WILEY

Measuring the Information Content of Stock Trades

Author(s): Joel Hasbrouck

Source: *The Journal of Finance*, Mar., 1991, Vol. 46, No. 1 (Mar., 1991), pp. 179-207

Published by: Wiley for the American Finance Association

Stable URL: https://www.jstor.org/stable/2328693

Measuring the Information Content of Stock Trades

JOEL HASBROUCK*

ABSTRACT

This paper suggests that the interactions of security trades and quote revisions be modeled as a vector autoregressive system. Within this framework, a trade's information effect may be meaningfully measured as the ultimate price impact of the trade innovation. Estimates for a sample of NYSE issues suggest: a trade's full price impact arrives only with a protracted lag; the impact is a positive and concave function of the trade size; large trades cause the spread to widen; trades occurring in the face of wide spreads have larger price impacts; and, information asymmetries are more significant for smaller firms.

CENTRAL TO THE ANALYSIS of market microstructure is the notion that in a market with asymmetrically informed agents, trades convey information and therefore cause a persistent impact on the security price. The magnitude of the price effect for a given trade size is generally held to be a positive function of the proportion of potentially informed traders in the population, the probability that such a trader is in fact informed (i.e., the probability that a private information signal has in fact been observed), and the precision of the private information. The close dependence of the price impact on these factors, which may be referred to as the extent of the information asymmetry, provides a strong motivation for the empirical determination of this impact. This paper strives to achieve such a determination in a framework that is robust to deviations from the assumptions of the formal models. In the process, the framework establishes a rich characterization of the dynamics by which trades and quotes interact.

The market considered here is a specialist market in which a designated market-maker exposes bid and ask quotes to the trading public. An extensive theory has evolved that analyzes the market-maker's exposure to traders with superior information.[1] Concerning the extent of the information asymmetry, this body of theory yields two important empirical predictions: first,

*Department of Finance, Leonard N. Stern School of Business, New York University. For comments on an earlier draft I am indebted to Larry Harris, Robert Wood, and seminar participants at Columbia University, Duke University, Pennsylvania State University, Southern Methodist University, and the Securities and Exchange Commission. I am especially indebted to the referee Larry Glosten for the illustrative microstructure model used in Section II and for his help in framing the argument of Section III. All errors are my own responsibility.

[1] See Bagehot (1971), Copeland and Galai (1983), Glosten and Milgrom (1985), Kyle (1985), Easley and O'Hara (1987), Glosten (1987, 1989), Admati and Pfleiderer (1988), and Foster and Viswanathan (1987).

that the asymmetry is positively related to the spread (between the bid and ask quotes), and second, that the asymmetry is positively related to the price impact of a trade. The first effect has led empirical researchers seeking measurable proxies for information asymmetry to examine posted bid-ask spreads.[2] The principal advantage of this procedure is that posted bid-ask spreads are relatively easy to observe. Nevertheless, there are certain difficulties. First, price discreteness sharply restricts the number of possible values that the spread can take. (In practice on the NYSE, the posted spread is often 1/8 or 1/4 point, and somewhat less frequently takes values 3/8 or 1/2.) This rounding may effectively obscure many cross-security differences. Secondly, spreads reflect a component related to transaction costs, such as clearing fees. Finally, the posted spread represents the prices available at a given time for transactions only up to some fixed (usually relatively small) amount. Glosten (1989) shows that the zero-quantity spread set by a monopolistic dealer may differ markedly from that prevailing in a competitive market. The size of the spread therefore reflects transaction costs and dealer monopoly power in addition to the asymmetric information effect, and based solely on observation of the spread there are no obvious ways to resolve the three components.

Other empirical studies have concentrated on the second effect of the information asymmetry, the size of the price impact of the trade.[3] These analyses typically assume: serial independence of transactions, no delay in the effect of a trade on the price, and a linear trade-price relationship with the intercept corresponding to fixed transaction costs. There are good reasons for questioning these assumptions. Inventory control considerations induce serial dependencies in trades, as do price pressure effects and order fragmentation.[4] Lagged adjustment to new information and exchange-mandated price smoothing may lead to distribution over time of the information impact.[5] Finally, the form of the functional relationship between trade size and information is a consequence of a fundamentally unobservable cost and information structure, and many plausible structures entail nonlinear effects.

Recently, Hasbrouck (1988) and Stoll (1989) have attempted to resolve information and inventory control effects. Hasbrouck suggests that the two effects can be partially resolved according to the persistence of their impact on the security price. Inventory control effects are inherently transient, while the information inferred from a trade due to asymmetric information is permanently impounded in the stock price. A key point of the present paper is that transience characterizes not only inventory control effects but most other noninformation imperfections (price discreteness, price pressure, order

[2] See McInish and Wood (1988) and Chiang and Venkatesh (1988).
[3] See Glosten and Harris (1988) and Foster and Viswanathan (1988).
[4] See Garman (1976), Stoll (1976, 1989), Amihud and Mendelson (1980, 1982), Ho and Stoll (1980, 1981, 1983), and O'Hara and Oldfield (1986).
[5] See Goldman and Beja (1979) and Beja and Goldman (1980).

fragmentation, price smoothing, etc.) as well. The practical implication is that the information effect of a trade should be measured as that which persists over a substantial period.

The Hasbrouck (1988) paper also noted that if there were to be any private information inferred from a trade, it must be inferred not from the total trade but from that component which was unanticipated—the trade innovation. To investigate a distinctive feature of the inventory control hypothesis, the earlier paper used the univariate trade innovation, i.e., that component of the trade which was unexpected, based solely on knowledge of the past trade history. The present paper generalizes this to incorporate broader information sets (such as histories of the quote revisions and nonlinear functions of the trade variables) and thereby achieves a fuller picture of the information effect.

In summary then, the present paper suggests that trades and quote revisions be considered from an econometric perspective as a system characterized by auto- and cross-correlations of a very general nature. Within this framework, the information impact of a trade may be formally defined as the ultimate impact on the stock price (or quote) resulting from the unexpected component of the trade, i.e., the persistent price impact of the trade innovation. The persistent price impact of a trade on the price is preferred to the immediate impact because the latter may be contaminated by transient liquidity effects. Use of the trade innovation (rather than the total trade) as the driving force has the effect of excluding the predictable portion of the trade, which by definition conveys no new information. Relating the trade innovation to the ultimate price impact effectively removes from the analysis the transient effects stemming from a multitude of market imperfections which need not be modeled explicitly.

The primary statistical technique employed here is vector autoregression applied to quote and trade data. While use of nonlinear functions of the variables permits a rich characterization of their dynamics, the estimation technique is essentially least squares. The basic computations are therefore linear and relatively tractable, important considerations when faced with systems characterized by large numbers of parameters and observations. Thus, the techniques and modeling philosophy may be easily generalized and applied in many other situations. This paper considers generalizations incorporating nonlinear trade effects and the role of the spread.

This paper is organized as follows. Section I heuristically describes a simple bivariate linear time series model of trades and quote revisions and then motivates and interprets the vector autoregression (VAR) model. Section II presents an explication of the VAR representation for a simple microstructure model. Section III is a more critical analysis of some of the problems that can befall VAR estimation in microstructure applications. The paper then turns to empirical results. The data are discussed in Section IV, and a simple bivariate VAR estimation is presented in Section V. Section VI presents a more elaborate model which incorporates nonlinear trading effects. The preliminary cross-sectional properties of the price impact measures

are analyzed in Section VII. Section VIII discusses the role of the spread in ascertaining the impact of trades. A brief summary concludes the paper in Section IX.

I. The Trades/Quote Revision System

This section presents a heuristic development of the basic econometric specification. The starting point is found in the empirical predictions of the structured theoretical models of the asymmetric information effect cited in the introduction. By relaxing the strong economic assumptions of these models and adding statistical assumptions of a fairly general nature, the discussion will arrive at a robust empirical specification in which the impact of a trade on price due to asymmetric information is both meaningful and measurable.

The notation and sequencing conventions in the model are similar to Hasbrouck (1988). The market-maker posts bid and ask quotes, prices at which he is willing to buy or sell shares. Denote these quotes as q_t^b and q_t^a, where the timing convention is that these are set after the trade has occurred at time t. Therefore, the quotes prevailing before trading has taken place are q_{t-1}^b and q_{t-1}^a. A transaction is characterized by its signed volume x_t which is positive if the trade is a purchase (buyer-initiated) and negative if a sale (seller-initiated). Based on x_t, the market-maker posts new quotes q_t^b and q_t^a. The formal models generally assume that these quotes satisfy a zero-expected-profit condition for the market-maker. If there are no transaction costs, and the only update to the public information set at time t is the announced trade, then the revision in the quotes at time t summarizes the information inferred from the observation of x_t.

