimization is undertaken, in which χ is constrained to be positive.

[20] Foster and Viswanathan (1995) also estimate the model, assuming that the number of informed traders is a real number. In the estimation process, this parameter is constrained to be larger than 1. With the number of the informed traders being less than one, the equilibrium is undefined.

[21] This statistic is analogous to the Likelihood Ratio statistic in the Maximum Likelihood Estimation.

5.1. Parameter estimates

5.1.1. Daily series

Table 3 reports the parameter estimates of the FV model using the daily series. This table also reports the correlation coefficient of the private signals, ρ, which is evaluated by the division of Ω_0 by Λ_0. Panel A presents the estimates for the sub-samples, in which sampling is conducted according to firm size. Panel B, on the other hand, provides estimates for the sub-samples, in which the sampling is conducted according to absolute price changes prior to trading rounds.

First, the number of informed traders is, generally, estimated to be larger than one. In fact, for all the samples except 'Large Move' the number of informed traders is significantly greater than one. When the results from the sub-samples are compared, the 'Large Firms' and the 'Large Move' samples are estimated to have the fewest informed traders in their respective groups.

Second, the estimates for initial price informativeness decreases with the firm size, suggesting that the initial level of informational asymmetry is inversely related with the firm size. This is reasonable in that the number of analysts whose activities produce information about the company increases with firm size. As a result, entering into the trading rounds, the traders face a lesser degree of an adverse selection problem as the firm size becomes larger.

Table 3
Parameter estimates of the general model for daily series

	Number of informed traders, M	Initial price informativeness, $(\Sigma_0)^{1/2}$	$\Lambda_0 - \Omega_0$, χ	Correlations of private signals, ρ	Intensity of liquidity trading, σ_u
Whole Sample	5.35*	12.95***	73.51***	−0.1200	337.32***
[2518]	(0.056)	(0.000)	(0.000)		(0.000)
Panel A: Sub-samples by firm size					
Large Firms	4.58***	0.12***	101.89***	−0.2792	950.12***
[839]	(0.000)	(0.000)	(0.000)		(0.000)
Medium Firms	7.66***	11.94***	36.07***	−0.0675	87.74***
[840]	(0.009)	(0.000)	(0.000)		(0.000)
Small Firms	6.49*	14.20***	70.46***	−0.0943	18.58***
[839]	(0.098)	(0.000)	(0.001)		(0.000)
Panel B: Sub-samples by $\lvert p_0 - p_{-1}\rvert$					
Large Move	2.15	13.75***	82.13***	0.0319	359.39***
[839]	(0.168)	(0.000)	(0.000)		(0.000)
Medium Move	7.91*	12.06***	101.71***	−0.1156	249.11***
[840]	(0.067)	(0.007)	(0.003)		(0.000)
Small Move	8.92*	13.69	246.84**	−0.1143	145.87***
[839]	(0.050)	(0.119)	(0.038)		(0.000)

Notes. This table provides the GMM estimates of the general model based on the report of FV for the daily series. The model is estimated using the normalized price volatility and the normalized trading volume. For details on normalization, refer to Section 2. Four parameters are to be estimated: the number of informed traders, M, the initial informativeness of price, Σ_0, the difference between initial variance and covariance of the signals, χ, and the amount of liquidity trading, σ_u. This table also provides the estimates of the correlation coefficient between any two private signals, ρ. The numbers in [ ] present the number of observations in the corresponding sample. The GMM estimators are asymptotically normally distributed. The numbers in ( ) are the p-values of the estimates. However, for M, the null hypothesis is whether the number of informed traders is larger than 1.

* The null hypothesis is rejected at a significance level of 10%.

** Idem, 5%.

*** Idem, 1%.

Third, the parameter estimates of intensity of liquidity trading increase with the price volatility leading into the trading rounds. For example, the liquidity trading is largest for the 'Large Move' sample. This is consistent with Bloomfield et al. (2006), which show in experimental settings that the noise trading is largest in a high volatility environment. It appears that liquidity traders want to execute trades when the stock price is moves significantly.

Fourth, unlike the other samples, only the 'Large Move' sample reveals positive correlation among the signals. As the correlation becomes closer to one, the model converges to HS in which the informed traders submit orders simultaneously, increasing the price volatility at the beginning of the trading periods. Therefore, the reason that a positive correlation for this sample is achieved is due to the high price volatility in the first period. If we investigate the realized price volatility (ΔP^2) in Table 2, we observe that the 'Large Move' sample has relatively large first period price volatility compared to that of the terminal period. The average ratio of the first period price volatility to the terminal period volatility for all the other 8 sub-samples is approximately 15%. The same ratio for the 'Large Move' sample is about 90%. The high first period price volatility for this sample appears to originate from the persistence in volatility. In other words, the previous day's high volatility is probably being spilled over into the first day.

Last, recall that the more negative the correlation between signals becomes, the greater the stock of information. The parameter estimate is smallest for the 'Large Firms' sample at -0.2792. This suggests that even though these companies have the least initial adverse selection problem and the fewest number of informed traders, they attract informed traders who can learn more from one another than the informed traders would about smaller firms.

5.1.2. Intra-day series

Table 4 presents the parameter estimates of the FV model using the intra-day series. It is important to note that the initial observation in the daily series is made six days prior to the announcements, whereas that in the intra-day series is made three hours prior to the announcements. Therefore, within the context of the estimation strategy, the changes of estimates from daily to intra-day estimations reflect changes in the trading environments, as the time of the announcements approaches.

