The Components of the Bid-Ask Spread: A General Approach

Author(s): Roger D. Huang and Hans R. Stoll

Source: *The Review of Financial Studies*, Winter, 1997, Vol. 10, No. 4 (Winter, 1997), pp. 995-1034

Published by: Oxford University Press. Sponsor: The Society for Financial Studies.

Stable URL: https://www.jstor.org/stable/2962337

The Components of the Bid-Ask Spread: A General Approach

Roger D. Huang
Hans R. Stoll
Vanderbilt University

A simple time-series market microstructure model is constructed within which existing models of spread components are reconciled. We show that existing models fail to decompose the spread into all its components. Two alternative extensions of the simple model are developed to identify all the components of the spread and to estimate the spread at which trades occur. The empirical results support the presence of a large order processing component and smaller, albeit significant, adverse selection and inventory components. The spread components differ significantly according to trade size and are also sensitive to assumptions about the relation between orders and trades.

The difference between the ask and the bid quotes — the spread — has long been of interest to traders, regulators, and researchers. While acknowledging that the bid-ask spread must cover the order processing costs incurred by the providers of market liquidity, researchers have focused on two additional costs of market making that must also be reflected in the spread.

We have benefited from the comments of seminar participants at Arizona State University, Louisiana State University, Rice University, University of California at Los Angeles, University of North Carolina at Chapel Hill, University of Southern California, Vanderbilt University, and the 1995 Asian Pacific Finance Association Conference. We are also grateful to Ravi Jagannathan (the editor) and two anonymous referees for their comments. This research was supported by the Dean's Fund for Research and by the Financial Markets Research Center at the Owen Graduate School of Management, Vanderbilt University. Address correspondence and send reprint requests to Roger D. Huang, Owen Graduate School of Management, Vanderbilt University, Nashville, TN 37203.

The Review of Financial Studies Winter 1997 Vol. 10, No. 4, pp. 995–1034

Amihud and Mendelson (1980), Demsetz (1968), Ho and Stoll (1981, 1983), and Stoll (1978) emphasize the inventory holding costs of liquidity suppliers. Copeland and Galai (1983), Easley and O'Hara (1987), and Glosten and Milgrom (1985) concentrate on the adverse selection costs faced by liquidity suppliers when some traders are informed.

Several statistical models empirically measure the components of the bid-ask spread. In one class of models pioneered by Roll (1984), inferences about the bid-ask spread are made from the serial covariance properties of observed transaction prices. Following Roll, other covariance spread models include Choi, Salandro, and Shastri (1988), George, Kaul, and Nimalendran (1991), and Stoll (1989). In another category of models, inferences about the spread are made on the basis of a trade indicator regression model. Glosten and Harris (1988) were the first to model the problem in this form, but they did not have the quote data to estimate the model directly. A recent article by Madhavan, Richardson, and Roomans (1996) also falls into this category. Other related articles include Huang and Stoll (1994), who show that short-run price changes of stocks can be predicted on the basis of microstructure factors and certain other variables, Lin, Sanger, and Booth (1995b), who estimate the effect of trade size on the adverse information component of the spread, and Hasbrouck (1988, 1991), who models the time series of quotes and trades in a vector autoregressive framework to make inferences about the sources of the spread.

Statistical models of spread components have been applied in a number of ways: to compare dealer and auction markets [Affleck-Graves, Hegde, and Miller (1994), Jones and Lipson (1995), Lin, Sanger, and Booth (1995a), Porter and Weaver (1995)], to analyze the source of short-run return reversals [Jegadeesh and Titman (1995)], to determine the sources of spread variations during the day [Madhavan, Richardson and Roomans (1996)], to test the importance of adverse selection for spreads of closed-end funds [Neal and Wheatley (1994)], and to assess the effect of takeover announcements on the spread components [Jennings (1994)]. Other applications, no doubt, will be found.

Most of the existing research provides neither a model nor empirical estimates of a three-way decomposition of the spread into order processing, inventory, and adverse information components. Further, much of the current research unknowingly uses closely related models to examine the issue. We show the underlying similarity of various models and we provide two approaches to a three-way decomposition of the spread.

This study's first objective is to construct and estimate a basic trade indicator model of spread components within which the various existing models may be reconciled. A distinguishing characteristic of trade indicator models is that they are driven solely by the direction of trade — whether incoming orders are purchases or sales. Covariance models also depend on the probabilities of changes in trade direction. We show that the existing trade indicator and covariance models fail to decompose the spread fully, for they typically lump order processing and inventory costs into one category even though these components are different.

The second objective is to provide a method for identifying the spread's three components — order processing, adverse information, and inventory holding cost. Inventory and adverse information components are difficult to distinguish because quotes react to trades in the same manner under both. We propose and test two extensions of the basic trade indicator models to separate the two effects. The first extension relies on the serial correlation in trade flows. Quote adjustments for inventory reasons tend to be reversed over time, while quote adjustments for adverse information are not. Trade prices also reverse (even if quotes do not), which is a measure of the order processing component. We use the behavior of quotes and trade prices after a trade to infer inventory and order processing effects that are distinct from adverse information effects. The second extension relies on the contemporaneous cross-correlation in trade flows across stocks. Because liquidity suppliers, such as market makers, hold portfolios of stocks, they adjust quotes in a stock in response to trades in other stocks in order to hedge inventory [Ho and Stoll (1983)]. We use the reaction to trades in other stocks to infer the inventory component as distinct from the adverse selection and order processing components. The empirical results yield separate inventory and adverse information components that are sensitive to clustering of transactions and to trade size as measured by share volume.

The basic and extended trade indicator models proposed and tested in this study have the advantage of simplicity. The essential features of trading are captured without complicated lag structures or other information.[1] Despite its simplicity, our approach is general enough to accommodate the many previous formulations while making no additional demands on the data. A second benefit is that the models can be implemented easily with a one-step regression procedure that provides added flexibility in addressing myriad statistical issues such as measurement errors, heteroskedasticity, and serial correlation.

[1] More involved econometric models of market mictorstucture require additional determinants. For example, Hasbrouck (1988) based inferences about the spread on longer lag structures. Huang and Stoll (1994) consider the simultaneous restrictions imposed on quotes and transaction prices by lagged variables such as prices of index futures. See also Hausman, Lo, and MacKinlay (1992).

A third benefit is that the trade indicator models provide a flexible framework for examining a variety of microstructure issues. One issue is the importance of trade size for the components of the spread. We adapt our trade indicator model to estimate the components of the bid-ask spread for three categories of trade size. We find that the components of the spread are a function of the trade size.

Another issue easily examined in our framework is time variation of spreads and spread components during the day. The trade indicator model can readily be modified to study this issue by using indicator variables for times of the day. Madhavan, Richardson, and Roomans (1996), in a model similar to ours, examine intraday variations in price volatility due to trading costs and public information shocks. They conclude that adverse information costs decline throughout the day and other components of the spread increase. However, they do not separate inventory and order processing components of the spread.

An issue that could also be examined within our framework is the observed asymmetry in the price effect of block trades. Holthausen, Leftwich, and Mayers (1987) and Kraus and Stoll (1972), for example, find that price behavior of block trades at the bid differ from those at the ask. In this article we focus on the spread midpoint, but the model can easily be modified to include indicator variables for the spread locations where a trade can occur. A covariance approach to estimating spread components, as in Stoll (1989), cannot be used to determine spread components for trades at the bid versus trades at the ask.