The primary price variable used in this paper is the midpoint of the quotes. To justify this as a meaningful variable, the initial assumption (subsequently relaxed) is that the quotes are set symmetrically about the expected value of the security conditional on all public information. More formally, define the value of the security at some convenient terminal time τ in the distant future as $\mathscr{X}_\tau$, and let Φ_t be the public information set at time t (including the quotes and trade at time t). Then the symmetry assumption is:

$$E\left[\left(q_t^b + q_t^a\right)/2 - \mathscr{X}_\tau \mid \Phi_t\right] = \left(q_t^b + q_t^a\right)/2 - E\left[\mathscr{X}_\tau \mid \Phi_t\right] = 0. \tag{1}$$

Under this assumption, the information inferred from the time t trade (x_t) may be conveniently summarized as the subsequent revision in the quote midpoint:

$$r_t = \left(q_t^b + q_t^a\right)/2 - \left(q_{t-1}^b + q_{t-1}^a\right)/2.$$

Transaction costs which are fixed and symmetric for purchases and sales are easily accommodated at this point: the identification of r_t with the information impact of x_t is not affected by a cost-based component of the spread.

The simplicity of the inference is, on the other hand, greatly affected by the presence of non-trade public information (news announcements). To maintain

the present notation, it will be assumed that the public information arrives after the *t*th trade but before the associated quote revision. (If the quotes are revised in the absence of a trade, let this be indicated by setting $x_t = 0$.) In this case, the quote revision reflects public as well as private information, and this vitiates the possibility of inferring the price impact of a particular trade. It is still possible, however, to assess the price impact of trades in some average sense. The assumption necessary to support inference in this new framework is that there is constancy over time in the function relating the trade (via the inferred private information) and the quote revision. This is tantamount to asserting that the quote revision is a stable function of the trade. To simplify certain aspects of the discussion, the initial assumption (subsequently modified) is that the dependence is linear. Thus, $r_t = bx_t + v_{1,t}$ where the disturbance $v_{1,t}$ reflects the public information. The price impact of the trade is measured by the coefficient b which may be estimated by regression.

This simple specification is characterized by a trade impact that is fully contemporaneous. To make this model more robust, it should be noted that many microstructure imperfections cause lagged effects. Price discreteness, for example, may induce threshold effects, since a quote revision may not be optimal until a series of trades of the same direction has occurred. (Additional effects of discreteness are discussed in Harris (1990).) Inventory control effects, lagged adjustment to information, and exchange-mandated price smoothing also involve serial dependencies. These considerations suggest a more flexible structure in which lagged values are included:

$$r_t = a_1 r_{t-1} + a_2 r_{t-2} + \cdots + b_0 x_t + b_1 x_{t-1} + \cdots + v_{1,t} \tag{2}$$

where the a's and b's are coefficients, and $v_{1,t}$ is the disturbance. In principle this representation may be of infinite order. As a practical empirical matter, it is truncated at some lag.

Serial correlation in the quote revisions is incompatible with the quote symmetry assumption of equation (1). The property of conditional symmetry at all times implied by (1) may instead be replaced by a weaker version of symmetry that holds for the expectation of quotes at some future time s conditional on the current (time t) information set:

$$\text{As } s \to \tau, \ E\big[\big(q_s^b + q_s^a\big)/2 - \mathscr{X}_\tau \,|\, \Phi_t\big]\big) \to 0. \tag{3}$$

The intuition behind this requirement is that while a quote revision model of the sort specified in equation (2) permits deviations between the quote-midpoint and efficient prices, these deviations are transient. With the passage of time, all rational agents expect quotes to revert (on average) to the fair value of the security. As will become apparent, the present analysis relies heavily on the expectation of the future quotes implied by the dynamic model. This assumption ensures that by carrying the projection out sufficiently far into the future, we arrive at the current efficient price.

To this point, trades have been assumed to be exogenous and wholly unanticipated. The same microstructure effects cited in the discussion of the quote revisions will also lead to serial dependencies in trades, and in consideration trades may be modeled in a similar fashion:

$$x_t = c_1 r_{t-1} + c_2 r_{t-2} + \cdots + d_1 x_{t-1} + d_2 x_{t-2} + \cdots + v_{2,t} \tag{4}$$

The disturbance in this equation, $v_{2,t}$, captures the unanticipated (innovative) component of the trade (relative to an expectation formed from linear projection on the trade and quote revision history). As such, if there is any private information to be inferred from a trade, it must reside in this innovation. Together equations (2) and (4) comprise a bivariate vector autoregression model (see Judge, Griffiths, Hill, Lutkepohl, and Lee (1985)). It is assumed that the disturbances have zero means and are jointly and serially uncorrelated, viz.

$$Ev_{1,t} = Ev_{2,t} = 0,$$
$$Ev_{1,t}v_{1s} = Ev_{2,t}v_{2s} = Ev_{1,t}v_{2,s} = 0, \quad \text{for } s \neq t. \tag{5}$$

Compared to the usual VAR specification, the model given by equations (2), (4), and (5) is slightly unusual in that the r_t specification includes the contemporaneous value of x_t. This implies, since the coefficients in (2) and (4) are linear projection coefficients, that $Ev_{1,t}v_{2,t} = 0$. This reflects the fact that the revision and trades are not determined simultaneously: the quote revision follows the trade, and r_t cannot contemporaneously influence x_t. One of the subsidiary purposes of this study is the analysis of Granger-Sims causality patterns in the data. There is, of course, a strong presumption of causality running from trades to quote revisions, and the structure permits this both contemporaneously and with lags. The structure also permits causality running from lagged quote revisions to trades. It does not permit, due to the above timing considerations, contemporaneous causality running from quote revisions to trades.

To see the role of public and private information in this model, consider first the trade representation of equation (4). Based on knowledge of the trade and price history, agents can use (4) to form an expectation of the *tth* trade. If there is any new information contained in x_t, it must reside in the innovation $v_{2,t}$, since the remaining component is entirely known. This does not imply that the innovation is a deterministic function of the new information. The presence of uninformed liquidity traders, for example, will introduce a noise component of $v_{2,t}$ that is uncorrelated with private information. Indeed unless some noise is present, the market-maker will only lose to arriving traders.

To assess the information content of a trade innovation in this model, suppose that at time $t = 0$ the system is in a stable state: all lagged values of r_t and x_t are zero, and the prevailing quotes are set to the unconditional expectation of the security's value, $E[(q^b_{-1} + q^a_{-1})/2] = E[\mathcal{Z}_\tau]$. At $t = 0$, a signed order $v_{2,0}$ arrives, making the current trade $x_0 = v_{2,0}$. By letting

$v_{1,t} = 0$ for all t and $v_{2,t} = 0$ for $t > 1$, and iterating on equations (2) and (4), it is possible to compute $E[r_t \mid v_{2,0}]$ and $E[x_t \mid v_{2,0}]$ for $t \geq 0$. The sum of the predicted quote revisions through step m,

$$\alpha_m(v_{2,0}) = \sum_{t=0}^{m} E[r_t \mid v_{2,0}], \tag{6}$$

gives the expected cumulative quote revision through the mth step. By (3), as m increases, $\alpha_m(v_{2,0}) = \sum_{t=0}^{m} E[(q_t^b + q_t^a)/2 - (q_{t-1}^b + q_{t-1}^a)/2 \mid v_{2,0}] \rightarrow E[\mathscr{L}_\tau \mid v_{2,0}] - E[\mathscr{L}_\tau]$. That is, the expected cumulative quote revision con- verges to the revision in the efficient price. For this reason, $\alpha_m(v_{2,0})$ pos- sesses an interpretation as the information revealed by the trade innovation and constitutes the basic construct underlying this paper.

In summary then, the VAR modeling strategy applied to trades and quotes allows, in principle at least, a resolution between private information (the trade innovation) and public information (the quote revision innovation). Furthermore, by explicitly reflecting the dynamic structure of the system, the VAR model permits calculation of both the contemporaneous and lagged impacts of an innovation. The persistent impact of a trade innovation on the quote has a natural interpretation, therefore, as the information content of the innovation net of transient effects.

II. The VAR Representation of a Simple Microstructure Model

Although the VAR framework proposed in the last section is very general, it is useful to consider the representation of a particular model. This section analyzes a simple microstructure model (suggested by Larry Glosten) which exhibits both asymmetric information and inventory control behavior and presents a numerical example which demonstrates the superiority of the VAR analysis over other plausible alternatives.

Consider the following framework. Let m_t be the efficient stock price, the expected value of the stock conditional on all public information. The evolu- tion of m_t is given by

$$m_t = m_{t-1} + z v_{2,t} + v_{1,t}, \tag{7}$$

where $v_{1,t}$ and $v_{2,t}$ are mutually and serially uncorrelated disturbance terms. Intuitively, $v_{1,t}$ stems from public nontrade information, and by what follows $v_{2,t}$ is the innovation in trades. The z coefficient reflects the private informa- tion conveyed by this innovation.

The quote-midpoint price is q_t, with evolution given by

$$q_t = m_t + a(q_{t-1} - m_{t-1}) + bx_t, \tag{8}$$

where x_t is the signed trade at time t, and a and b are adjustment coefficients with $0 < a \leq 1$ and $b > 0$. This process has an inventory control interpretation. To see this, assume that at time $t = 0$, $q_0 = m_0$. If $x_1 > 0$ (a purchase), then the market-maker will raise the quotes in excess of m_1 in an

attempt to elicit sales. The case of $a < 1$ is associated with imperfect inventory control: competition from public limit order traders, for example, forces q_t to move closer to m_t with the passage of time. If $a = 1$, then trades are completely reversed over time. The final equation in the model describes the evolution of trades:

$$x_t = -c(q_{t-1} - m_{t-1}) + v_{2,t}, \tag{9}$$

where $c > 0$ defines a downward sloping demand schedule, and $v_{2,t}$ is the innovation (unexpected component) of the trade. From equation (7), the response of the efficient price to the trade innovation is $z v_{2,t}$, and so z is the measure of information asymmetry in this model.