First, as the time of the announcements approaches, the number of informed traders increases dramatically. If the 'Whole' samples are compared, the number of informed traders increases, from approximately 5 to 300. This result provides evidence that the number of informed traders is not constant as the time of the announcements approaches, and that more traders are acquiring private signals. Therefore, a more realistic theoretical model would take into consideration the endogenizing of the number of informed traders, as a function of time to the revelation of information.

FV show that, even though the informed traders are endowed with positively correlated signals, as time approaches maturity, the conditional correlation among the signals decreases gradually. This decrease occurs because the market maker learns more about the 'average' of the signals than about individual signals. In fact, s/he will never learn about the individual trader's signal. As a result, the correlated component of the signal is revealed to the market maker quickly since it causes informed traders to trade similarly leading to large volumes, but that the uncorrelated component is more difficult for market makers to detect in volume. Eventually, the remaining component of the signal is negatively correlated.

Learning by the market maker, and its implications with regard to the correlation among private signals, does not appear to be confirmed by the empirical results. Except for the 'Large Move' sample, for which the positive correlation becomes negative, the estimated correlations

Table 4
Parameter estimates of the general model for intra-day series

	Number of informed traders, M	Initial price informative-ness, $(\Sigma_0)^{1/2}$	$\Lambda_0 - \Omega_0$, χ	Correlations of private signals, ρ	Intensity of liquidity trading, σ_u
Whole Sample	291.75**	7.40***	0.55***	−0.0023	6.29***
[2,491]	(0.020)	(0.000)	(0.000)		(0.000)
Panel A: Sub-samples by firm size					
Large Firms	332.56***	3.99***	0.18***	−0.0022	15.31**
[830]	(0.000)	(0.000)	(0.004)		(0.000)
Medium Firms	142.98***	5.59***	2.00*	−0.0063	2.43***
[831]	(0.001)	(0.000)	(0.089)		(0.000)
Small Firms	259.29***	13.99***	82.57	−0.0148	0.68***
[830]	(0.005)	(0.000)	(0.283)		(0.000)
Panel B: Sub-samples by $\lvert p_0 - p_{-2}\rvert$					
Large Move	212.53	6.04***	1.06***	−0.0040	8.84**
[830]	(0.201)	(0.000)	(0.000)		(0.011)
Medium Move	350.05***	5.31	0.64***	−0.0025	8.19***
[831]	(0.000)	(0.366)	(0.000)		(0.000)
Small Move	259.29***	8.29	3.79***	−0.0036	3.12***
[830]	(0.000)	(0.443)	(0.001)		(0.000)

Notes. This table provides the GMM estimates of the general model, based on the work of Foster and Viswanathan (FV) for the intra-day series. The model is estimated using the normalized price volatility and the normalized trading volume. For details on normalization, refer to Section 2. Four parameters are to be estimated: the number of informed traders, M, the initial informativeness of price, Σ_0, the difference between initial variance and covariance of the signals, χ, and the amount of liquidity trading, σ_u. This table also provides the estimates of the correlation coefficient between any two private signals, ρ. The numbers in [ ] present the number of observations in the corresponding sample. The GMM estimators are asymptotically normally distributed. The numbers in ( ) are the p-values of the estimates. However, for M, the null hypothesis is whether the number of informed traders is larger than 1.

* The null hypothesis is rejected at a significance level of 10%.
** Idem, 5%.
*** Idem, 1%.

for intra-day series are larger than those for daily series. In other words, it appears that the conditional correlation increases rather than decreases as time approaches the announcements. For example, for the 'Whole' sample, the correlation increased from −0.12 to −0.0023. One way of explaining this seemingly contradictory result is to explain the increase in correlation with the increase in number of informed traders. In the daily series a fewer number of informed traders endowed with more negatively correlated signals trade. The stock of information is larger in this case, and the informed traders have less incentive to trade aggressively at earlier trading rounds. However, as time passes, more traders receive signals, and accordingly in the intra-day series, a larger number of informed traders have less negatively correlated signals, implying that the stock of information is less than in the daily series. Again, the parameters of the models change over time, and this is another area that the theory needs to address.

Next, the level of liquidity trading decreases as the time of the announcements approaches. Recall that σ_u for the daily series estimates the level of liquidity trading over a trading day, while for intra-day series, it estimates the level of liquidity trading over a thirty minute interval. As there are 13 30-minute intervals in a trading day, in order to heuristically compare the two sets of estimates,[22] the estimates from the intra-day series should be multiplied by 3.6, which is the square root of 13. Then, for all samples the daily amount of intra-day liquidity trading is much smaller than the liquidity trading from the daily series. In fact, the ratios of the former to the latter range from 6 to 13%. In the case of the 'Whole' sample, the daily amount of intra-day liquidity trading is approximately 23, while the liquidity trading from the daily series is 337, yielding 7% for the ratio. It is evident that, near the time of the announcements, the liquidity traders reduce their orders, thereby rendering the market less liquid. This confirms the findings reported by Krinsky and Lee (1996), in which the adverse selection component of the bid-ask spread increases near the time of the announcements. In addition, as in the case of the numbers of informed traders, the intensity of liquidity trading is also altered as time progresses. This may indicate a market in which the likes of Admati and Pfleiderer's (1988) discretionary liquidity traders time their trades in order to minimize their expected losses. These timing activities conducted by the liquidity traders may also contribute to the cross-sectional variations of the estimates over the sub-samples. For both daily and intra-day series, the intensity of liquidity trading is positively related to firm size. As was the case with the daily series, when the sub-samples sorted on preceding absolute price change are analyzed, the intensity of liquidity trading is positively related to the preceding volatility for both daily and intra-day series.