The remainder of the article is organized as follows. Section 1 constructs a basic trade indicator model and shows how one may derive from this model existing covariance models of the spread and existing trade indicator models. A variant of the basic model that incorporates different trade size categories is also presented. While the basic model (and the existing models implied by it) provides important insights into the sources of short-term price variability, we show that it is not rich enough to separately identify adverse information from inventory effects. Section 2 describes the dataset which consists of all trades and quotes for 20 large NYSE stocks in 1992. Section 3 describes the econometric methodology. In Section 4 the results of estimating the basic model are presented, including the effect of trade size. Section 5 introduces the first extended model in which the three components of the spread are decomposed on the basis of reversals in quotes. We also show how the components are affected by the observed sequence of trade sizes. Section 6 decomposes the spread on the basis of information on marketwide inventory pressures. Conclusions are in Section 7.

1. A Basic Model

In this section we develop a simple model of transaction prices, quotes, and the spread within which other models are reconciled. We adopt the convention that the time subscript (t) encompasses three separate and sequential events. The unobservable fundamental value of the stock in the absence of transaction costs, V_t, is determined just prior to the posting of the bid and ask quotes at time t. The quote midpoint, M_t, is calculated from the bid-ask quotes that prevail just before a transaction. We denote the price of the transaction at time t as P_t. Also define Q_t to be the buy-sell trade indicator variable for the transaction price, P_t . It equals +1 if the transaction is buyer initiated and occurs above the midpoint, −1 if the transaction is seller initiated and occurs below the midpoint, and 0 if the transaction occurs at the midpoint.

We model the unobservable V_t as follows:

$$V_t = V_{t-1} + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t, \tag{1}$$

where S is the constant spread, α is the percentage of the half-spread attributable to adverse selection, and ε_t is the serially uncorrelated public information shock. Equation (1) decomposes the change in V_t into two components. First, the change in V_t reflects the private information revealed by the last trade, $\alpha(S/2)Q_{t-1}$, as in Copeland and Galai (1983) and Glosten and Milgrom (1985). Second, the public information component is captured by ε_t.

While V_t is a hypothetical construct, we do observe the midpoint, M_t, of the bid-ask spread. According to inventory theories of the spread, liquidity suppliers adjust the quote midpoint relative to the fundamental value on the basis of accumulated inventory in order to induce inventory equilibrating trades [Ho and Stoll (1981) and Stoll (1978)]. Assuming that past trades are of a normal size of one, the midpoint is, under these models, related to the fundamental stock value according to

$$M_t = V_t + \beta \frac{S}{2} \sum_{i=1}^{t-1} Q_i, \tag{2}$$

where β is the proportion of the half-spread attributable to inventory holding costs, where $\sum_{i=1}^{t-1} Q_i$ is the cumulated inventory from the market open until time $t-1$, and Q_1 is the initial inventory for the day. In the absence of any inventory holding costs, there would be a one-to-one mapping between V_t and M_t. Because we assume that the spread is constant, Equation (2) is valid for ask or bid quotes as well as for the midpoint.

The first difference of Equation (2) combined with Equation (1) implies that quotes are adjusted to reflect the information revealed by the last trade and the inventory cost of the last trade:

$$\Delta M_t = (\alpha + \beta)\frac{S}{2}Q_{t-1} + \varepsilon_t, \tag{3}$$

where Δ is the first difference operator.

The final equation specifies the constant spread assumption:

$$P_t = M_t + \frac{S}{2}Q_t + \eta_t, \tag{4}$$

where the error term η_t captures the deviation of the observed half-spread, $P_t - M_t$, from the constant half-spread, $S/2$, and includes rounding errors associated with price discreteness.

The spread, S, is estimated from the data and we refer to it as the traded spread. It differs from the observed posted spread, S_t, in that it reflects trades inside the spread but outside the midpoint. Trades inside the spread and above the midpoint are coded as ask trades, and those inside the spread and below the midpoint are coded as bid trades. If trades occur between the midpoint and the quote, S is less than the posted spread, which is the case in the data we analyze. If trades occur only at the posted bid or the posted ask, S is the posted spread. The estimated S is greater than the observed effective spread defined as $|P_t - M_t|$ because midpoint trades coded as $Q_t = 0$ are ignored in the estimation.[2]

Combining Equations (3) and (4) yields the basic regression model

$$\Delta P_t = \frac{S}{2}(Q_t - Q_{t-1}) + \lambda\frac{S}{2}Q_{t-1} + e_t, \tag{5}$$

where $\lambda = \alpha + \beta$ and $e_t = \varepsilon_t + \Delta\eta_t$. Equation (5) is a nonlinear equation with within-equation constraints. The only determinant is an indicator of whether trades at t and $t-1$ occur at the ask, bid, or midpoint. This indicator variable model provides estimates of the traded spread, S, and the total adjustment of quotes to trades, $\lambda(S/2)$. On the basis of Equation (5) alone, we cannot separately identify the adverse selection (α) and the inventory holding (β) components of the half-spread. However, we can estimate the portion of the half-spread not due to adverse information or inventory as $1 - \lambda$. This

[2] By contrast, estimates of S derived from the serial covariance of trade prices, as in Roll (1984), are influenced by the number of trades at the midpoint. Harris (1990) shows using simulations that the Roll (1984) estimator can be seriously biased. For estimates of the effective spread, $|P_t - M_t|$, see Huang and Stoll (1996a, 1996b).

remaining portion is an estimate of order processing costs, such as labor and equipment costs.

1.1 Comparison with covariance spread models

Serial covariance in the trade flow, Q_t, plays an important role in earlier covariance models of the spread. Specifically the covariance is a function of the probability of a trade flow reversal, π, or a continuation, $1 - \pi$. A reversal is said to occur if after a trade at the bid (ask), the next trade is at the ask (bid). Equation (5) accounts for reversals but does not assume a specific probability of reversal. Instead it relies on the direction of individual trades and the magnitude of price and quote changes.

Roll (1984) proposes a model of the bid-ask spread that relies exclusively on transaction price data and assumes $\pi = 1/2$:

$$S = 2\sqrt{-\operatorname{cov}(\Delta P_t, \Delta P_{t-1})}, \tag{6}$$

His model assumes the existence of only the order processing cost, for the stock's value is independent of the trade flow and there are no inventory adjustments. To derive Roll's model from Equation (5), set $\alpha = \beta = 0$ in Equation (5) to obtain

$$\Delta P_t = \frac{S}{2}\Delta Q_t + e_t, \tag{7}$$

Calculate the serial covariances of both sides of Equation (7), using the fact that $\operatorname{cov}(\Delta Q_t, \Delta Q_{t-1})$ equals -1 when $\pi = 1/2$, to produce Roll's estimator of Equation (6).[3]

Choi, Salandro, and Shastri (CSS) (1988) extend Roll's (1984) model to permit serial dependence in transaction type. Serial covariance in trade flows can occur if large orders are broken up or if "stale" limit orders are in the book. When π is not constrained to be one-half, Equation (7) implies the CSS's estimator

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = -\pi^2 S^2, \tag{8}$$

which is the Roll model if $\pi = 1/2$. It is important to emphasize that in the CSS's model, the deviation of π from one-half is not due to inventory adjustment behavior of liquidity suppliers.