From the econometrician's viewpoint, the efficient price m_t is unobservable. The econometric model, therefore, must involve only x_t and q_t, or (to be consistent with the VAR framework of the last section) x_t and r_t, where $r_t = q_t - q_{t-1}$. To expedite the discussion, it is useful to consider $a < 1$ and $a = 1$ as separate cases. Taking the first, it may be shown that the model implies the following autoregressive representation:

$$
\begin{aligned}
r_t = (z + b)x_t &+ \left[zbc - (1-a)b \right] x_{t-1} \\
&+ a \left[zbc - (1-a)b \right] x_{t-2} + \cdots + v_{1,t}
\end{aligned} \tag{10}
$$

and

$$x_t = -bcx_{t-1} - abcx_{t-2} - a^2 bcx_{t-3} + \cdots + v_{2,t}. \tag{11}$$

The order of this vector autoregression is infinite, but as long as $a < 1$, the coefficients are declining.

Given the innovations representations for r_t and x_t, it is possible to compute (to any degree of accuracy) the autocovariances for and cross-covariances between the two series. From these, one may in turn compute the regression coefficients that would be obtained in truncated specifications corresponding to equations (10) and (11). These are highly nonlinear and complicated functions of the original parameters. Accordingly, for illustration purposes, consider a specific numerical example. Let $a = 0.6$, $b = 1$, $c = 0.5$, $\text{Var}(v_{1,t}) = \text{Var}(v_{2,t}) = 1$, and let the value of asymmetric information coefficient be given as $z = 1$. The following discussion considers three estimation strategies.

Suppose first that the econometrician believes in the complete absence of inventory control or other transient liquidity considerations. In this case, the impact of a trade on price could be assessed by estimating the regression $r_t = \hat{z}x_t + \hat{v}_{1,t}$, where the "$\,\hat{}\,$" symbol is used to indicate that the parameter now belongs to a model that is incorrectly specified. The regression coefficient $\hat{z} = \text{Cov}(r_t, x_t)/\text{Var}(x_t)$, and for the particular numerical values used here $\hat{z} = 1.960$. This is a gross overestimate of the true value of unity and arises as a consequence of the contemporaneous "overreaction" of the price due to inventory considerations.

Now suppose that the econometrician still disavows any inventory effects but believes that the price exhibits lagged adjustment to trades. This sug-

Table I

The Behavior of Prices and Trades in a Simple Model

This table is based on the microstructure model presented in Section II. The truncated reduced-form vector-autoregressive specification corresponding to this model is:

$$r_t = 1.998\, x_t + .096\, x_{t-1} + .053\, x_{t-2} + .025\, x_{t-3} + \hat{v}_{1,t}$$

$$x_t = -.486\, x_{t-1} - .269\, x_{t-2} - .124\, x_{t-3} + \hat{v}_{2,t}$$

where r_t is the quote-revision, x_t is the trade, $\hat{v}_{1,t}$ is the quote-revision innovation, $\hat{v}_{2,t}$ is the trade innovation, and t indexes transactions. Based on an initial innovation at transaction $t = 0$ of $\hat{v}_{1,t} = 0$ and $\hat{v}_{2,t} = 1$ (i.e., a purchase of one share), the table gives the implied response of r_t and x_t. Quantity α_m is the cumulative quote revision through transaction m: $\alpha_m = \sum_{i=0}^{m} r_i$.

t	r_t	x_t	α_m
0	1.998	1.0	1.998
1	$-.874$	$-.486$	1.124
2	$-.060$	$-.033$	1.064
3	.041	.023	1.105
4	.105	.058	1.210
5	$-.054$	$-.030$	1.155
⋮			
10	-6×10^{-4}	-3×10^{-4}	1.156

gests regressing r_t against current and lagged x_t in an autoregression similar to equation (10) but truncated (for practicality) at some finite lag. If the regression is truncated at lag 3, the regression coefficients are given by:

$$r_t = 1.998\, x_t + 0.096\, x_{t-1} + 0.053\, x_{t-2} + 0.025\, x_{t-3} + \hat{v}_{1,t}. \qquad (12)$$

Assuming (in the absence of inventory considerations), that trades are uncorrelated, the derived impact of a unit trade on the price is given by the sum of the trade coefficients, 2.172. As an estimate of z this is even worse than that of the previous analysis.

Finally, suppose that the econometrician adopts the full VAR modeling strategy advocated here, and also estimates a trade autoregression (truncated at lag 3). The regression coefficients for the latter representation are:

$$x_t = -0.486\, x_{t-1} - 0.269\, x_{t-2} - 0.124\, x_{t-3} + \hat{v}_{2,t}. \qquad (13)$$

To estimate the asymmetric information effect using the dynamic strategy, suppose that all lagged x_t are initially zero and that at time $t = 0$, $x_0 = 1$. Table I describes the dynamic motion of the system through time $t = 10$. In accordance with the inventory effects, the initial shocks to both x_t and r_t are partially reversed in subsequent periods. Taking time $t = 10$ as the "limiting point" of the system, it would be concluded that one unit trade causes a 1.156 unit increase in the stock price. This is much closer to the true value of unity than either of the first two approaches. Although this demonstration is based on one set of parameter values, it does indicate that VAR estimation may be

markedly superior to estimation methods which do not adequately take into account the dynamic structure of the market.[6]

Turning now to the case where the adjustment coefficient in the price equation (8) is $a = 1$, it may be shown that the sum of the coefficients in the moving average representation for trades is zero. In this case, the trade series does not possess an autoregressive representation. The market-maker inventory series does possess such a representation, however, and may be used in lieu of trades.

As a final point, the VAR representation for this simple example does not exhibit dependence on lagged quote revisions. With a fair increase in the complexity of the model, it would be possible to introduce such dependence by permitting q_t in (8) to depend explicitly on contemporaneous and lagged $v_{1,t}$ (public information) or by permitting x_t in (9) to depend on lagged $v_{1,t}$. To permit x_t to depend on contemporaneous $v_{1,t}$, however, would be tantamount to impounding public information in the unexpected component of the trade. As will be shown in the next section, this causes problems for the analysis.

III. The Model Specification: A Critique

Section I proposed a bivariate VAR model of trades and quote revisions comprising the linear specifications of equations (2) and (4) and the disturbance properties in (5). As a measure of asymmetric information, the section proposed the cumulative quote revision caused by a trade innovation, $\alpha_m(v_{2,0})$. While the supporting discussion motivated many of the features of the model in a heuristic fashion, the construction was not a formal development from first principles. The model is therefore almost surely an approximation, and it is necessary to reflect on its adequacy. The present discussion will consider two sorts of issues. The first problems are econometric and concern the potential for biased or inconsistent estimates of the VAR system. The second set of issues is interpretive and concerns the conditions under which $\alpha_m(v_{2,0})$ does in fact measure asymmetric information.

A. Econometric Concerns

From an estimation perspective, the crucial feature of the model, which follows from the disturbance structure assumed in equation (5), is that the disturbances ($v_{1,t}$ and $v_{2,t}$) are uncorrelated with the regressors (the lagged, and in one instance contemporaneous, values of x_t and r_t). This claim rests on the assumption of stationarity. If the quote revisions and trades $\{r_t, x_t\}$ are jointly covariance stationary, then the Wold theorem ensures that the model may be written as a joint moving average process of infinite order. If this moving average representation is invertible (that is, if its current

---

[6] The estimation error in the VAR approach is due to the truncation of the autoregressions, and with longer lags better accuracy is achieved. If the regressions are truncated at lag 5, the implied price change through time $t = 10$ is 1.053. If the regressions are truncated at lag 7, this value becomes 1.019.

innovation can be computed as a convergent weighted sum of the current and past realizations), then by definition the series possess an infinite autoregressive representation, i.e., one of the form (2) and (4) in which the coefficients converge.

B. Interpreting $\alpha_m(v_{2,0})$ as a Measure of Private Information

In the formal models of asymmetric information, the trade is driven partially by private information and partially by liquidity needs, but in no part is the trade driven by public information which is relevant to forecasting the value of the security. This intuition underlies the dichotomy in the present VAR model of equations (2) and (4) that identifies all public information with the quote revision innovation ($v_{1,t}$) and all private information (plus an uncorrelated liquidity component) with the trade innovation ($v_{2,t}$). This dichotomy asserts that $v_{2,t}$ reflects no public information and so justifies the interpretation of the impulse response function $\alpha_m(v_{2,0})$ as a measure of the private information. In practice, of course, the dichotomy is not likely to be as clean, and estimated values of α_m may in consequence either fail to capture the response of the quote to private information or else may reflect the response of the quote to public information. The following discussion is directed at some of the imperfections which might upset the dichotomy. Although the remarks specifically address concerns in the present VAR model, it should be emphasized that they pervade most analyses which seek to relate trading activity and security prices.