Now, the issue of initial informativeness is discussed. In five out of seven cases, this parameter is significantly estimated. This value roughly measures the level of uncertainty heading into trading rounds. A larger estimate implies a less informative initial price. First, if the sub-samples are investigated, the informativeness increases with the firm size. This is reasonable in that a larger number of analysts produce more information for larger firms. Second, if we compare the results between the daily and intra-day analyses, we find that the adverse selection sharply increases near the announcements. Note that like the case of liquidity trading, the estimates for the intra-day estimation measures volatility over 30 minute intervals. Therefore, in order to compare the values fairly, the intra-day estimates must be multiplied by the square root of 13. Then, for all samples, the daily amounts of intra-day estimates are greater than that of daily estimates, implying an increase in the adverse selection problem. More specifically, for all samples except, the 'Large Firms' sample, the daily amounts of intra-day estimates are approximately 1.5 times to 3.5 times larger than the daily estimates. For the 'Large Firms' sample, due to the low estimate for the daily series, the intra-day estimate is literally more than 100 times greater than the daily estimate.

In summation, the results from Table 3 and 4 characterize the market before the announcements as the one in which, with the time of announcements drawing near, a multiple and increasing number of informed traders trade on their private signals, and learn about the other traders' private signals, while the adverse selection intensifies, and the liquidity trading decreases.

5.2. Test of models: heterogeneous vs. homogeneous information

In this section, we test the validity of the assumptions imposed in the two restricted models: HS and Kyle. The HS model is derived from the FV model, via the assumption that the informed traders receive homogeneous information, whereas Kyle is derived by further assuming that the informed trader is monopolistic. The results of the tests using intra-day series are reported in

[22] This is exact if the liquidity trading over each of the thirty-minute intervals is independent, which is precisely the assumption of the theoretical model.

Table 5.[23] The panels A and B present the *DM* statistics for the sub-samples, in which sampling is conducted according to firm size and absolute price changes, respectively.

For all of the samples tested, both HS and Kyle are soundly rejected against FV. When the Kyle model is tested against the HS model, Kyle is not rejected. Moreover, in this case, the DM statistics are uniformly small, ranging from 0.002 to 0.066, implying that merely changing the number of traders from one to many, coupled with the retention of information homogeneity, improves performance of the model only slightly. In other words, immediately prior to the earnings announcements, relaxing the assumption of the monopolistic informed trader does not improve the effectiveness of explaining the data, unless it is accompanied by the relaxation of the assumption of homogeneous information. Therefore, it is the structure of the information, i.e. homogeneous vs. heterogeneous, that adds empirical value to market microstructure modeling, rather than the number of informed traders present at least three hours before the announcements. The primary reason HS does not outperform Kyle appears to rely on the fact that a rat race does not exist among informed traders at the beginning of trading periods. In fact, Table 1 indicates that, for all the samples, Δp^2 increases as the time of the announcements approaches.

Table 5
Test of models for intra-day series

	Kyle vs. HS $\chi^2(1)$ under Kyle	HS vs. FV $\chi^2(1)$ under HS	Kyle vs. FV $\chi^2(2)$ under Kyle
Whole Sample	0.005	223.714***	223.720***
[2491]	(0.943)	(0.000)	(0.000)
Panel A: Sub-samples by firm size			
Large Firms	0.002	148.515***	148.516***
[830]	(0.968)	(0.000)	(0.000)
Medium Firms	0.003	89.812***	89.815***
[831]	(0.954)	(0.000)	(0.000)
Small Firms	0.066	72.698***	72.765***
[830]	(0.797)	(0.000)	(0.000)
Panel B: Sub-samples by $\lvert p_0 - p_{-2}\rvert$			
Large Move	0.005	42.620***	42.625***
[830]	(0.943)	(0.000)	(0.000)
Medium Move	0.007	147.730***	147.737***
[831]	(0.934)	(0.000)	(0.000)
Small Move	0.034	31.753***	31.787***
[830]	(0.854)	(0.000)	(0.000)

Notes. This table provides the *DM* (Distance Metric) statistics for the testing of the models of homogeneous information against the general model of heterogeneous information. There are two models of homogeneous information: HS and Kyle. In HS, the informed traders receive an identical signal. In Kyle, there is only one informed trader. We first estimate FV, HS, and Kyle, and obtain $DM = 2n[Q(\overline{Y}) - Q(\widehat{Y})]$, where n is the number of observations, $Q(\overline{Y})$, the minimized value of the criterion function evaluated at the GMM estimates for the restricted model, and $Q(\widehat{Y})$ is the minimized value of the criterion function evaluated at the GMM estimates for the unrestricted model. In Kyle vs. HS, HS serves as the unrestricted model. In the other two cases, FV is the unrestricted model. The numbers in [ ] are the number of observations in the sample, while those in ( ) are the p-values of the estimates.

*** The null hypothesis is rejected at a significance level of 1%.

23 The tests are conducted using the daily series. The results are similar to those using the intra-day series, and will not be reported for the sake of parsimony. The results can be provided upon request.