More generally, the probability of a trade flow reversal (continuation) is greater (less) than 0.5 when liquidity suppliers adjust bid-ask spreads to equilibrate inventory. Stoll (1989) models this aspect of market making and allows for the presence of adverse selection

[3] The covariance in trade changes is $\operatorname{cov}(\Delta Q_t, \Delta Q_{t-1}) = -4\pi^2$, which is -1 when $\pi = 1/2$. The covariance in trades is $\operatorname{cov}(Q_t, Q_{t-1}) = (1 - 2\pi)$, which is zero when $\pi = 1/2$.

costs, inventory holding costs, and order processing costs. Buy and sell transactions are no longer serially independent and their serial covariance provides information on the components of the spread. The model consists of two equations:

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = S^2[\delta^2(1-2\pi) - \pi^2(1-2\delta)], \tag{9}$$

$$\operatorname{cov}(\Delta M_t, \Delta M_{t-1}) = \delta^2 S^2(1-2\pi), \tag{10}$$

where $\delta = \frac{\Delta P_{t+1}|P_t=A_t, P_{t+1}=A_{t+1}}{S} = \frac{-\Delta P_{t+1}|P_t=B_t, P_{t+1}=B_{t+1}}{S}$ is the magnitude of a price continuation as a percentage of the spread.[4] The two equations are used to estimate the two unknowns δ and π. Stoll's covariance estimators, Equations (9) and (10), result directly from the covariances of Equations (5) and (3), respectively, when one uses the transformation $\delta = \lambda/2$. Stoll shows that the expected revenue earned by a supplier of immediacy on a round-trip trade is $2(\pi - \delta)S$. This amount is compensation for order processing and inventory costs. The remainder of the spread, $[1 - 2(\pi - \delta)]S$, is the portion of the spread not earned by the supplier of immediacy, and this amount reflects the adverse information component of the spread.[5]

George, Kaul, and Nimalendran (GKN) (1991) ignore the inventory component of the bid-ask spread and assume no serial dependence in transaction type so that $\pi = 1/2$.[6] Their model in our notation is Equation (5) with $\beta = 0$. Under GKN's assumptions, Equation (5) implies the GKN's covariance estimator:

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = -(1-\alpha)\frac{S^2}{4}, \tag{11}$$

where $1-\alpha$ is the order processing component of the bid-ask spread. Equation (11) is observationally equivalent to Stoll's, Equations (9) and (10), under GKN's assumptions that $\pi = 1/2$, $\beta = 0$.

[4] Under the assumption of a constant spread, writing the covariance in terms of quote midpoints as in Equation (10) is equivalent to writing it in terms of the bid or ask as Stoll does.

[5] Stoll further decomposes the revenue component, $2(\pi - \delta)S$ into order processing and inventory components by arguing that $\pi = 0.5$ and $\delta = 0.0$ for order processing and $\pi > 0.5$ and $\delta = 0.5$ for inventory holding, but this decomposition is ad hoc.

[6] George, Kaul, and Nimalendran (1991) use daily data and consider changing expectations in their model. Their formulation of time-varying expectations may be incorporated into our setup by expressing the trade price and the fundamental stock value in natural logarithms and by including a linearly additive term for an expected return over the period $t-1$ to t in Equation (1). Since our analysis focuses on microstructure effects at the level of transactions data where changing expectations are likely to be unimportant, we ignore this complication in the article.

1.2 Comparison with trade indicator spread models

The basic model, Equation (5), also generalizes some existing trade indicator spread models. We provide two examples.

Glosten and Harris (GH) (1988) develop a trade indicator variable approach to model the components of the bid-ask spread. Their basic model under our timing convention is[7]

$$\Delta P_t = Z_t Q_t + C_t \Delta Q_t + e_t, \tag{12}$$

where Z_t is the adverse selection spread component, C_t is the transitory spread component reflecting order processing and inventory costs, and e_t is defined as in Equation (5). GH use Fitch transaction data, which contains transaction prices and volumes but no information on quotes. Consequently, they are unable to observe Q_t and cannot estimate Equation (12) directly. Instead, they estimate Z_t and C_t by conditioning them on the observed volume at time t. We are able to observe Q_t and can estimate Equation (12) directly. Under GH's assumption that there are two components to the spread and making our assumption of a constant spread, GH's adverse selection component is $Z_t = \alpha(S/2)$ and their order processing component is $C_t = (1-\alpha)(S/2)$. They assume that $\beta = 0$. Making these substitutions in Equation (12) and rearranging terms yields a restricted version of Equation (5). They do not provide estimates of the spread. We detail the derivation of the GH model in Appendix A.

Madhavan, Richardson, and Roomans (MRR) (1996) also provide a trade indicator spread model along the lines of GH. Using our timing convention and assuming serially uncorrelated trade flows, their model is[8]

$$\Delta P_t = (\phi + \theta) Q_t - \phi Q_{t-1} + e_t, \tag{13}$$

where θ is the adverse selection component, ϕ is the order processing and inventory component, and e_t is as defined in Equation (5). Upon rearranging, Equation (13) becomes

$$\Delta P_t = \theta Q_t + \phi \Delta Q_t + e_t, \tag{14}$$

which has the same form as the GH model [Equation (12)]. As in GH, MRR assume that $\beta = 0$. As we do later in this article, with respect to our basic model [Equation (5)], MRR extend their model to allow the surprise in trade flow to affect estimated values.[9]

---

[7] Equation 2 in Glosten and Harris (1988, p. 128).

[8] Equation 3 in Madhavan, Richardson, and Roomans (1996, p. 7).

[9] MRR also provide estimates of the unconditional probability of a trade that occurs within the quoted spreads.

1.3 Trade size

Equation (5) generalizes existing spread models as described in Sections 1.1 and 1.2. We show below in Section 3 that the regression setup implied by Equation (5) makes it easier to account for a variety of econometric issues. Equation (5) can also easily be generalized to numerous new applications merely by introducing indicator variables that are 1 under certain conditions and 0 otherwise. For example, the model can be used to estimate S and λ for different times of the trading day by the introduction of time indicator variables. This is the principal objective of Madhavan, Richardson, and Roomans (1996). It can also be used to estimate S and λ at different spread locations to determine issues such as whether spread components for trades at the ask differ from those for trades at the bid.

In this article we generalize Equation (5) to allow different coefficient estimates by trade size category. We choose three trade size categories, although any number of categories is possible. The model is then developed by writing Equations (1) and (2) with indicator variables for each size category as shown in detail in Appendix B. The result is

$$\begin{aligned}\Delta P_t = {} & \frac{S^s}{2}D_t^s + (\lambda^s - 1)\frac{S^s}{2}D_{t-1}^s + \frac{S^m}{2}D_t^m + (\lambda^m - 1)\frac{S^m}{2}D_{t-1}^m \\ & + \frac{S^l}{2}D_t^l + (\lambda^l - 1)\frac{S^l}{2}D_{t-1}^l + e_t, \end{aligned} \tag{15}$$

where

$$\begin{aligned} D_t^s &= Q_t && \text{if share volume at } t \le 1000 \text{ shares} \\ &= 0 && \text{otherwise} \\ D_t^m &= Q_t && 1000 \text{ shares} < \text{if share volume at } t < 10,000 \text{ shares} \\ &= 0 && \text{otherwise} \\ D_t^l &= Q_t && \text{if share volume at } t \ge 10,000 \text{ shares} \\ &= 0 && \text{otherwise.} \end{aligned}$$

Equation (15) allows the coefficient estimates for small (s), medium (m), and large (l) trades to differ. The estimate of λ depends on the trade size at time $t-1$, which determines the quote reaction, and the estimate of S depends primarily on the trade size at t, which determines where the trade is relative to the midpoint. The parameter estimates do not depend on the sequence of trades. In extensions of the basic model provided later, the sequence of trades does matter.