The first part of the dichotomy asserts that the quote revision innovation reflects only public information. Since the quotes are provided by dealers or limit-order traders, this assertion may be violated if either of these parties possesses superior information. While specialists do not typically engage in analysis of the sort that is likely to directly uncover corporate financial information, they do possess an informational advantage in their knowledge of the limit order book. The extent of information production by limit order traders is unknown. (However, given that the value of private information generally dissipates over time, there are incentives for informed traders to prefer the execution certainty of market orders.) In this event, the impulse response function measures the private information inferred from the trade relative to the information impounded in the prevailing quotes. The latter may reflect the private information of the quote setters, which of course will not be captured by the impulse response measure.

The other aspect of the dichotomy, that the trade innovation contains no public information, may be formalized as the requirement that public information is not useful in predicting the trade innovation. That is, letting Φ_t be the public information immediately subsequent to the time t quote revision,

$$E\left[v_{2,t+k} \mid \Phi_{t-1}\right] = 0, \qquad \text{for } k \geq 0. \tag{14}$$

Since the public information set contains variables such as earnings and merger announcements which may affect trading decisions, it is not immedi-

ately clear that this assumption is even approximately correct. But equation (14) does not rule out such relations: public information may affect trades via the dependence of x_t on lagged r_t.

In actuality, it appears that violations of (14) are most likely when there are market features which impair the quote revision process and thereby constrain the quote revisions from fully reflecting public information. This discussion considers effects resulting from specialist smoothing, stale quotes, and stale limit orders.

The New York and American Stock Exchanges include as part of the market-maker's affirmative obligation a requirement that prices not be "too discontinuous". If the market-maker must show a smooth price transition path, then quotes may not be immediately revised to reflect public information. Suppose, for example, that all agents have homogenous information and that a positive earnings forecast becomes public knowledge. The prior quotes do not reflect this information, and after the announcement traders would have an incentive to buy at the prior ask price (if they could). It is usually assumed that the market-maker can immediately adjust the quotes to reflect the announcement and that at these new quotes the liquidity traders would have no greater incentive to buy than they did before. In the event of such an immediate adjustment, the econometrician analyzing this sample would detect no trading correlated with the quote revision and would correctly infer the absence of asymmetric information.

But if the market-maker is compelled to set the quotes in such a way as to ensure a smooth price adjustment path, then he may not be able to revise the quotes as fully and as immediately as unconstrained use of the news would imply. To take an extreme case, suppose that the market-maker is enjoined from raising the quotes *except* in response to a trade. He would be forced to maintain the prior quotes until a trader submitted a buy order. He would then move the quotes up by the prescribed maximum amount (say 1/8 of a point). If this price were still "too low" (relative to the public information), then the next trade would also tend to be a buy order. The quotes would then be raised by another 1/8 and so on until the correct quotes were reached. In this sample, the econometrician would detect a relation between trades and the subsequent quote revisions and would incorrectly attribute this relation to asymmetric information.

Although price smoothing is a likely feature of the trading process, there are several considerations which are likely to mitigate the severity of the problem. The first point is one of clarification. The term "price smoothing" is sometimes taken to describe any regime which exhibits lagged price adjustment. In this connection, it should be emphasized that equation (14) is not inevitably violated whenever there is lagged price adjustment. The model discussed in Section III, for example, exhibits lagged adjustment (in (10) the current quote revision depends on lagged trades), but the VAR analysis nevertheless leads to correct inference. The inferential problem is a consequence instead of the market-maker's inability to use public information optimally in setting the quotes and the consequent public information component of the trade innovation.

Secondly, the problem of smoothing in response to public information is potentially severe only for relatively small trades. In bridging a large price jump by transactions at prices separated by 1/8 of a point, the market-maker is not required to satisfy total demand at the intermediate prices. Small trades at these prices suffice to satisfy the smoothing obligation. As a final and third consideration, note that when faced with major news or announcements, the market-maker has the option (with the permission of the exchange floor governor) of closing and reopening the market. The obligation to smooth prices does not extend over the period of closure.

Another apparent stickiness in the quote revision process arises from stale reported quotes. In the reporting process, priority is given to trades over quotes. While the traders in the crowd on the floor may know (and base their orders on) the correct quotes, these quotes will not be reflected in the data base. These verbal quotes constitute part of the public information set but are not reflected in the posted quote history. As such, a trade that appears to have been unexpected from the researcher's viewpoint may in fact have been a response to an unreported quote.

Stale limit orders might appear to cause similar problems. A trader placing a limit order at some distance from the market price is committing to trade at the limit price. Presumably due to the costs of continuously monitoring the market, the trader is in effect employing a trading rule based only on the security price. This is a subset of the total public information set, and it must be admitted that an update to nonprice public information may cause a limit order to be distinctly suboptimal relative to what the absent limit-order trader would prefer.

From these remarks it should be clear that in the present application (as in many others) the distinction between public and private information is more clearly visible in formal models than in practice. The announcement over the Dow-Jones wire of a merger bid, for example, is conventionally taken as an update to the public information set. Nevertheless, if public information is considered to be that upon which all agents condition their actions, then a merger bid which leaves an absent trader's limit sell order exposed at a ridiculously low price is still private information: it was not incorporated into at least one trader's decision. Similar considerations apply to the quote representations which are made in the trading crowd but never get disseminated off the floor.

As a final point, it should be emphasized that the VAR innovations are computed on the basis of an assumed linear model. While the model will be generalized in later sections to accommodate certain nonlinearities, the computation of the innovations will always be contingent on the specification. Since interactions between trades and quote revisions may be picked up by one specification and not by another, misspecification may further upset the dichotomy.

In summary, market imperfections which blur the identification of public and private information with the quote revision and trade innovations, respectively, do not lead to inconsistent econometric estimates. They may lead, for a variety of reasons, to incorrect attribution of private and public

information effects. Such considerations are likely to permeate all analyses which attempt to relate trading and price movements, however, and are not idiosyncratic to VAR models.

IV. Data Sources

The results reported in this paper are based on a sample of transactions data from the Institute for the Study of Security Markets (ISSM).[7] These data comprise a sequenced trades and quote record from the New York and American Stock Exchanges and the consolidated regional exchanges, over the 62 trading days in the first quarter of 1989. The sample of firms for the present study was selected in the following manner. For all issues present on the ISSM and 1988 daily CRSP files, I computed total market values (end of 1988), formed quartile market value subsamples, and sorted each quartile by sequence on the ISSM tape (roughly alphabetical). I applied all analyses to the first 20 firms in each quartile which had at least 500 transactions, average price greater than $5, and less than 10% change in the number of shares outstanding over the sample period.

In constructing the time series of trade and quote revisions, every attempt was made to keep the notation consistent with the model discussed in earlier sections, subject to certain inherent limitations. The clock ("t") for the model is transaction sequence. Thus, the model is specified in "transaction time". The first observation is the first trade or quote for the first trading day of 1989, and thereafter t is incremented each time a trade or quote is posted. There are several exceptions to this rule. First, the appearance of a quote revision within 5 seconds prior to a transaction was viewed as a reporting anomaly, and, in the analysis, the quote is resequenced after the trade. Second, if a quote revision occurs within 15 seconds following a trade, I assign it the same t subscript. Finally, I ignore quotes posted by the regional exchanges, which tend to follow those of the principal exchange quite closely. Determination of the direction of a trade is made on the basis of the prevailing quote. A transaction price above the prevailing quote-midpoint is assumed to be a purchase (positive sign), and vice versa for prices below the midpoint. If the transaction price is exactly at the midpoint, the direction of the trade is indeterminate, and x_t is set to zero. At the daily opening, x_t is also set to zero.

With respect to the method of clocking the observations, it should be emphasized that the present study is conditioned on the interpretation of t as an index of events (trades or quotes). Ideally, one would prefer to incorporate wall-clock time dependencies as well, but this would necessitate a substantial increase in econometric complexity. For the VAR model to be estimated

---

[7] An earlier version of this paper employed a different data source, a transaction record compiled by a firm under contract to the American Stock Exchange, and used by Hasbrouck and Ho (1987), Hasbrouck and Schwartz (1988), and Hasbrouck (1988). The results based on the ISSM data set are highly similar to those obtained from the earlier data source.

consistently, it is only necessary that the time series be covariance stationary with respect to the time index used. The crucial assumption here is that data are stationary in transaction time. This assumption does not rule out the existence of conditional nonstationarities due to time-of-day effects, although they are of course ignored in the modeling.

V. Estimation of a Simple Trade/Quote Revision Model

Much of the appeal of the VAR approach lies in its potential for generalization to encompass subtleties of the trade-quote interactions. While later sections will present some of these generalizations, it is useful for expositive purposes to begin the empirical analysis with a simple bivariate model similar to the one outlined in the preceding sections. For reasons of brevity here and throughout the discussion of the empirical work, I will present detailed results for a single representative stock and summary results for the larger sample.