In summary, the assumption of homogeneous information is strongly rejected. The HS model possesses little empirical value. We should not, however, jump to the conclusion that the HS model is insufficient to explain the information dissemination process at times proximal to the earnings announcements. Unlike the testing of Kyle or FV, knowledge of the exact timing of the informed traders' acquisition of information is critical in testing HS. This is because, in HS, the majority of the trading activities occur at the beginning of the trading periods. If the timing of the information acquisition is miss-specified, and set either too early or too late, the HS model is very likely to be rejected. If the time is set too early, the rat race would appear in the middle of trading rounds, or if the time is set too late, it would never shows up at all. For the Kyle or FV models, however, as long as heavy trading activities occur within the last trading round, these models are unlikely to be rejected, even if the timing of information acquisition is miss-specified.

5.3. Decomposition trading volume and price volatility

One of the benefits of structural model estimation is the possibility not only of the identification of the underlying sources which bring about changes in observable variables, but also the quantification of the source's relative contributions to changes in the observable variables. In this section, the relative contributions of each of the market participants to the expected trading volume and price volatility are documented.

First, we analyze trading volume. As an illustration, from the point estimates of FV for the 'Whole' sample using intra-day series, the expected trading volume for each period is evaluated using Eq. (2). The first term in Eq. (2) is the portion contributed by the informed traders, the second term, by the liquidity traders, and the final term is the portion contributed by the market maker. We then quantify, as a percentage, the relative contributions of the informed traders, liquidity traders, and the market maker to the expected trading volume. The results of these evaluations are reported in Table 6.

As demonstrated, the relative contribution of the informed traders increases as the time approaches maturity. This is due to the fact that the informed traders become more aggressive as the announcement approaches. However, contributions by both the liquidity traders and the market maker decrease as the time approaches maturity. The decrease in the contribution of the liquidity

Table 6
Trading volume decomposition

Trading periods	Sample moments	Estimated moments	Informed traders	Market maker	Liquidity traders
1	8.36	8.49	0.756	0.123	0.121
2	8.50	9.14	0.773	0.115	0.112
3	9.76	10.05	0.793	0.105	0.102
4	11.57	11.43	0.818	0.093	0.090
5	18.38	13.86	0.848	0.078	0.074
6	25.16	19.65	0.890	0.058	0.052

Notes. This table shows the intertemporal changes in the proportion of expected trading volumes contributed to by the informed traders, the market maker, and the liquidity traders. From the point estimates of FV for the 'Whole' sample using intra-day series, the expected trading volume for each of the periods is evaluated using Eq. (2). The first term in Eq. (2) is the portion contributed to by the informed traders, the second term, by the liquidity traders, and the final term, the market maker. We then quantify, as a percentage, the relative contributions of the informed traders, liquidity traders, and the market maker to the expected trading volumes. The second column of the table provides the sample moments of trading volume. They are identical to those provided in Table 1. The third column provides estimated theoretical moments, which serve as the basis of decomposition.

Table 7
Price volatility (Δp^2) decomposition

Trading periods	Sample moments	Estimated moments	Informationally motivated trades	Liquidity trades
1	0.68	2.02	0.044	0.956
2	0.64	2.23	0.047	0.953
3	0.82	2.56	0.056	0.944
4	0.78	3.08	0.072	0.928
5	4.48	4.06	0.102	0.898
6	9.11	6.59	0.184	0.816

Notes. This table provides the intertemporal changes in the proportion of price volatility contributed to by the informationally motivated trades and the liquidity trades. From the point estimates of FV for the 'Whole' sample using intra-day series, the price volatility for each of the periods is evaluated using Eq. (1). The first two terms in Eq. (1) are the portion contributed by informationally motivated trades, and the third term, by the liquidity trades. We then quantify, as a percentage, the relative contributions of the informationally motivated trades and the liquidity trades to the price volatility. The second column of the table provides the sample moments of Δp^2. They are identical to those provided in Tables 1. The third column shows the estimated theoretical moments, which serve as the basis of decomposition.

traders is rather unsurprising, as the absolute amount of liquidity trading is assumed to remain constant, whereas the overall expected trading volume increases. What is interesting, however, is the fact that the contribution by the market maker also decreases. Note that the market maker fills the net orders. Therefore, the fact that the proportion of the net orders decreases with liquidity trading held constant indicates that the proportion of orders crossed by the informed traders must be increasing. This stems from the fact that, as time progresses, informed traders learn from one another, reducing the correlation among the signals. In this process, they eventually reach opposite positions.

Second, with regard to price volatility, we conduct a similar exercise. According to the point estimates of FV for the Whole sample using intra-day series, the price volatility, Δp^2, for each period is evaluated using Eq. (1). The first two terms in Eq. (1) represent the portion contributed by the informationally motivated trades, and the second term, by the liquidity trades. We then quantify, as a percentage, the relative contributions of the informationally motivated trades and the liquidity trades to the price volatility. The results of these evaluations are presented in Table 7.

Similar to the case of trading volume, the relative contribution by the informationally motivated trades increases as time approaches maturity. This, again, is due to the fact that the informed traders become more aggressive, whereas the liquidity traders submit constant orders.

6. Conclusion

In this paper, we estimate a theoretical model of strategic informed trading developed by Foster and Viswanathan (FV), in which multiple informed traders receive heterogeneous private signals regarding the terminal value of the asset. This model nests two simpler models: that of (i) Holden and Subrahmanyam (HS) in which multiple informed traders receive homogeneous private signals, and that of (ii) Kyle (1985), in which the informed trader is assumed to be monopolistic. From the equilibria of these models, we derive the moment restrictions on price volatility, trading volume, and inverse market depth parameter. Using the price and volume data for the six trading days—for the daily series—, and three hours—for the intra-day series—leading up to the earnings announcements, we estimate and test these models via the Generalized Method of Moments (GMM).