1.4 Summary

We have integrated existing spread models driven solely by a trade indicator variable. Most models simply seek to identify the adverse selection component and assume the remainder of the spread reflects inventory and order processing. In fact, estimates of adverse information probably include inventory effects as well since existing procedures cannot distinguish the two. Our basic model [Equation (5)], which we have used as a framework to integrate existing work, also cannot make that distinction. It can only identify the order processing component and the sum of inventory and adverse information. We now describe the data and the econometric procedures, and we estimate Equation (5) and the generalization [Equation (15)] that accounts for trade size categories. In later sections of the article we propose and test two alternative extensions that provide a full three-way decomposition of the spread.

2. Data Description

Trade and quote data are taken from the data files compiled by the Institute for the Study of Security Markets (ISSM). We use a ready-made sample of the largest and the most actively traded stocks by examining the 20 stocks in the Major Market Index for all trading days in the calendar year 1992. The securities are listed in Appendix C.

To ensure the integrity of the dataset, the analysis is confined to transactions coded as regular trades and quotes that are best bid or offer (BBO) eligible. All prices and quotes must be divisible by 16, be positive, and asks must exceed bids. We restrict the dataset to NYSE trades and quotes. Each trade is paired with the last quote posted at least 5 seconds earlier but within the same trading day.

The NYSE often opens with a call market and operates as a continuous market the remainder of the trading day. To avoid mixing different trading structures, we ignore overnight price and quote changes. We also exclude the first transaction price of the day if it is not preceded by a quote, which will be the case if the opening is a call auction based on accumulated overnight orders.

Table 1 presents the summary statistics for the 20 firms in the sample. The number of observations range from a low of 15,682 (62 trades per day) for USX (X) to a high of 181,663 (715 trades a day) for Philip Morris (MO). The next lowest number of observations belong to 3M (MMM) which averages about 165 trades a day. Given the wide disparity in trading activity in USX relative to the other firms in the sample, we exclude it from further analysis.

Table 1 also contains statistics on share price and posted spread. These statistics are provided for all trades, for trade sizes less than or equal to 1000 shares (small), for trade sizes between 1000 and 10,000 shares (medium), and for trade sizes greater than or equal to 10,000 shares (large). The share price varies considerably across stocks and

Table 1
Descriptive statistics

Company	# of Obs.	Trade Size	Mean Price	Std. Dev. Price	Mean Spread	Std. Dev. Spread	Company	# of Obs.	Trade Size	Mean Price	Std. Dev. Price	Mean Spread	Std. Dev. Spread
AXP	69271	All	22.403	1.171	0.146	0.047	KO	124195	All	55.525	18.453	0.159	0.056
	40077	Small	22.389	1.176	0.142	0.043		86571	Small	55.975	18.556	0.156	0.054
	22769	Medium	22.407	1.160	0.150	0.050		31780	Medium	55.308	18.418	0.167	0.059
	6425	Large	22.473	1.175	0.160	0.056		5844	Large	50.041	16.069	0.170	0.061
CHV	48489	All	68.305	3.711	0.190	0.084	MMM	41893	All	96.706	4.611	0.220	0.102
	30691	Small	68.195	3.729	0.186	0.080		28733	Small	96.721	4.596	0.216	0.098
	16056	Medium	68.518	3.669	0.195	0.089		12306	Medium	96.725	4.621	0.229	0.109
	1742	Large	68.294	3.699	0.204	0.090		854	Large	95.946	4.898	0.240	0.111
DD	72748	All	48.924	2.486	0.165	0.063	MO	181663	All	78.075	3.101	0.161	0.068
	42520	Small	48.916	2.478	0.165	0.062		122214	Small	78.115	3.087	0.157	0.062
	26246	Medium	48.920	2.485	0.165	0.065		49939	Medium	78.016	3.132	0.169	0.078
	3982	Large	49.039	2.582	0.162	0.061		9510	Large	77.861	3.098	0.178	0.082
DOW	63087	All	56.837	2.803	0.175	0.065	MOB	58654	All	63.032	2.360	0.184	0.068
	44064	Small	56.814	2.794	0.176	0.066		37481	Small	62.938	2.360	0.177	0.066
	16866	Medium	56.909	2.830	0.171	0.064		18275	Medium	63.182	2.358	0.195	0.071
	2157	Large	56.740	2.778	0.176	0.068		2898	Large	63.296	2.304	0.208	0.069
EK	72040	All	42.770	3.143	0.163	0.062	MRK	146800	All	85.259	50.528	0.207	0.098
	46644	Small	42.662	3.132	0.160	0.059		96392	Small	89.139	51.384	0.207	0.096
	20908	Medium	42.961	3.173	0.167	0.065		44535	Medium	80.479	49.031	0.210	0.101
	4488	Large	43.000	3.062	0.174	0.074		5873	Large	57.835	32.870	0.184	0.088

Table 1
(continued)

Company	# of Obs.	Trade Size	Mean Price	Std. Dev. Price	Mean Spread	Std. Dev. Spread	Company	# of Obs.	Trade Size	Mean Price	Std. Dev. Price	Mean Spread	Std. Dev. Spread
GE	123727	All	77.720	2.771	0.162	0.059	PG	71084	All	72.596	24.944	0.196	0.099
	83392	Small	77.740	2.818	0.160	0.057		48923	Small	73.179	24.974	0.193	0.095
	36530	Medium	77.683	2.676	0.167	0.061		19978	Medium	72.098	24.907	0.205	0.107
	3805	Large	77.632	2.619	0.170	0.065		2183	Large	64.100	22.905	0.200	0.103
GM	104527	All	36.077	4.121	0.165	0.059	S	56762	All	42.555	2.287	0.168	0.061
	56168	Small	35.996	4.094	0.157	0.055		32000	Small	42.541	2.272	0.163	0.058
	37213	Medium	36.193	4.144	0.174	0.062		20546	Medium	42.556	2.311	0.174	0.063
	11146	Large	36.102	4.169	0.179	0.063		4216	Large	42.652	2.282	0.181	0.065
IBM	145157	All	79.809	14.174	0.159	0.057	T	145476	All	42.590	3.421	0.147	0.048
	77141	Small	79.455	13.865	0.163	0.058		107170	Small	42.600	3.450	0.143	0.044
	59680	Medium	80.331	14.422	0.155	0.055		29920	Medium	42.564	3.331	0.157	0.055
	8336	Large	79.359	15.070	0.152	0.056		8386	Large	42.550	3.363	0.160	0.057
IP	44418	All	68.374	4.802	0.212	0.088	X	15682	All	26.482	1.870	0.176	0.062
	25843	Small	68.287	4.754	0.207	0.087		9668	Small	26.447	1.859	0.173	0.061
	16531	Medium	68.482	4.867	0.218	0.090		4870	Medium	26.582	1.896	0.178	0.064
	2044	Large	68.603	4.843	0.217	0.086		1144	Large	26.358	1.829	0.183	0.064
JNJ	129636	All	72.069	25.881	0.168	0.061	XON	73223	All	60.227	2.601	0.155	0.054
	95301	Small	72.716	25.974	0.164	0.059		43145	Small	60.170	2.623	0.148	0.049
	30992	Medium	70.955	25.694	0.178	0.063		25271	Medium	60.314	2.570	0.163	0.058
	3343	Large	63.979	23.102	0.187	0.068		4807	Large	60.273	2.553	0.173	0.064

The table presents descriptive statistics for share price and quoted spread of Major Market Index securities by trade size for the sample period 1992. A small trade size has 1000 shares or less, a medium trade size has greater than 1000 but less than 10,000 shares, and a large trade size has 10,000 or more shares.

within the year for certain stocks. The mean posted spread, which is a trade-weighted mean, always exceeds 12.5 cents but is generally less than 20 cents. The mean posted spread also tends to increase with trade size; IBM and MRK being notable exceptions.