The VAR model proposed in Section I is linear in the signed trade x_t. In fact, when such model is estimated for different firms, the impact of trades on quote revisions is found to be generally positive (as expected) but of highly variable magnitude. This appears to be a consequence of a small number of large trades. A better-behaved bivariate model results from replacing x_t with an indicator variable which simply measures the direction (purchase or sale) of the trade. This variable is $x_t^0 = +1$ (if $x_t > 0$), 0 (if $x_t = 0$), -1 (if $x_t < 0$). The VAR model therefore involves equations in which r_t and x_t^0 are regressed against lagged values. This (and all VAR models discussed here) are estimated using ordinary least squares estimates of the multiple-equation system.[8]

The representative stock is Ames Department Stores (ADD) which has an average market value of \$581 million (placing it in the second highest quartile subsample) and average share price over the sample period of \$15.53. Table II presents the coefficients of the bivariate VAR estimated through five lags. The most important set of coefficients are those of $x_t^0, \cdots, x_{t-5}^0$ in the r_t equation. The coefficient of x_t^0 implies that on average the quote midpoint is raised by roughly \$0.014 immediately subsequent to a purchase order. The coefficients at longer lags are generally positive but decreasing.

---

[8] As defined, x_t^0 is a limited dependent variable. This is in no sense incompatible with the assumptions required for specification and consistent estimation: if $\{x_t^0, r_t\}$ are jointly covariance stationary and invertible, then a VAR model of the form given in equations (2), (4), and (5) exists. The discreteness of x_t^0 does impose restrictions on the distribution and moments of the disturbances, however. For example, although the $v_{2,t}$ are serially uncorrelated, they are not in general serially independent. The disturbance properties given in (5) also ensure that the coefficients in (2) and (4) may be consistently estimated using ordinary least squares. Of course, the VAR coefficient estimates obtained in any given sample may quite possibly imply $Ev_{2,t} \neq 0$ for certain out-of-sample data points. This is a consequence of estimation error.

Table II
Estimates of the Bivariate Vector Autoregressive Model for
Ames Department Stores

The model is:

$$r_t = \sum_{i=1}^{5} a_i r_{t-i} + \sum_{i=0}^{5} b_i x_{t-i}^0 + v_{1,t} \quad x_t^0 = \sum_{i=1}^{5} c_i r_{t-i} + \sum_{i=1}^{5} d_i x_{t-i}^0 + v_{2,t}$$

where r_t is the price (quote-midpoint) change in dollars; x_t^0 is a trade indicator variable ($+1$ for a buy order and -1 for a sell order); and t indexes transactions. The linear regression coefficient estimates are based on all transactions for the first quarter of 1989.

	Coeff.	T-Stat.		Coeff.	T-Stat.
a_1	$-.118$	-7.22	c_1	-3.647	-13.44
a_2	$-.011$	$-.67$	c_2	-1.689	-6.05
a_3	$-.013$	$-.17$	c_3	$-.505$	-1.80
a_4	$-.022$	-1.31	c_4	$-.130$	$-.46$
a_5	$-.002$	$-.14$	c_5	$.115$	$.41$
b_0	$.014$	15.15			
b_1	$.007$	6.83	d_1	$.167$	10.16
b_2	$.000$	$.46$	d_2	$.120$	7.20
b_3	$.001$	$.87$	d_3	$.078$	4.66
b_4	$-.000$	$-.30$	d_4	$.021$	1.24
b_5	$.001$	$.94$	d_5	$.034$	2.03
$R^2(r_t)$	$.096$		$R^2(x_t^0)$	$.085$	

The second important feature in the estimation is the strong positive autocorrelation in trades reflected by the lagged x_t^0 coefficients in the x_t^0 estimation. This is consistent with the findings of Hasbrouck and Ho (1987) and Hasbrouck (1988) and suggests simply that purchases tend to follow purchases, and similarly for sales. This pattern of positive autocorrelation at low lags is highly typical. There is, in particular, a striking absence at low and moderate lags of the trade reversal consistent with inventory control mechanisms. The short-run predominance of positive autocorrelation is more consistent with lagged adjustment to new information.

The nature and duration of the autocorrelation in trades is an important consideration in the specification and estimation of these models. While evidence for trade reversal is found in Hasbrouck (1988), it arises only in large pooled samples of relatively low market-value stocks. Furthermore, even in these samples, the reversal is very weak and is distributed over very long lags (up to 200 transactions). For individual stocks, the reversal pattern is simply too weak to estimate reliably in samples of the present size. This consideration underlies the decision to truncate the specifications at lags below, which reversal might be presumed to operate.

The data limitations do not preclude, however, a consideration of how trade effects at long lags might affect the analysis. In the first place, given the weakness of the trade reversal, it is unlikely that addition of trades at long

lags would add significant explanatory power to the trade or quote revision estimations. It is unlikely, therefore, that the estimated coefficients would be significantly altered. The omission may be more serious, however, for the implied cumulative quote revisions. Since α_m is cumulative, over long periods even very weak effects can add up to significance. When long-run trade reversal is present, therefore, the quote impact calculations implied by VAR estimates of low order are likely to overstate the long-run price impact: the model will catch the initial positive impact of a trade on the quote, but will miss the subsequent long-run reversion. If this problem exists, it is most likely to be present in the estimations for stocks of low market values and may cause the estimates of the price impact to be biased upwards.

The positive impact of trades on quotes and the positive autocorrelation of trades are the primary factors determining the model's dynamic adjustment path. Two other related features, however, are also noteworthy. These are the patterns of negative lagged r_t coefficients in both the r_t and x_t^0 specifications. From a purely descriptive viewpoint, this pattern in the r_t estimation implies negative autocorrelation in the quote revisions. In the x_t^0 estimation it implies Granger-Sims causality running from quote revisions to trades.[9]

In the present application, there are several ways of interpreting these findings. In the first place, they may simply arise from measurement error. The sign of the trade variables is determined by comparing the transaction price to the prevailing quotes. If there is a reporting error in the quotes, subsequent trades may be misclassified, and the observed causality may simply reflect correlation of measurement errors. The nature of the correlation depends on the nature of the reporting errors. The most likely source of errors is stale quotes. That is, in the face of rapid trading activity the transactions are reported promptly, while the quotes are not. While this may cause misclassified trades, it should not lead to any obvious dependence of the (mismeasured) trade on the lagged (mismeasured) quote revision nor should the delay introduce negative autocorrelation in the quote revisions. On the other hand, quote reporting errors which are of a random nonpersistent nature may be capable of generating the observed behavior.

The causality is, on the other hand, consistent with certain aspects of current microstructure theory. A negative relation between trades and lagged quote revisions is consistent with inventory control effects since a monopolistic market-maker with an inventory surplus would reduce his quotes to elicit more purchases. It is also consistent with the price experimentation hypothesis suggested by Leach and Madhavan (1989) in which the market-maker sets quotes to extract information optimally from the traders. These possibilities are deserving of further study.

---

[9] When dealing with two jointly covariance stationary time series x and y, x is said by Granger-Sims to cause y if knowledge of past x and y leads to better predictions of y than would result from knowledge of past y alone. See Geweke, Meese, and Dent (1983). In the r_t estimation, a formal Wald test rejects the null hypothesis that the quote revision coefficients are zero with a tail probability $p < 0.0001$.
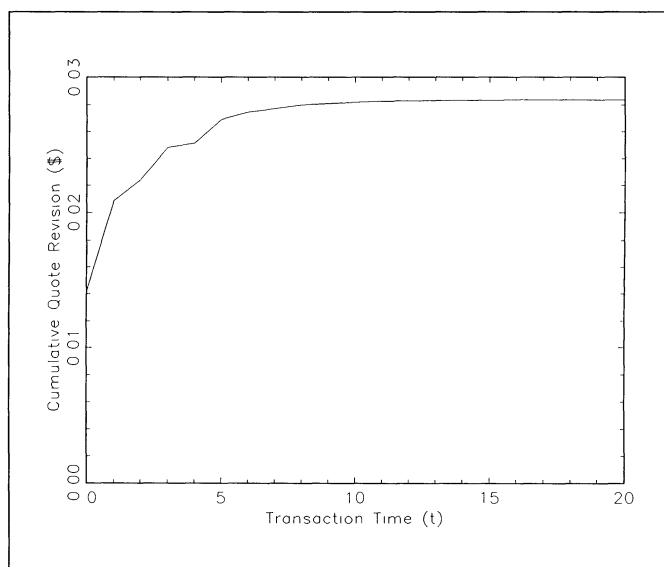
Figure 1. The quote revision process for Ames Department Stores implied by a bivariate model. The figure depicts the cumulative quote revision for a share of Ames Department Stores implied by the bivariate vector autoregression model presented in Table II, subsequent to an initial buy order. The index t refers to transactions.

The dynamic properties of the system may be illustrated by examining the response of the quote midpoint to a purchase at time $t = 0$ ($x^0_{t=0} = +1$). The cumulative quote revision is graphed in Figure 1 through $t = 20$, at which time it has converged to roughly \$0.028. The convergence is rapid (after $t = 5$, the preponderance of the adjustment is complete), but it is far from instantaneous. This suggests that transient dynamic considerations, to which VAR models are particularly well-suited, are important.

VI. Nonlinearities in the Response of Quotes to Trades

The VAR systems discussed to this point have been bivariate models based on the assumption that quote revisions are linear either in trade size or trade sign. This assumed linearity is a tenuous approximation at best. Not only are the underlying cost and information functions unlikely to be linear but the order entry and trade negotiation processes are almost certainly dependent on trade size. Models which are more realistic but still amenable to linear estimation may be developed by moving to multivariate VAR specifications which are linear in nonlinear transformations of the trade variable. This section discusses the development and estimation of such a model.