The principal findings of this study are (i) multiple informed traders with heterogeneous signals appear to trade close to the earnings announcements, (ii) as the time of the announcements approaches, the number of informed traders increases, the adverse selection problem intensifies, and the liquidity trading decreases, (iii) the structure of the information, i.e. whether it is homogeneous or heterogeneous, is of greater importance than simply the number of informed traders, in explaining the trading activities close to the time of the announcements. Finally, (iv) the contribution of informed trading to trading volume, as well as price volatility, increases as the time of the announcements approaches.

As a concluding remark, it is important to emphasize that the process of structural model estimation is a limited one, in that theoretical models are likely to be miss-specified as a result of heavy structure imposed by the simplifying assumptions. For example, the Kyle-variant models estimated in this paper do not directly take into consideration the inventories maintained by the specialists, nor do they consider the role of public information regarding new events or up-coming announcements. In addition, the Kyle-variant models assume that exogenous parameters, including the number of informed traders, liquidity trading, and other information-related parameters are constant during the trading rounds. These issues are extremely difficult to handle, but, in this paper, the constant parameter assumption is addressed, at least in part, via the estimation of the models over two different time horizons. Evidence is found to suggest that these parameters do, indeed, change as the time of the announcements approaches. In fact, the number of informed traders appears to increase with time, and we believe that the provision of empirical evidence regarding this effect constitutes an important contribution to the relevant literature. The job of theoretical modeling of endogenizing the number of informed traders, as well as determining its impact on equilibrium order-submitting and price-setting strategies, is left as future research.

Acknowledgments

I am grateful to Kerry Back, Bruno Biais, Sugato Bhattacharyya, Eric Chang, Richard Green, Craig Holden, Murugappa Krishnan, Ananth Madhavan, Edward Nelling, Duane Seppi, Rajdeep Singh, Fallaw Sowell, Chester Spatt, Avnidhar Subrahmanyam, S. Viswanathan, and the seminar participants in Atlanta Finance Workshop, Carnegie Mellon University, Georgia Institute of Technology, University of Waterloo and Yale School of Management for the many helpful comments and suggestions. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-003-B00109). I would especially like to thank the anonymous referee for the insightful comments and suggestions. The usual disclaimer applies.

Appendix A. Derivation of Foster and Viswanathan (FV)

In this appendix, we discuss more in detail the structure of the general model, FV, and the equilibrium thereof. For the formal proofs, please refer to Foster and Viswanathan (1996).

A.1. Structure

As was mentioned in this paper text, in FV, three types of risk-neutral traders, a market maker, M informed traders and a number of liquidity traders trade in the market over T trading periods.

Prior to trading, i.e. at time 0, each informed trader receives a private signal which is correlated to, but may well be different from the others'.

Let the liquidating value be $\tilde{S}$. The ith informed trader receives a signal s_{i0} at the start of trading. Further, it is assumed that $\tilde{S} = p_0 + \sum_{i=1}^{M} s_{i0}$, where p_0 is the common prior about the mean of the terminal value of the asset. The signal vector for all informed traders, $(s_{10}, s_{20}, \ldots, s_{M0})$, is drawn from a multivariate normal distribution. The mean of each signal is zero. Both the variance of each signal and the covariance between any two signals are identical across the informed traders. The variance and covariance of signals are denoted by Λ_0, and Ω_0, respectively. In addition, there are a number of liquidity traders who submit a random liquidity demand, u_t, at period t. We assume that u_t is normally distributed with mean 0 and variance σ_u^2. We further assume that the liquidity trading is independent of all the other random variables.

At the beginning of period t, the ith informed trader updates valuation of the asset's liquidating value, and submits orders representing request for changes in holdings, $\Delta x_{it} = x_{it} - x_{it-1}$. The market maker then observes the net order flow, $\Delta y_t = \sum_{i=1}^{M} \Delta x_{it} + u_t$, and sets the period-$t$ price so that

$$p_t = E\big(\tilde{S} \bigm| \Delta y_1, \Delta y_2, \ldots, \Delta y_1\big). \tag{A.1}$$

In choosing Δx_{it} the ith informed trader uses his initial private signal and the history of net order flows so that $E(\pi_{it} \mid s_{i0}, \Delta y_1, \Delta y_2, \ldots, \Delta y_1)$ is maximized, where

$$\pi_{it} = \sum_{\tau=t}^{T} \big(\tilde{S} - p_\tau\big)\Delta x_{i\tau} \quad \left(= \tilde{S}(x_{iT} - x_{it}) - \sum_{\tau=t}^{T} p_\tau \Delta x_{i\tau}\right). \tag{A.2}$$

It is necessary to provide definitions to several important variables and parameters before proceeding with the equilibrium. FV is unique due to the heterogeneity of private signals observed by each informed trader, both the market maker and the informed traders create inferences about the signals. In particular, at the beginning of period $t+1$, after having observed the order flows $(\Delta y_1, \Delta y_2, \ldots, \Delta y_t)$, the market maker updates his estimate of the signal of the ith trader, s_{i0}, to

$$t_{it} = E(s_{i0} \mid \Delta y_1, \Delta y_2, \ldots, \Delta y_t). \tag{A.3}$$