3. Estimation Procedure

Equation (5) may be estimated by procedures that impose strong distributional assumptions such as maximum-likelihood (ML) or least-squares (LS) methods. For example, the ML approach taken by Glosten and Harris (1988) illustrates the practical difficulties of using an ML technique when the model is predicated on the specification of price discreteness. We opt for a generalized method of moments (GMM) procedure which imposes very weak distributional assumptions. This is especially important since the error term, e_t, includes rounding errors. The GMM procedure also easily accounts for the presence of conditional heteroskedasticity of an unknown form.

Define $f(x_t, \omega)$ to be a vector function such that for estimating the basic model [Equation (5)], it is

$$f(x_t, \omega) = \begin{bmatrix} e_t Q_t \\ e_t Q_{t-1} \end{bmatrix} \tag{16}$$

where $\omega = (S\lambda)'$ is the vector of parameters of interest, and for estimating the basic model with size categories [Equation (15)], it is

$$f(x_t, \omega) = [e_t D_t^s \quad e_t D_{t-1}^s \quad e_t D_t^m \quad e_t D_{t-1}^m \quad e_t D_t^l \quad e_t D_{t-1}^l]' \tag{17}$$

where $\omega = (S^s \; \lambda^s \; S^m \; \lambda^m \; S^l \; \lambda^l)'$. The basic models imply the orthogonality conditions $E[f(x_t, \omega)] = 0$. Under the GMM procedure, the parameter estimates chosen are those that minimize the criterion function

$$J_T(\omega) = g_T(\omega)' S_T g_T(\omega), \tag{18}$$

where $g_T(\omega)$ is the sample mean of $f(x_t, \omega)$, and S_T is a sample symmetric weighting matrix. Hansen (1982) proves that, under weak regularity conditions, the GMM estimator $\hat{\omega}_T$ is consistent and

$$\sqrt{T}(\hat{\omega}_T - \omega_0) \rightarrow N(0, \Omega) \tag{19}$$

where

$$\Omega = (D_0' S_0^{-1} D_0)^{-1}$$
$$D_0 = E\left[\frac{\partial f(x_t, \omega)}{\partial \omega}\right]$$
$$S_0 = E[f(x_t, \omega) f(x_t, \omega)'].$$

The basic models [Equations (5) and (15)] are exactly identified.

We also test overidentified models where the number of orthogonality conditions exceed the number of parameters that need to be estimated. An attractive feature of the GMM procedure is that it provides a test for overidentifying restrictions. Specifically, Hansen proves that T times the minimized value of Equation (18) is asymptotically distributed as chi-square, with the number of degrees of freedom equal to the number of orthogonality conditions minus the number of estimated parameters.

It is worth noting the differences between the proposed procedure and Stoll's (1989) estimation of his covariance model. Our procedure yields estimates of the spreads whereas Stoll relies on posted spreads. In addition, we avoid the criticism raised by George, Kaul, and Nimalendran (1991) that Stoll's estimates may be biased since they are nonlinear transformations of the linear parameters obtained from regressing covariances of price changes and quote revisions on mean spreads. Our GMM estimation procedure provides consistent estimates of the nonlinear parameters directly. Finally, the GMM procedure easily accommodates conditional heteroskedasticity of an unknown form and serial correlation in the residuals.

4. Two-Way Decomposition of the Spread

Table 2 presents the GMM estimates of our indicator variable model [Equation (5)]. Although it is not possible to separate out the adverse selection and inventory holding costs, the model separates order processing costs, $1 - \lambda$, from other sources of the spread, and it provide estimates of the traded spread, S. Moreover, estimation of Equation (5) makes no assumption about the conditional probability of trades, π.

The estimates of the traded spread given in Table 2 range from a low of 9.9 cents for IBM to a high of 13.5 cents for Procter and Gamble. A comparison to the average posted spread in Table 1 shows that the estimated traded spread is less for each stock than the posted spread, as expected.

The proportion of the traded spread that is due to adverse information and inventory effects, λ, ranges from a low of 1.9% of the traded spread for ATT to a high of 22.3% of the traded spread for 3M. The remaining part of the traded spread, 98.1% and 77.7%, respectively, is the order processing component. The order processing component averages 88.6% across all stocks. Given the presumption in numerous models that adverse information is a large component of the spread, the relatively small fraction of the spread estimated for both adverse information and inventory is surprising. The estimates are in line with those of George, Kaul, and Nimalendran (1991). They are smaller than those of Lin, Sanger, and Booth (1995b), and Stoll (1989). The

Table 2
Traded spread and order processing component

Company	Traded Spread, S		Adverse Selection and Inventory Holding, λ	
	Coefficient	Standard Error	Coefficient	Standard Error
AXP	0.1178	0.0002	0.0272	0.0016
CHV	0.1177	0.0006	0.1546	0.0039
DD	0.1254	0.0003	0.1663	0.0026
DOW	0.1348	0.0004	0.1992	0.0031
EK	0.1252	0.0003	0.0782	0.0020
GE	0.1167	0.0003	0.1060	0.0016
GM	0.1174	0.0002	0.0402	0.0014
IBM	0.0993	0.0003	0.1562	0.0019
IP	0.1350	0.0008	0.2127	0.0050
JNJ	0.1236	0.0003	0.0892	0.0019
KO	0.1202	0.0002	0.0604	0.0014
MMM	0.1239	0.0009	0.2229	0.0051
MO	0.1237	0.0002	0.0560	0.0015
MOB	0.1298	0.0004	0.1240	0.0033
MRK	0.1271	0.0004	0.1224	0.0018
PG	0.1352	0.0005	0.1689	0.0031
S	0.1160	0.0004	0.0916	0.0025
T	0.1214	0.0001	0.0186	0.0008
XON	0.1111	0.0003	0.0613	0.0021
AVG.	0.1222	0.0004	0.1135	0.0024

The table presents the results of estimating Equation (5). The estimated dollar traded spread and the proportion of traded spread due to adverse selection and inventory holding cost are shown. The order processing proportion is 1 minus the proportion due to adverse selection and inventory holding cost. The last row reports the average statistics for all stocks.

low adverse information component is also consistent with the result of Easley et al. (1996) that the risk of information-based trading is lower for active securities than for infrequently traded securities. This is because the presence of relatively more uninformed traders in an active stock reduces the probability that a market maker would end up trading with an informed trader.

The source of the estimate of $\alpha + \beta$ is the change in quotes in response to trades as specified in Equation (3). If many trades occur at the same quotes, the estimated adverse information and inventory effects will be small. The low estimates of $\alpha+\beta$ may be partly spurious if trades bunch at the bid or offer because large trades are broken up or because buying or selling programs cause several transactions to be at an unchanged bid or ask. In the next section, in addition to modifying our basic model to provide a decomposition of adverse information and inventory effects, we also provide estimates with and without trade bunching.

Before decomposing the adverse selection and the inventory holding costs in the next section, we consider the basic model with size categories [Equation (15)]. The results of the estimations are presented in Table 3. The differences in traded spreads between the small and medium-size trades are generally economically insignificant, but large trades experienced traded spreads that are almost 1.5 cents higher on average. The variation in the λ component across trade sizes is much more dramatic. Adverse selection and inventory holding components account for 3.3% of the traded spread on average for small trades. This increases to 21.7% for medium-size trades and further doubles to almost 43% for large trades. The small λ coefficients reported in Table 2 suggest that the estimates are heavily influenced by the more frequent occurrences of small trades.