In view of the fact that a general function may be approximated by a polynomial expansion, one obvious specification involves regressing the quote revision against current and lagged signed powers of the trade variable. The

model analyzed here is quadratic. In addition to the signed trade variable x_t and the indicator variable x_t^0 introduced earlier, I add $x_t^2 = x_t^0 \mid x_t \mid^2$, a large-trade variable. The full quadratic VAR model comprises four linear equations in which each variable in the set $\{r_t, x_t^0, x_t, x_t^2\}$ is regressed against lagged values of the entire set. In practice, however, x_t^2 was found to be statistically significant only in the r_t estimations. Furthermore, none of the other variables were found to be significant linear predictors of x_t^2. Therefore, no equation was estimated for this variable; in effect, it was treated as exogenous.

The quadratic VAR model was estimated for all 80 issues in the sample. While the cross-sectional properties of these estimations will be discussed in the next section, specific estimation results are reported here for ADD. For reasons of brevity, these estimates are further summarized by equation and groups of like variables. Table III reports for each equation and each type of variable: the sum of the coefficients across all lags, the *T*-statistic associated with this sum, and the *p*-value (right-hand tail value) associated with the Wald test of null hypothesis that all coefficients in the group are zero. Full estimation results are available on request.

In the r_t estimation, the coefficient sums for x_t^0 and x_t are positive, and that for x_t^2 is negative: as a function of trade size, the price impact is positive, increasing, and concave. This concavity is quite typical of all the stocks in the sample and may arise in part from differences in market operation for trades of extreme sizes. The asymmetric information problem may be different for large block trades (which are negotiated between participants in a loose ongoing business relationship) and small trades (which are relatively anonymous). In the trade variable regressions, the pattern of positive auto-correlation is most pronounced in the indicator variable x_t^0 and less striking in the x_t. The coefficients of lagged quote revisions are generally negative. The Granger-Sims causality finding discussed in the previous section is strongly confirmed here: across the entire sample, the null hypothesis that r_t did not Granger-cause x_t^0, x_t, or x_t^2 was rejected at the 0.05 level in 73 of the 80 cases using a Wald test. Finally, to illustrate the dynamic and size-dependent properties of the system, Figure 2 presents the cumulative quote revision subsequent to purchases of 1,000, 10,000, and 50,000 shares.

VII. Cross-Sectional Effects

While the assessment of the extent of informational asymmetry for a particular stock is a problem of considerable interest in its own right, it is a precursor to a larger class of issues that concern the relations between the extent of the asymmetry and other characteristics of the stock and the market structure. This section presents a simple cross-sectional analysis.

It was noted in the introduction that the information asymmetry measure employed here is a summary measure and that the theoretical models suggest that it should be a function of the relative proportion of potentially

Table III

Estimates of the Quadratic Vector Autoregressive Model for Ames Department Stores

The set of variables in the model is $\{r_t, x_t^0, x_t, x_t^2\}$ where r_t is the price (quote midpoint) change; x_t^0 is a trade indicator variable; x_t is the signed trade volume; $x_t^2 = x_t^0 \mid x_t \mid^2$; and t indexes transactions. The general specification for each equation is:

$$(*)_t = \sum_i a_i r_{t-i} + \sum_i b_i x_{t-i}^0 + \sum_i c_i x_{t-i} + \sum_i d_i x_{t-i}^2 + v_{*,t}$$

where $(*)$ indexes the set of variables in the model. No equation was estimated for x_t^2. The table contains summary statistics for each group of linear regression coefficients. In each group, "*lags*" specifies the lags used in the specification, "p" is the right-tail probability for a Wald test of the null hypothesis that all coefficients in the group are zero, "Σ" is the sum of the coefficients in the group, and a T-statistic for this sum is reported in parentheses. The estimates were based on all transactions for the first quarter of 1989.

Dependent Variable (*): Coefficient Group		r_t	and x_t^0	x_t
a_i	*lags*	1–5	1–5	1–5
	p	< .001	< .001	.28
	Σ	−.210	−6.18	−83.76
		(−5.01)	(−8.77)	(−.83)
b_i	*lags*	0–5	1–5	1–5
	p	< .001	< .001	.003
	Σ	.0160	.402	11.07
		(8.21)	(13.15)	(2.52)
c_i	*lags*	0–5	1–5	1–5
	p	< .001	.267	.718
	Σ	.00022	.00055	.0511
		(8.15)	(2.12)	(1.37)
d_i	*lags*	0–5		
	p	< .001		
	Σ	-6.59×10^{-8}		
		(−3.95)		
R^2		.125	.086	.006

informed traders, the probability that someone has in fact observed a private information signal, and the precision of the private signal. These determinants are not directly observable, and so one research problem is the construction of plausible proxy variables. A second problem arises in the fact that these determinants are almost certainly not exogenous, since the private information production is sensitive to the returns to this activity. A related problem is the identification of the exogenous determinants of the system which are amenable to adjustment by regulation of market structure or information disclosure.

A thorough treatment of these problems lies beyond the scope of this paper. Some preliminary conclusions may be drawn, however, from an examination
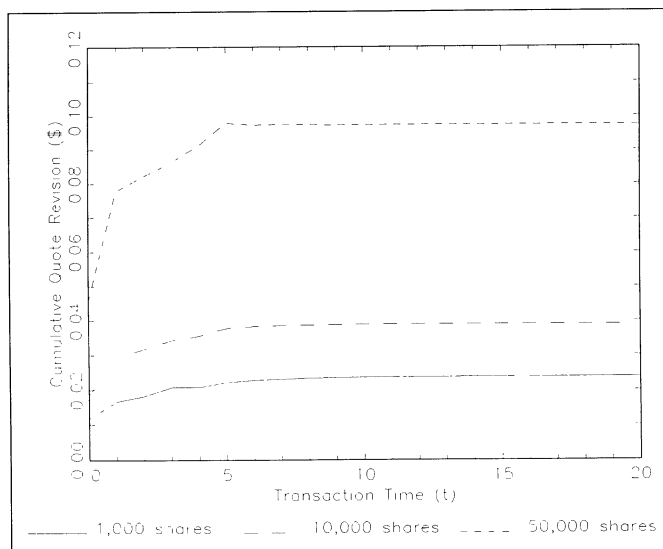
Figure 2. The quote revision process for Ames Department Stores implied by a quadratic model. The figure depicts the cumulative quote revision for a share of Ames Department Stores implied by the quadratic vector autoregression model presented in Table III, subsequent to initial buy orders of the indicated sizes. The index t refers to transactions.

of the relation between price impact measures and firm value. The interest in this link stems from the central role that market value seems to play as a determinant of return characteristics. Furthermore, the nature of the relation is not immediately obvious. The level of information production for large firms is almost certainly higher than that for small firms, but the shares of larger firms are also more widely held by liquidity traders, and it is the relative proportion of liquidity and informed traders that determines the price impact. It is also widely known that the immediate price impact of a trade on small issue is larger than the impact of a trade of the same size on a large issue. But much of the small firm impact might plausibly be transient, leaving a negligible persistent effect.

To assess this issue, four standardized price impact measures for each of the 80 firms in the sample were constructed in the following manner. I first computed the persistent price impact for a 1,000-share purchase implied by the quadratic VAR model discussed in the last section through the $20th$ step, that is, $\alpha_{m=20}(v_{2,0} = 1,000 \text{ shares})$. Since the units of this are dollars/(shares)^2, this measure is not immediately comparable across firms. A related measure is the average proportional price impact of a 1,000-share purchase, defined as

$$\delta_1 = \alpha_{m=20}(v_{2,0} = 1,000 \text{ shares})/\overline{P},$$

where $\overline{P}$ is the average price of the stock over the sample period. The units of δ_1 are (shares)$^{-1}$: the numerator is a dimensionless proportional increase.

In comparing this measure across firms with different market characteristics, it might be argued that a 1,000-share purchase is in some sense more significant for a small firm than a large one simply because the purchase is larger relative to the typical trade size: large firms tend to have larger trades. To partially alleviate this problem, I define two measures which take as the driving innovation trades lying at the 50th and 90th percentiles of the sample distribution function of the trades sizes for the particular firm:

$$\delta_2 = \alpha_{m=20}(v_{2,0} = \text{a 50th percentile purchase})/\bar{P},$$

$$\delta_3 = \alpha_{m=20}(v_{2,0} = \text{a 90th percentile purchase})/\bar{P}.$$

The three measures proposed thus far attempt to assess the absolute importance of the information revealed in the trades. It would also be useful to possess measures of the importance of trade impact relative to all public information. The motivation for this lies in the fact that private information is essentially prior knowledge of public information. Two firms with the same long-run return variance, for example, are in some sense equal in the intensity of information arrival. But they may still differ in the proportion of this long-run variance that is explained by trades (i.e., attributable to private information). While this proportion may be calculated from the VAR representation, the problems of specification and estimation are formidable (see Hasbrouck (1990)). As an alternative, the relative measure used here is the ratio of δ_3 to the standard deviation of the daily return for the stock:

$$\delta_4 = \delta_3/\sigma,$$

where σ is the standard deviation of daily returns computed over 1988.