Let the estimation error be $s_{it} = s_{i0} - t_{it}$. Finally, we define the variables that measure the remaining information at the end of period-t:

$$\Sigma_t = \mathrm{Var}\big(\tilde{S} \bigm| \Delta y_1, \Delta y_2, \ldots, \Delta y_1\big) = \mathrm{Var}\big(\tilde{S} - p_t\big), \tag{A.4}$$

$$\Lambda_t = \mathrm{Var}(s_{i0} \mid \Delta y_1, \Delta y_2, \ldots, \Delta y_1) = \mathrm{Var}(s_{i0} - t_{it}) = \mathrm{Var}(s_{it}), \tag{A.5}$$

$$\Omega_t = \mathrm{Cov}(s_{i0}, s_{j0} \mid \Delta y_1, \Delta y_2, \ldots, \Delta y_1) = \mathrm{Cov}(s_{i0} - t_{it}, s_{j0} - t_{jt}) = \mathrm{Cov}(s_{it}, s_{jt}). \tag{A.6}$$

Here, Σ_t is the conditional variance of the liquidation value; Λ_t, the conditional variance of trader i's signal, s_{i0}; and Ω_t, the conditional covariance between any two signals, s_{i0} and s_{j0}. It can be demonstrated that these measures are inter-related, and in particular,

$$\Lambda_t - \Omega_t = \chi \quad \text{for all } t. \tag{A.7}$$

A.2. Equilibrium

The equilibrium notion originates from Kyle (1985) and is provided below. First, the ith informed trader's period-t order submitting strategy is defined by

$$x_{it} = X_{it}(s_{i0}, \Delta y_1, \Delta y_2, \dots, \Delta y_{t-1}), \tag{A.8}$$

and the market maker's period-t pricing rule is defined by

$$p_t = P_t(\Delta y_1, \Delta y_2, \dots, \Delta y_t). \tag{A.9}$$

Definition. A Bayesian Nash equilibrium of a trading game is an $M+1$ vector of strategies, $(X_1, \dots, X_T, P)$, where $X_i = \langle X_{i1}, \dots, X_{iT}\rangle$ with $X_{i0} = 0$, and $P = \langle P_1, \dots, P_T\rangle$, such that the following *Profit Maximization* and *Market Efficiency* conditions are satisfied; i.e.,

(1) for any $I = 1, \dots, M$ and all $t = 1, \dots, T$ we have for all $X_i' = \langle X_{i1}' \cdots X_{iT}'\rangle$,

$$\begin{aligned} &E\{\pi_t(X_1, \dots, X_i, \dots, X_M, P) \mid s_{i0}, \Delta y_1, \dots, \Delta y_{t-1}\} \\ &\quad \geqslant E\{\pi_t(X_1, \dots, X_i', \dots, X_M, P) \mid s_{i0}, \Delta y_1, \dots, \Delta y_{t-1}\}; \end{aligned}$$

(2) for all $t = 1, \dots, T$,

$$p_t = E(\tilde{S} \mid \Delta y_1, \Delta y_2, \dots, \Delta y_1).$$

Given the structure of the model and the definition of the equilibrium, we solve for a linear Markov equilibrium. In other words, we search for equilibrium where an informed trader's order submitting strategy, market maker's learning about the signal vector, and market maker's learning about the true value of the asset take the form:

$$x_{it} = \beta_t s_{it-1}, \tag{A.10}$$

$$t_{it} = t_{it-1} + \zeta_t \Delta y_t, \tag{A.11}$$

$$p_t = p_{t-1} + \lambda_t \Delta y_t. \tag{A.12}$$

Now, we are in the position to give the necessary and sufficient condition for a recursive linear Markov equilibrium described above. They are stated in the following proposition. For the sake of parsimony, the proof and detailed discussion on the equilibrium is not provided.

Proposition A.1. *The necessary and sufficient conditions for a recursive linear Markov equilibrium are*:

$$\beta_t = \frac{\eta_t - \lambda_t \psi_t}{\lambda_t\{1 + (1 + \phi_t(M-1))(1 - (\lambda_t \psi_t/\theta))\}}, \tag{A.13}$$

$$\gamma_t = \frac{(1 - 2\mu_t\lambda_t)(1 - (\lambda_t \beta_t (M-1)/\theta))}{2\lambda_t(1 - \mu_t\lambda_t)}, \tag{A.14}$$

$$\lambda_t = \frac{\beta_t M \Sigma_t}{\theta \sigma_u^2}, \tag{A.15}$$

plus the value function variables

$$\alpha_{t-1} = [\eta_t - \lambda_t\beta_t\{1 + (M-1)\phi_t\}]\beta_t + \alpha_t\left[1 - \frac{\lambda_t\beta_t}{\theta}\{1 + (M-1)\phi_t\}\right]^2, \tag{A.16}$$

$$\psi_{t-1} = \big[\eta_t - \lambda_t\beta_t\{1+(M-1)\phi_t\}\big]\gamma_t - \lambda_t\gamma_t\beta_t + \beta_t\left\{1 - \frac{\lambda_t\beta_t(M-1)}{\theta}\right\}$$
$$+ \psi_t\left[1 - \frac{\lambda_t\beta_t}{\theta}\{1+(M-1)\phi_t\}\right]\left[1 - \frac{\lambda_t\beta_t(M-1)}{\theta} - \lambda_t\gamma_t\right], \tag{A.17}$$

$$\mu_{t-1} = -\lambda_t\gamma_t^2 + \gamma_t\left\{1 - \frac{\lambda_t\beta_t(M-1)}{\theta}\right\} + \mu_t\left\{1 - \frac{\lambda_t\beta_t(M-1)}{\theta} - \lambda_t\gamma_t\right\}^2, \tag{A.18}$$