To examine formally the variation in the estimates across trade size categories, we consider two constraints that impose overidentifying restrictions on Equation (15). The first constraint (Constraint 1) requires the traded spreads but not the order processing costs to be the same across size categories:

$$\Delta P_t = \frac{S}{2}(D_t^s + D_t^m + D_t^l) + (\lambda^s - 1)\frac{S}{2}D_{t-1}^s + (\lambda^m - 1)\frac{S}{2}D_{t-1}^m + (\lambda^l - 1)\frac{S}{2}D_{t-1}^l + e_t. \tag{20}$$

The second constraint (Constraint 2) requires both the traded spread and the order processing costs to be the same across size categories, which is the basic model [Equation (5)].

The results of overidentifying tests are presented in panel A of Table 4. The chi-square statistics reject both Constraints 1 and 2 at the usual significance levels. However, the magnitude of the chi-square statistics are much bigger for Constraint 2 than for Constraint 1, reflecting the small differences in traded spreads reported in Table 3. The results highlight the importance of considering the composition of the spread by trade size. Panel B of Table 4 presents the average across companies of the constrained parameter estimates for the two constraints. Under Constraint 1, average estimates of λ differ substantially across trade size categories and are comparable to average estimates in Table 3. Under Constraint 2, λ is 8.4%, reflecting the dominance of small trades in the sample. This estimate of adverse selection and inventory holding costs is somewhat smaller than the average estimate (11.4%) reported for the basic model without size categories in Table 2.

5. Three-Way Decomposition of the Spread Based on Induced Serial Correlation in Trade Flows

To distinguish the adverse selection (α) and inventory (β) components of the traded spread, we first make use of the fact that, under

Table 3
Traded spread and order processing component by trade size

		Traded Spread, S			Adverse Selection and Inventory Holding, λ		
Company	Estimate	Small	Medium	Large	Small	Medium	Large
AXP	Coeff.	0.1194	0.1138	0.1141	−0.0133	0.0368	0.2524
	Std. Error	0.0002	0.0004	0.0009	0.0022	0.0037	0.0102
CHV	Coeff.	0.1148	0.1192	0.1526	0.0372	0.3068	0.5682
	Std. Error	0.0006	0.0010	0.0041	0.0049	0.0086	0.0229
DD	Coeff.	0.1261	0.1207	0.1297	0.0620	0.2736	0.4471
	Std. Error	0.0004	0.0006	0.0015	0.0035	0.0051	0.0135
DOW	Coeff.	0.1329	0.1358	0.1511	0.0727	0.4238	0.6366
	Std. Error	0.0004	0.0008	0.0029	0.0035	0.0068	0.0194
EK	Coeff.	0.1245	0.1237	0.1265	0.0075	0.1524	0.3936
	Std. Error	0.0003	0.0005	0.0013	0.0023	0.0048	0.0124
GE	Coeff.	0.1176	0.1102	0.1301	0.0407	0.2273	0.4002
	Std. Error	0.0003	0.0005	0.0017	0.0020	0.0042	0.0131
GM	Coeff.	0.1177	0.1153	0.1192	−0.0189	0.0748	0.2530
	Std. Error	0.0003	0.0004	0.0008	0.0019	0.0035	0.0076
IBM	Coeff.	0.0995	0.0975	0.1077	0.0185	0.2646	0.4590
	Std. Error	0.0004	0.0004	0.0012	0.0029	0.0039	0.0104
IP	Coeff.	0.1337	0.1342	0.1602	0.1320	0.2916	0.5648
	Std. Error	0.0009	0.0013	0.0042	0.0070	0.0095	0.0239
JNJ	Coeff.	0.1236	0.1216	0.1377	0.0475	0.1948	0.3888
	Std. Error	0.0003	0.0006	0.0024	0.0021	0.0055	0.0182
KO	Coeff.	0.1199	0.1186	0.1256	0.0035	0.1631	0.3554
	Std. Error	0.0002	0.0005	0.0012	0.0016	0.0041	0.0110
MMM	Coeff.	0.1181	0.1334	0.1886	0.1065	0.4133	0.5953
	Std. Error	0.0009	0.0015	0.0083	0.0066	0.0107	0.0349
MO	Coeff.	0.1252	0.1176	0.1205	0.0149	0.1210	0.2698
	Std. Error	0.0002	0.0004	0.0011	0.0018	0.0039	0.0105
MOB	Coeff.	0.1288	0.1291	0.1454	0.0347	0.2621	0.5398
	Std. Error	0.0005	0.0009	0.0028	0.0039	0.0077	0.0206
MRK	Coeff.	0.1236	0.1317	0.1396	0.0408	0.2388	0.3712
	Std. Error	0.0004	0.0006	0.0022	0.0024	0.0042	0.0117
PG	Coeff.	0.1342	0.1346	0.1556	0.1013	0.2773	0.5604
	Std. Error	0.0005	0.0009	0.0041	0.0039	0.0071	0.0218
S	Coeff.	0.1132	0.1160	0.1319	−0.0247	0.1783	0.4629
	Std. Error	0.0004	0.0006	0.0017	0.0032	0.0053	0.0136
T	Coeff.	0.1217	0.1187	0.1186	−0.0104	0.0633	0.2398
	Std. Error	0.0001	0.0003	0.0005	0.0008	0.0027	0.0072
XON	Coeff.	0.1100	0.1088	0.1229	−0.0348	0.1664	0.3906
	Std. Error	0.0003	0.0005	0.0014	0.0029	0.0050	0.0128
AVG.	Coeff.	0.1213	0.1211	0.1357	0.0325	0.2174	0.4289
	Std. Error	0.0004	0.0007	0.0023	0.0031	0.0056	0.0156

The table presents the results of estimating Equation (15). The estimated dollar traded spread and the proportion of traded spread due to adverse selection and inventory holding cost by trade size are shown. The order processing proportion is 1 minus the proportion due to adverse selection and inventory holding cost. A small trade has 1000 shares or less, a medium trade has greater than 1000 but less than 10,000 shares, and a large trade has 10,000 or more shares. The last two rows report the average statistics for all stocks.

inventory models, changes in quotes affect the subsequent arrival rate of trades. After a public sale (purchase) at the bid (ask), the dealer lowers (raises) the bid (ask) relative to the fundamental stock price in order to increase the probability of a subsequent public purchase (sale) [see, e.g., Ho and Stoll (1981)]. The dealer is then compensated

Table 4
Restricted models of traded spread and order processing component by trade size

Panel A

Company	Constraint 1: Traded spread is the same but order processing costs are allowed to vary across trade sizes.		Constraint 2: Traded spread and order processing costs are constrained to be the same across trade sizes.	
	$\chi^2(2)$	*P*-value	$\chi^2(4)$	*P*-value
AXP	171.12	0.0000	1347.00	0.0000
CHV	91.96	0.0000	917.39	0.0000
DD	78.18	0.0000	2395.00	0.0000
DOW	46.63	0.0000	2428.00	0.0000
EK	4.96	0.0839	1697.00	0.0000
GE	300.12	0.0000	3004.00	0.0000
GM	43.07	0.0000	1973.00	0.0000
IBM	77.98	0.0000	4305.00	0.0000
IP	38.08	0.0000	369.62	0.0000
JNJ	44.49	0.0000	1166.00	0.0000
KO	31.77	0.0000	2378.00	0.0000
MMM	147.66	0.0000	535.86	0.0000
MO	319.51	0.0000	2475.00	0.0000
MOB	33.92	0.0000	1144.00	0.0000
MRK	215.56	0.0000	1998.00	0.0000
PG	27.24	0.0000	795.29	0.0000
S	124.95	0.0000	1794.00	0.0000
T	137.53	0.0000	2466.00	0.0000
XON	94.37	0.0000	2203.00	0.0000