Table IV reports summary statistics for the total sample and the market-value subsamples. The average proportional price change for a 1,000-share purchase is given by δ_1 at 0.00299 (0.299%), and in moving from lower to higher market-value subsamples, this measures does in fact decline. The trade distribution points (0.50 and 0.90) document the tendency of trade sizes to be larger for larger firms. It is therefore interesting that the impact measures δ_2 and δ_3, which are based on relative trade sizes, also exhibit the monotonic decline across subsamples. Finally, note that even though σ declines in moving from lower to higher value subsamples, the relative measure δ_4 still declines in the same fashion as the others. (F-tests rejected the null hypothesis of constant subsample means with p values < 0.001.) These results are quite preliminary but are consistent with information asymmetries being larger for firms with smaller market values.

VIII. Spreads, Quote Revisions, and Trading Activity

This paper has taken the position that as a measure of the average extent of information asymmetry in a security market, the average response of the price to trades is superior to use of the posted spread. Nevertheless, analysis of the dynamic properties of the spread may still yield valuable insights as to

<div align="center">

Table IV

Summary Statistics

</div>

The total sample comprises 80 NYSE/AMEX stocks; the subsamples are quartiles based on equity market value as of the end of 1988. Values in the table are sample means and (in brackets) standard deviations. *Average Share Price* is computed over the sample period (1989, first quarter); $\sigma(daily\ return)$ is computed over 1988; the 50th and 90th percentile trades are the .50 and .90 points of the sample distribution function of trade sizes for the individual firms. Quantity δ_1 is the average proportional impact on the quote midpoint of a 1,000 share purchase; δ_2 is the impact of a buy with size equal to the .50 point; δ_3 is the impact of a buy with size equal to the .90 point; $\delta_4 = \delta_3 / \sigma(daily\ return)$.

	Total Sample	Market Value Subsamples			
		1 (lowest)	2	3	4 (highest)
Average Share Price ($) ($\bar{P}$)	19.46 [15.44]	9.96 [5.27]	13.06 [7.42]	18.12 [9.75]	36.7 [18.99]
Market value ($ Million)	1,087. [2,214.]	25. [7.]	91.6 [28.]	374. [173.]	3,859. [3,081.]
σ(daily return)	.0236 [.0085]	.0288 [.0092]	.0261 [.0075]	.0225 [.0076]	.0168 [.0034]
50%ile trade (100 shares)	4.9 [1.7]	4.6 [1.2]	5.2 [1.5]	4.7 [1.8]	5.4 [2.1]
90%ile trade (100 shares)	32.4 [21.8]	19.1 [7.8]	30.2 [17.5]	36.9 [26.6]	43.5 [23.5]
δ_1	.00299 [.00234]	.00594 [.00201]	.00349 [.00115]	.00178 [.00102]	.00072 [.00029]
δ_2	.00256 [.00200]	.00486 [.00194]	.00311 [.00108]	.00164 [.00099]	.00064 [.00028]
δ_3	.00378 [.00285]	.00717 [.00266]	.00459 [.00141]	.00233 [.00123]	.00103 [.00047]
δ_4	.1540 [.0983]	.2575 [.0859]	.1897 [.0773]	.1066 [.0455]	.0620 [.0290]

how the market's perception of the information asymmetry changes over time as a result of trading activity. Easley and O'Hara (1987) describe a model in which the market-maker faces uncertainty both as to whether or not a trader is potentially informed and also as to whether a potentially-informed trader is in fact informed (that is, whether a private information signal has been observed). In their separating equilibrium, informed traders make large trades. Subsequent to a large trade, therefore, other participants (including the market-maker) increase their assessments of the probability that an information event has occurred. This may reasonably lead the market-maker to widen the spread.[10]

---

[10] In the Easley and O'Hara example, the spread for small trades is actually eliminated (as a consequence of the separation) while the spread for large trades increases.

This conjecture may be examined empirically using a VAR equation system for the spread, defined here as $s_t = q_t^a - q_t^b$. The dynamic properties of the spread are presumably dependent on the magnitudes of the trades rather than the (signed) directions. Accordingly, the VAR specification involves $\{s_t, |x_t^0|, |x_t|, |x_t^2|\}$, in which $|x_t^0|$ is a trade indicator variable ($= 1$ if a trade occurred), $|x_t|$ is the unsigned trade volume, and $|x_t^2|$ is the (trade volume)^2. Table V summarizes the spread estimation for ADD. The coefficients of lagged s_t are highly positive, indicating strong persistence. The significance of the individual trade coefficients is lower than in the r_t estimation (Table III). Nevertheless, a formal Wald test of the null hypothesis that all the trade variable coefficients are zero leads to rejection at a confidence level better than 0.001. For the entire sample, rejection at a 0.05 confidence level was implied in 77 of the 80 firms. Figure 3 graphs the response of the spread subsequent to purchases of 1,000, 10,000, and 50,000 shares. For the largest trade size, the spread temporarily widens, a finding consistent with the Easley and O'Hara hypothesis.

A dynamic response of the spread to trades suggests that the market's perception of information asymmetry is not uniform over time. As noted in the introduction, the extent of the asymmetry is a positive determinant of both the spread and the impact of a trade on price. This suggests that trades which occur when the spread is high should have a larger impact than those which occur when the spread is low. This conjecture may be examined in the VAR framework using trade variables which are functions of the spread. A particularly simple functional form is the product of the trade variable and the spread. Thus, corresponding to the original trade variables, I include three spread-trade variables: $s_t x_t^0$, $s_t x_t$, and $s_t x_t^2$. For a given trade size, of course, these variables are increasing in the spread. The VAR system then consists of linear regressions involving the set $\{r_t, x_t^0, x_t, x_t^2, s_t x_t^0, s_t x_t, s_t x_t^2\}$.

The r_t estimation for ADD is summarized in Table VI. The coefficient sums for the spread-trade variables have the same sign as the corresponding trade variables in the original quadratic estimation (Table III), but the statistical significance of these sums is substantially lower. Nevertheless, based on a formal Wald test, the null hypothesis that the coefficients of all the spread-trade variables are zero is rejected at a p-level < 0.001. In the full sample, rejection of this null at the 0.05 level occurred in 76 of the 80 cases. A representative impulse response analysis for ADD implied that the cumulative quote revision (through step $t = 20$) for a 10,000-share purchase was \$0.026 when the prevailing spread was 1/8, and \$0.058 when the spread was 1/4: trades that occur when the spread is wide have a relatively higher impact on the quote than those that occur when the spread is narrow.

IX. Conclusions and Extensions

The empirical analysis presented here is motivated by theoretical models which suggest that the response of security prices to trading activity is a

<div align="center">

Table V

Estimates of the Spread Equation for
Ames Department Stores
</div>

The set of variables in the model is $\{s_t, |x_t^0|, |x_t|, |x_t^2|\}$ where s_t is the prevailing spread (in dollars); $|x_t^0|$ is a trade indicator variable (1 if a trade has just occurred, 0 otherwise); $|x_t|$ is the trade volume; $|x_t^2| = |x_t|^2$; and t indexes transactions. The specification is:

$$s_t = \text{constant} + \sum_i a_i s_{t-i} + \sum_i b_i |x_{t-i}^0|$$

$$+ \sum_i c_i |x_{t-i}| + \sum_i d_i |x_{t-i}^2| + v_t.$$

The table contains summary statistics for each group of linear regression coefficients. In each group, "*lags*" specifies the lags used in the specification, "*p*" is the right-tail probability for a Wald test of the null hypothesis that all coefficients in the group are zero, "Σ" is the sum of the coefficients in the group, and a T-statistic for this sum is reported in parentheses. The estimates are based on all transactions for the first quarter of 1989.

RHS Variable Group		
constant		.065
		(6.77)
a_i	*lags*	1–5
	p	< .001
	Σ	.725
		(32.29)
b_i	*lags*	0–5
	p	.151
	Σ	.004
		(.98)
c_i	*lags*	0–5
	p	< .001
	Σ	.00020
		(5.89)
d_i	*lags*	0–5
	p	< .001
	Σ	-4.7×10^{-8}
		(−3.24)
$R^2 = .967$		

consequence of asymmetric information. The principal difficulty in making this concept operational is that the actual mechanism linking trades and quote responses is complicated by deviations from the idealized theoretical constructs including but not limited to liquidity effects, inventory control behavior, and price discreteness.
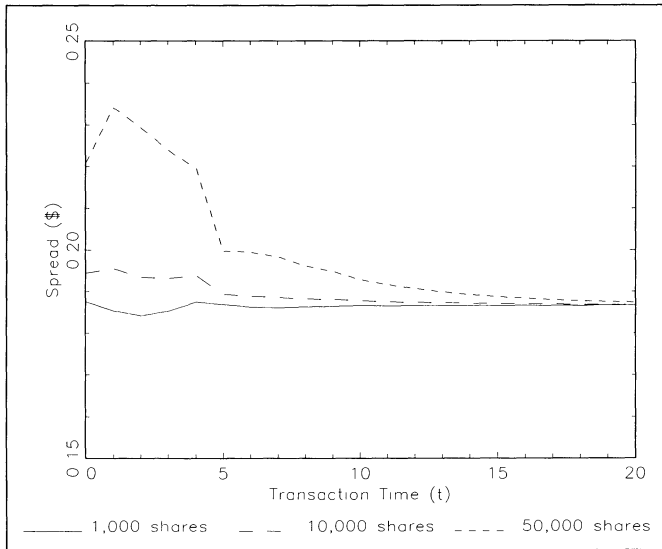
Figure 3. The spread adjustment process for Ames Department Stores. The figure depicts the spread quoted for a share of Ames Department Stores implied by the spread model presented in Table V, subsequent to initial buy orders of the indicated sizes. The index t refers to transactions.