$$\delta_{t-1} = \delta_t + \frac{\alpha_t\lambda_t^2}{\theta^2}\sigma_u^2 + \alpha_t\frac{\lambda_t^2\beta_t^2}{\theta^2}\big\{(M-1)\operatorname{Var}(s_{jt-1} \mid s_{i0}, y_1, y_2, \ldots, y_1)\big\}$$
$$+ \alpha_t\frac{\lambda_t^2\beta_t^2}{\theta^2}\big\{(M-2)(M-1)\operatorname{Cov}(s_{jt-1}, s_{kt-1} \mid s_{i0}, y_1, y_2\cdots y_1)\big\}, \tag{A.19}$$

where $\alpha_T = \psi_T = \mu_T = \delta_T = 0$ and

$$\phi_t = \frac{\Omega_{t-1}}{\Lambda_{t-1}}, \tag{A.20}$$

$$\eta_t = \frac{\theta}{M}\big\{1+(M-1)\phi_{t-1}\big\}, \tag{A.21}$$

$$\operatorname{Var}(s_{jt-1} \mid s_{i0}, y_1, y_2, \ldots, y_1) = \frac{\Lambda_{t-1}^2 - \Omega_{t-1}^2}{\Lambda_{t-1}}, \tag{A.22}$$

$$\operatorname{Cov}(s_{jt-1}, s_{kt-1} \mid s_{i0}, y_1, y_2, \ldots, y_1) = \frac{\Omega_{t-1}(\Lambda_{t-1} - \Omega_{t-1})}{\Lambda_{t-1}}, \tag{A.23}$$

with

$$\theta = \kappa M, \quad \textit{and} \tag{A.24}$$

$$\kappa = \frac{c_0\chi}{(\Lambda_0[\Lambda_0 + (M-2)\Omega_0] - (M-1)\Omega_0^2)}, \tag{A.25}$$

where c_0 is the covariance between the true value of the asset and each signal.

Furthermore, the second order condition,

$$\lambda_t(1 - \mu_t\lambda_t) > 0 \quad \textit{and} \tag{A.26}$$

the following recursive equations on the information variances must hold:

$$\Sigma_t = \left(1 - \frac{M\lambda_t\beta_t}{\theta}\right)\Sigma_{t-1}, \tag{A.27}$$

$$\Lambda_t = \Lambda_{t-1} - \frac{M}{\theta^2}\frac{\lambda_t\beta_t}{\theta}\Sigma_{t-1}, \tag{A.28}$$

$$\Omega_t = \Omega_{t-1} - \frac{M}{\theta^2}\frac{\lambda_t\beta_t}{\theta}\Sigma_{t-1}. \tag{A.29}$$

In this equilibrium, there are five exogenous parameters: the number of the informed traders, M, the variance of private signal, Λ_0, the covariance between two signals, Ω_0, the unconditional mean of liquidating value, p_0, and the amount of liquidity trading, σ_u^2. In the estimation part, however, p_0 will be replaced by the previous period's price. This means the remaining four parameters are left to estimate. Given the values of these exogenous parameters and the number of trading period, T, all the other endogenous parameters are specified completely.

Appendix B. Proof of Proposition 1

First, Eq. (1) is proven. From Eq. (18) in FV, $\Delta p_t^2 = (p_t - p_{t-1})^2 = \lambda_t^2 \Delta y_t^2$, where y_t is the net order. In addition, as price changes and net order flows are related in a linear fashion, a price change fully reveals the net order flow. Therefore, since the net order is the sum of the orders from the informed traders and liquidity traders,

$$E\left(\Delta p_t^2 \mid p_1, \dots, p_{t-1}\right) = E\left[\lambda_t^2 \left(\sum_{i=1}^{M} x_{it}^2 + u_t^2 + 2\sum\sum_{i \neq j} x_{it} x_{jt}\right) \Bigm| p_1, \dots, p_{t-1}\right]. \tag{A.30}$$

From Eq. (24) in FV

$$\begin{aligned} E(x_{it}x_{jt} \mid p_1, \dots, p_{t-1}) &= E\left(\beta_t^2 s_{it-1} s_{jt-1} \mid p_1, \dots, p_{t-1}\right) \\ &= E\left[\beta_t^2 s_{it-1} E(s_{jt-1} \mid s_{it-1}) \mid p_1, \dots, p_{t-1}\right], \end{aligned} \tag{A.31}$$

where s is the private signal, and the last equation is derived via the law of iterative expectations. Then, using the definition of ϕ_t given in Proposition 1 in FV, the following is obtained:

$$\begin{aligned} E\left[\beta_t^2 s_{it-1} E(s_{jt-1} \mid s_{it-1}) \mid p_1, \dots, p_{t-1}\right] &= \beta_t^2 E\left(\phi_t s_{it-1}^2 \mid p_1, \dots, p_{t-1}\right) \\ &= \beta_t^2 \phi_t \Lambda_{t-1} = \beta_t^2 \Omega_{t-1}. \end{aligned} \tag{A.32}$$

Second,

$$E\left(x_{it}^2 \mid p_1, \dots, p_{t-1}\right) = E\left(\beta_t^2 s_{it-1}^2 \mid p_1, \dots, p_{t-1}\right) = \beta_t^2 \Lambda_{t-1}. \tag{A.33}$$

Substituting (A.31) and (A.33) into (A.30) yields (1).