Panel B

	Traded Spread		Adverse Selection and Inventory Holding					
	All Sizes		Small		Medium		Large	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constraint 1	0.1215	0.0004	0.0327	0.0028	0.2207	0.0051	0.3809	0.0162
	All Sizes				All Sizes			
Constraint 2	0.1237	0.0004			0.0838	0.0023		

Panel A presents tests of overidentifying constraints on the basic model [Equation (15)] of traded spread and order processing component estimated in Table 3. The reported chi-square statistics have the number of degrees of freedom equal to the number of orthogonality conditions minus the number of parameters to be estimated.
Panel B presents the averages of the constrained estimates across companies for the constraints defined in panel A.

for inventory risk because the expected midquote change is positive after a dealer sale and negative after a dealer purchase. The probability of a purchase (sale) is greater than 0.5 just after a sale (purchase). In other words, under an inventory model, negative serial covariance in trades (Q_t) is induced. As trades reverse, quotes reverse. Consequently, under inventory models negative serial correlation in quote changes (as well as in trades) is induced, and this implication can be used, at least in principle, to identify the inventory component. The negative serial correlation in trades, Q_t, and quote changes, ΔM_t is separate from, and in addition to, the negative serial correlation in transaction price changes, ΔP_t resulting from the bid-ask bounce (as in the Roll model, for example).

5.1 The extended model with induced serial correlation in trade flows

Equations (1)–(5) make no assumption about the probability of trades and therefore cannot distinguish inventory and adverse information effects. We modify the model to reflect the serial correlation in trade flows. The conditional expectation of the trade indicator at time $t-1$, given Q_{t-2} is easily shown to be[10]

$$E(Q_{t-1}|Q_{t-2}) = (1-2\pi)Q_{t-2}, \tag{21}$$

where π is the probability that the trade at t is opposite in sign to the trade at $t-1$. Once we allow π to differ from one-half, Equation (1) must be modified to account for the predictable information contained in the trade at time $t-2$. On the assumption that the market knows Equation (21), the change in the fundamental value will be given by

$$\Delta V_t = \alpha\frac{S}{2}Q_{t-1} - \alpha\frac{S}{2}(1-2\pi)Q_{t-2} + \varepsilon_t, \tag{22}$$

where the second term on the right-hand side subtracts the information in Q_{t-1} that is not a surprise. When $\pi = 0.5$, the sign of the trade is totally unpredictable and Equation (22) reduces to Equation (1). Notice that changes in the fundamental value, ΔV_t, are serially uncorrelated and unpredictable since the changes are induced by trade innovations (the first two terms) and unexpected public information releases (the last term). Consider, for example, the expectation of ΔV_t conditional on the information after Q_{t-2} is observed but before Q_{t-1} is observed:

$$E(\Delta V_t | V_{t-1}, Q_{t-2}) = 0.$$

One cannot predict the change in underlying value from past public or past trade information.

By combining Equations (22) and (2) we obtain

$$\Delta M_t = (\alpha+\beta)\frac{S}{2}Q_{t-1} - \alpha\frac{S}{2}(1-2\pi)Q_{t-2} + \varepsilon_t. \tag{23}$$

Note that in arriving at Equation (23) we used Equation (2) directly without modification for the expected sign of the trade. What matters for inventory costs is the actual inventory effect, not the unexpected portion. There is inventory risk only when inventory is acquired (even if the inventory change was expected), and there is no inventory risk if inventory is not acquired (even if the lack of inventory change was

[10] The conditional expectation may be readily calculated from the fact that $Q_{t-1} = Q_{t-2}$ with probability $(1-\pi)$, and $Q_{t-1} = -Q_{t-2}$ with probability π.

unexpected). Consequently quote adjustments for inventory reasons depend on actual trades, not trade surprises. This distinction is what allows us to estimate separately the inventory and adverse information components.

Unlike the expected change in underlying value, the expected change in the quote midpoint can be predicted on the basis of past trades. Consider the expectation of ΔM_t conditional on the information available after M_{t-1} (and therefore Q_{t-2}) is observed, but before Q_{t-1} and M_t are observed:

$$E(\Delta M_t | M_{t-1}, Q_{t-2}) = \beta\frac{S}{2}(1 - 2\pi)Q_{t-2}. \tag{24}$$

First, the expected quote midpoint change does not depend on the adverse information component. Although the observed quote midpoint change does depend on the adverse information component, the expected change does not because the change in the true value V_t is serially uncorrelated. Second, the conditional expectation highlights the important result that the expected change in quotes is very much smaller than and of opposite sign from the immediate change in quotes following a trade. In the absence of any change in the fundamental value of the stock, the immediate inventory response of quotes to a trade is $\beta(S/2)Q_{t-2}$, while the expected change in quotes is the right-hand side of Equation (24), which is much smaller. Equation (24) reflects the fact, noted by several authors, that inventory adjustments are long lived and difficult to observe [see, e.g., Hasbrouck (1988) and Hasbrouck and Sofianos (1993)]. When inventories are slow to revert to their desired levels, π is close to one-half, which lowers the expected reversal in Equation (24). The equation indicates that what is being measured by the expected quote change is how much of the inventory-induced quote adjustment is expected to be reversed in the subsequent trade, not the amount of the spread that is due to inventory. For example, if $\beta = 0.25$, $S/2 = 10$ cents, $\pi = 0.7$, the immediate inventory response of quotes to a trade at the bid is $\beta(S/2)Q_{t-2} = -2.5$ cents, but the expected change in quotes in the subsequent trade is $\beta(S/2)(1 - 2\pi)Q_{t-2} = +1$ cent.

Combining Equations (23) and (4) yields

$$\Delta P_t = \frac{S}{2}Q_t + (\alpha + \beta - 1)\frac{S}{2}Q_{t-1} - \alpha\frac{S}{2}(1 - 2\pi)Q_{t-2} + e_t, \tag{25}$$

which is the analog of Equation (5).[11] Estimation of the traded spread S, the three components of the spread α, β, and $1 - \alpha - \beta$, and

[11] The expected price change per share to the supplier of immediacy who buys or sells at time $t - 1$

the probability of a trade reversal π can then be accomplished by estimating Equations (21) and (25) simultaneously.

It is possible to estimate the components of the spread directly from the quote-change equation if estimates of traded spread are not needed. Specifically, do not combine Equation (23) with the specification of a traded spread in Equation (4) as derived in Equation (25), but instead consider the following variant of Equation (23):

$$\Delta M_t = (\alpha + \beta)\frac{S_{t-1}}{2}Q_{t-1} - \alpha(1 - 2\pi)\frac{S_{t-2}}{2}Q_{t-2} + e_t, \qquad (26)$$

where the constant traded spreads in Equation (23) are replaced with observed posted spreads.[12] We estimate the extended model consisting of Equations (21) and (26). The parameter space is reduced by not estimating the traded spread, something that simplifies the empirical implementation, particularly for the model with trade size categories.