To circumvent these difficulties, this paper suggests that quote revisions and trades be analyzed as a vector time series model. In such a model, the information content of a trade may be meaningfully measured as the persistent price impact of the unexpected component of the trade, i.e., the ultimate impact of the trade innovation. By focusing on the trade innovation rather than the trade itself, we may avoid misleading inferences due to inventory control or other transient liquidity effects. By considering the persistent impact of the innovation, we concentrate on the information ultimately impounded in the price after transient liquidity effects have died out.

This paper implements these principles by estimating general vector autoregressive models for trade and quote variables. The results are consistent with the following conclusions.

1. The full impact of a trade on the security price is not felt instantaneously but with a protracted lag.
2. As a function of trade innovation size, the ultimate impact of the innovation on the quote is nonlinear, positive, and increasing, but concave.
3. The order flow is affected by prior quote revisions: there is Granger-Sims causality running from quote revisions to trades.
4. Spread size exhibits a response to trading activity. Large trades in particular are associated with a widening of the bid-ask spread. This finding is consistent with a market-maker who infers from a large trade an increased likelihood that an information event has occurred.

<div align="center">

Table VI

Estimates of the Expanded intraday Quote Revision Equation
for Ames Department Stores

</div>

The set of variables in the model $\{r_t, x_t^0, x_t, x_t^2, s_t x_t^0, s_t x_t, s_t x_t^2\}$, where r_t is the price (quote midpoint) change; x_t^0 is a trade indicator variable; x_t is the signed trade volume; $x_t^2 = x_t^0 \mid x_t \mid^2$; s_t is the spread; and t indexes transactions. The estimated equation is:

$$r_t = \sum_i a_i r_{t-i} + \sum_i b_i x_{t-i}^0 + \sum_i c_i x_{t-i} + \sum_i d_i x_{t-i}^2 + \sum_i e_i s_{t-i} x_{t-i}^0$$
$$+ \sum_i f_i s_{t-i} x_{t-i} + \sum_i g_i s_{t-i} x_{t-i}^2 + v_t$$

The table contains summary statistics for each group of linear regression coefficients. In each group, "*lags*" specifies the lags used in the specifications, "*p*" is the right-tail probability for a Wald test of the null hypothesis that all coefficients in the group are zero, "Σ" is the sum of the coefficients in the group, and a *T*-statistic for this sum is reported in parentheses. Estimates are based on all transaction data for the first quarter of 1989.

	RHS Variable Group			*RHS* Variable Group	
a_i	*lags*	1–5			
	p	$< .001$			
	Σ	$-.208$			
		(-4.97)			
b_i	*lags*	0–5	e_i	*lags*	0–5
	p	$. < .001$		p	$< .001$
	Σ	$-.024$		Σ	.246
		(-4.10)			(7.08)
c_i	*lags*	0–5	f_i	*lags*	0–5
	p	.007		p	.04
	Σ	.0003		Σ	$-.0004$
		(2.77)			$(-.77)$
d_i	*lags*	0–5	g_i	*lags*	0–5
	p	.40		p	.50
	Σ	-1.8×10^{-7}		Σ	5.0×10^{-7}
		(-1.83)			(1.21)
$R^2 = .197$					

5. Trades which occur in the face of a relatively wide spread have a larger price impact than those which occur when the spreads are narrow.

6. Across firms, the price impact and (by implication) the extent of the information asymmetry appear more significant for firms with smaller market values.

The first point suggests that analyses which assume that the full impact of a trade on price is instantaneous are seriously incomplete. The second is a tentative characterization of the trade size-price impact relation. The third point is open to a number of interpretations, the most intriguing of which is the possibility of price experimentation by the dealer. The fourth and fifth points are together consistent with a plausible interpretation of the induced

effect of large trades. Large trades lead to large price impacts but also induce an increase in the spread consistent with a perceived increase in the extent of the information asymmetry, implying that subsequent trades will have larger price impacts.

There are a number of directions for further research along the lines presented here. These potential studies fall into two general groups: studies which use the trade impact measures defined here to study cross-security and cross-time differences, and those which refine the present method. In the first group lie more extensive comparative analyses of the trade impact across firms which have different ownership and trading patterns. Another important set of issues concerns intertemporal variation of the trade impact. Study of time-of-day and day-of-week patterns in the trade-impact measures can yield valuable insights into the timing of information production and trading activity.

The second class of studies concerns refinements. The linearized vector autoregression models studied here are naive in failing to capture some obvious features of the market process. Specific concerns include: What specifications appears to best characterize the price impact as a function of trade size? How should the model incorporate the distinctive features of market openings and closings? What are the effects of varying interevent times?

REFERENCES

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

Amihud, Yakov and Haim Mendelson, 1980, Dealership markets: Market making with uncertainty, *Journal of Financial Economics* 8, 31–54.

—— and Haim Mendelson, 1982, Asset price behavior in a dealership market, *Financial Analysts Journal* 38, 50–59.

Bagehot, Walter (pseudonym), 1971, The only game in town, *Financial Analysts Journal* 8, 31–53.

Beja, Avraham and M. Barry Goldman, 1980, On the behavior of prices in disequilibrium, *Journal of Finance* 35, 235–48.

Chiang, Raymond and P. C. Venkatesh, 1988, Insider holdings and perceptions of information asymmetry: A note, *Journal of Finance* 43, 1041–1048.

Copeland, Thomas and Dan Galai, 1983, Information effects on the bid-ask spread, *Journal of Finance* 38, 1457–69.

Easley, David and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics* 18, 69–90.

Foster, Douglas F. and S. Viswanathan, 1987, Intraday variations in volumes, spreads and variances: I. Theory, Working paper, Fuqua School of Business, Duke University.

—— and S. Viswanathan, 1988, Variations in volumes, spreads and variances, Working paper, Futures and Options Research Center, Duke University.

Garman, Mark B., 1976, Market microstructure, *Journal of Financial Economics* 3, 257–75.

Geweke, J., R. Meese, and W. Dent, 1983, Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence, *Journal of Econometrics*, 21, 161–194.

Glosten, Lawrence R., 1987, Components of the bid-ask spread and the statistical properties of transaction prices, *Journal of Finance* 42, 1293–1307.

——, 1989, The role of the monopolistic specialist, *Journal of Business* 62, 211–235.

—— and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a market-maker market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

—— and Lawrence E. Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123–142.

Goldman, M. Barry and Avraham Beja, 1979, Prices vs. equilibrium prices: Returns variances, serial correlation and the role of the specialist, *Journal of Finance* 34, 595–607.

Harris, Lawrence, 1990, Estimation of stock price variances and serial covariances from discrete observations, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Hasbrouck, Joel, 1988, Trades, quotes, inventories and information, *Journal of Financial Economics* 22 229–252.

——, 1990, The summary informativeness of stock trades: An econometric analysis, Working paper, Stern School, New York University.

—— and Thomas S. Y. Ho, 1987, Order arrival, quote behavior and the return generating process, *Journal of Finance* 42, 1035–1048.

—— and Robert A. Schwartz, 1988, An assessment of stock exchange and over-the-counter markets, *Journal of Portfolio Management* 14, 10–16.

Ho, Thomas S. Y. and Hans R. Stoll, 1980, On dealer markets under competition, *Journal of Finance* 35, 259–67.

—— and Hans R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics* 9, 47–73.

—— and Hans R. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance* 38, 1053–74.

Judge, George G., William E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee, 1985, *The Theory and Practice of Econometrics* (John Wiley, Somerset, NJ).

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–36.

Leach, Christopher and Ananth Madhavan, 1989, Price experimentation and market structure, Working paper, Wharton School, University of Pennsylvania.

McInish, Thomas H. and Robert A. Wood, 1988, An analysis of intraday patterns in bid/ask spreads for NYSE stocks, Working paper, Penn State University.

O'Hara, Maureen and George S. Oldfield, 1986, The microeconomics of market making, *Journal of Financial and Quantitative Analysis* 21, 361–76.

Stoll, Hans R., 1976, Dealer inventory behavior: An empirical investigation of NASDAQ/NMS stocks, *Journal of Financial and Quantitative Analysis*, 359–380.

——, 1989, Inferring the components of the bid-ask spread: Theory and empirical tests, *Journal of Finance* 44, 115–134.

Wood, Robert L., Thomas H. McInish, and J. Keith Ord, 1985, An investigation of transactions data for NYSE stocks, *Journal of Finance* 40, 723–738.