As to the expected trading volume in (2), the trading volume is first defined as

$$\begin{aligned} &E(TV_t | p_1, \dots, p_{t-1}) \\ &\quad \equiv \frac{1}{2} E\left[\left(\sum_{i=1}^{M} |\Delta x_{it}| + |u_t| + \left|\sum_{i=1}^{M} \Delta x_{it} + u_t\right|\right) \Bigm| p_1, \dots, p_{t-1}\right] \\ &\quad = \frac{1}{2} E\left[\left(\sum_{i=1}^{M} |\beta_t s_{it-1}| + |u_t| + \left|\sum_{i=1}^{M} \beta_t s_{it-1} + u_t\right|\right) \Bigm| p_1, \dots, p_{t-1}\right], \end{aligned} \tag{A.34}$$

in which the first term measures the trading volume contributed by the informed traders, the second term, by the liquidity traders, and the final term, by the market maker. Note that, conditionally upon past prices, all of the random variables inside the absolute value operator are distributed normally, with a mean of zero. Therefore, using the property of normal random variables,

$$\begin{aligned} &E(TV_t | p_1, \dots, p_{t-1}) \\ &\quad = \sqrt{\frac{1}{2\pi}} \left\{ \sqrt{\operatorname{Var}\left(\sum_{i=1}^{M} \beta_t s_{it-1} \mid p_1, \dots, p_{t-1}\right)} + \sqrt{\operatorname{Var}(u_t \mid p_1, \dots, p_{t-1})} \right. \end{aligned}$$

$$+\sqrt{\operatorname{Var}\left(\sum_{i=1}^{M}\beta_t s_{it-1}+u_t \mid p_1,\ldots,p_{t-1}\right)}\Bigg\}. \tag{A.35}$$

Note that

$$\begin{aligned}
&\operatorname{Var}\left(\sum_{i=1}^{M}\beta_t s_{it-1}+u_t \mid p_1,\ldots,p_{t-1}\right)\\
&\quad=\operatorname{Var}\left(\sum_{i=1}^{M}\beta_t s_{it-1} \mid p_1,\ldots,p_{t-1}\right)+\sigma_u^2\\
&\quad=M\beta_t^2\operatorname{Var}(s_{it-1}\mid p_1,\ldots,p_{t-1})+M(M-1)\beta_t^2\operatorname{Cov}(s_{it-1},s_{jt-1}\mid p_1,\ldots,p_{t-1})+\sigma_u^2\\
&\quad=M\beta_t^2\Lambda_{t-1}+M(M-1)\beta_t^2\Omega_{t-1}+\sigma_u^2.
\end{aligned} \tag{A.36}$$

Then, a substitution of (A.33) and (A.36) into (A.35) proves (2).

References

Admati, A.-R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Rev. Finan. Stud. 1, 3–40.

Back, K., Cao, H., Willard, G., 2000. Imperfect competition among informed traders. J. Finance 55, 2117–2155.

Bloomfield, R., Saar, G., O'Hara, M., 2006. The limits of noise trading: An experimental analysis. Unpublished manuscript. The Johnson Graduate School of Management, Cornell University.

Caballé, J., Krishnan, M., 1995. The sources of volatility in a dynamic financial market with insider trading. Unpublished manuscript. University of Minnesota.

Chae, J., 2005. Trading volume, information asymmetry, and timing information. J. Finance 60, 413–442.

Cho, J.-W., 2004. State-space representation and estimation of market microstructure models. Asia-Pacific J. Finan. Stud. 33, 249–305.

Cho, J.-W., Krishnan, M., 2000. Prices as aggregators of private information: Evidence from S&P 500 Futures data. J. Finan. Quant. Anal. 35, 111–126.

Easley, D., Kiefer, N., O'Hara, M., Paperman, J., 1996. Liquidity, information, and infrequently traded stocks. J. Finance 51, 1405–1436.

Easley, D., Kiefer, N., O'Hara, M., 1997. One day in the life of a very common stock. Rev. Finan. Stud. 10, 805–835.

Foster, F.-D., Viswanathan, S., 1993a. The effect of public information and competition on trading volume and price volatility. Rev. Finan. Stud. 6, 23–56.

Foster, F.-D., Viswanathan, S., 1993b. Variation in trading volume, return volatility, and trading costs: Evidence from recent price formation models. J. Finance 48, 187–211.

Foster, F.-D., Viswanathan, S., 1995. Can speculative trading explain the volume-volatility relation? J. Bus. Econ. Statist. 13, 379–396.

Foster, F.-D., Viswanathan, S., 1996. Strategic trading when agents forecast the forecasts of others. J. Finance 51, 1437–1478.

Holden, C.-W., Subrahmanyam, A., 1992. Long-lived private information and imperfect competition. J. Finance 47, 247–270.

Hughson, E., 1990. Intraday trade in dealership markets. Unpublished manuscript. Carnegie Mellon University.

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

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

Lee, C.-M.-C., Mucklow, B., Ready, M., 1993. Spreads, depths, and the impact of earnings information: An intraday analysis. Rev. Finan. Stud. 6, 345–374.

Morse, D., Ushman, N., 1983. The effect of information announcements on the market microstructure. The Accounting Review 58, 247–258.

Newey, W.-K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics, vol. 4. Elsevier Science, pp. 2111–2245.

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

Seppi, D.-J., 1992. Block trading and information revelation around quarterly earnings announcements. Rev. Finan. Stud. 5, 281–305.

Venkatech, P.-C., Chiang, R., 1986. Information asymmetry and the dealer's bid–ask spread: A case study of earnings and dividend announcements. J. Finance 41, 1089–1102.