We now incorporate size categories into the extended model. When trade size categories are considered, the π estimate will differ according to the trade size categories at time $t-2$ and $t-1$. For example, when a small trade is observed at $t-1$, the probability of a reversal that is expected will depend on whether the previous trade at $t-2$ was a small, medium, or large trade. We denote the reversal probabilities as π^{ij}, where the superscript i refers to the trade size category at $t-2$ and j refers to the trade size category at $t-1$. The extended model with size categories is

$$Q_t^j = (1 - 2\pi^{ij})Q_{t-1}^i + \zeta_t^{ij} \qquad (27)$$

$$\Delta M_t^{ij} = (\alpha^{ij} + \beta^{ij})\frac{S_{t-1}^j}{2}Q_{t-1}^j - \alpha^{ij}(1 - 2\pi^{ij})\frac{S_{t-2}^i}{2}Q_{t-2}^i + e_t^{ij}, \qquad (28)$$

---

is from Equation (25):

$$E(\Delta P_t | Q_{t-1}) = \frac{S}{2}(1 - 2\pi)Q_{t-1} - \frac{S}{2}Q_{t-1} + \beta\frac{S}{2}Q_{t-1} + \alpha\frac{S}{2}Q_{t-1}^*$$

where $Q_{t-1}^* = Q_{t-1} - (1-2\pi)Q_{t-2}$ is the unexpected trade sign. The first term is the trade reversal induced if $\pi > 0.5$. The second term is the usual reversal. The third term is an attenuation of the reversal due to the adjustment of quotes in response to inventory effects. The fourth term is the attenuation of the reversal due to permanent changes in quotes to reflect the information contained in the trade at $t-1$. The expected reversal can be shown to be the same as Stoll's (1989) expected reversal of $(\pi - \delta)S$ (where $\delta = (\alpha + \beta)/2$) except for the difference between Q_{t-1}^* and Q_{t-1}. Stoll assumes that the trade is totally unanticipated so that $Q_{t-1}^* = Q_{t-1}$.

[12] Since the posted spreads are coupled with trade indicator variables that code midpoint trades as zeroes, midpoint trades are still ignored as in a traded spread.

where

$Q_t^s = Q_t$	if share volume at $t \leq 1000$ shares
	is censored otherwise
$Q_t^m = Q_t$	1000 shares $<$ if share volume at $t < 10,000$ shares
	is censored otherwise
$Q_t^l = Q_t$	is share volume at $t \geq 10,000$ shares
	is censored otherwise.

The model, Equations (27) and (28), is fairly complex for it requires the estimation of nine different values of α and β for each of the nine trade size transitions between $t-2$ and $t-1$. We use data points for each possible transition and censor the data otherwise.

5.2 Empirical results based on serial correlation in trade flows

The GMM procedure is easily modified to estimate the extended model by considering an expanded set of orthogonality conditions. The empirical results are in panel A of Table 5 for the extended model without size categories, that is, Equations (21) and (26). The order processing component $(1-\alpha-\beta)$ is surprisingly large, averaging 84% of the traded spread across all stocks, slightly higher than the results of the basic model in Table 2. Most striking in panel A is the negative value of the adverse selection component. This result may be traced to estimates of π that are less than 0.5. When π is less than 0.5, changes in ΔV_t are attenuated, which can be seen by examining Equation (22). If the change in the stock's underlying value in reaction to a trade is reduced (because the sign of the trade is anticipated), the change in the stock's quote midpoint ascribed to inventory effects is increased. Consequently the net effect is to reduce α and raise β. However, a negative value of α seems unreasonable.

The estimation results for the extended model with size categories are in panel B. The results reported are the averages across companies for each ij size transition. As in panel A, the adverse selection components are negative, with the exception of large-to-large category, and the π estimates are all less than 0.5.

While a negative value of α and a value of π less than 0.5 are empirically possible, such values are theoretically impermissible under the class of spread models we consider. Adverse information cannot be negative so long as at least one investor can be better informed than those investors or dealers setting quotes. Spread models also specify the lower bound for π as 0.5 when there are no inventory holding costs. Since a market maker recovers inventory holding costs from trade and quote reversals, negative serial covariance in trade flows is required. However, estimates of π less than 0.5 indicate positive serial covariance in trade flows.

Table 5
Components of the spread: Estimates based on serial correlation in trade flows

Panel A

Company	No. of Obs.	α Coeff.	α Std. Err.	β Coeff.	β Std. Err.	π Coeff.	π Std. Err.
AXP	68,583	−0.0526	0.0024	0.1209	0.0029	0.2080	0.0018
CHV	47,753	−0.0298	0.0043	0.2388	0.0045	0.1466	0.0019
DD	71,913	−0.0732	0.0030	0.2669	0.0034	0.1732	0.0016
DOW	62,298	−0.0508	0.0040	0.2886	0.0044	0.2181	0.0018
EK	71,290	−0.0415	0.0026	0.1499	0.0029	0.1986	0.0017
GE	122,930	−0.0314	0.0018	0.1861	0.0020	0.0896	0.0011
GM	103,817	−0.0247	0.0017	0.1114	0.0019	0.0939	0.0012
IBM	144,362	−0.0222	0.0020	0.2627	0.0022	0.0711	0.0009
IP	43,746	−0.0216	0.0061	0.2611	0.0064	0.2094	0.0022
JNJ	128,960	−0.0233	0.0027	0.1674	0.0028	0.1878	0.0013
KO	123,450	−0.0130	0.0017	0.1207	0.0019	0.1198	0.0012
MMM	41,154	−0.0109	0.0055	0.2502	0.0054	0.1667	0.0021
MO	180,840	−0.0463	0.0015	0.1404	0.0017	0.1476	0.0010
MOB	57,981	−0.0177	0.0052	0.1915	0.0053	0.2401	0.0020
MRK	145,952	0.0170	0.0022	0.1524	0.0019	0.0838	0.0010
PG	70,356	−0.0334	0.0043	0.2166	0.0043	0.2211	0.0018
S	55,960	−0.0222	0.0031	0.1641	0.0033	0.1284	0.0018
T	144,646	−0.0311	0.0009	0.0684	0.0012	0.1540	0.0012
XON	72,649	−0.0673	0.0031	0.1920	0.0036	0.1910	0.0017
AVG.	92,560	−0.0314	0.0031	0.1868	0.0033	0.1605	0.0015

Panel B

Trade Size Category	Avg. No. of Obs.	α Coeff.	α Std. Err.	β Coeff.	β Std. Err.	π Coeff.	π Std. Err.
Large to large	649	0.0512	0.0815	0.3803	0.0497	0.0843	0.0152
Large to medium	1790	−0.0676	0.0280	0.3430	0.0295	0.1373	0.0104
Large to small	2039	−0.0525	0.0341	0.2297	0.0371	0.2565	0.0110
Medium to large	1555	−0.0513	0.0387	0.4476	0.0377	0.1464	0.0126
Medium to medium	11,036	−0.0410	0.0092	0.3015	0.0093	0.0879	0.0035
Medium to small	15,386	−0.0480	0.0076	0.1674	0.0082	0.1894	0.0038
Small to large	2298	−0.0311	0.0438	0.4442	0.0446	0.2597	0.0114
Small to medium	15,176	−0.0299	0.0087	0.2677	0.0091	0.1703	0.0038
Small to small	42,631	−0.0157	0.0037	0.0928	0.0039	0.1717	0.0024

The table presents the results of using the serial correlation in trade flows to estimate the components of the spread. α is the estimated adverse selection component of the spread, β is the estimated inventory holding component of the spread, and π is the estimated probability of a trade reversal.

In Panel A, the model consists of Equations (21) and (26) and does not distinguish between trade size categories. AVG denotes the average statistics for all stocks.

In Panel B, the model consists of Equations (27) and (28) and is estimated for each trade size category separately and averages across companies are reported. Trade size categories refer to trade sizes at time $t-2$ and $t-1$ when changes in midpoint quotes occur over $t-1$ to t. A small trade size has 1000 shares or less, a medium trade size has greater than 1000 but less than 10,000 shares, and a large trade size has 10,000 or more shares.

A source of positive serial correlation is that orders are broken up as they are executed. A large order may, for example, be negotiated at a single price but be reported in a series of smaller trades. Alternatively, a single large limit order may be executed at a single price against various incoming market orders. In other words, orders could be negatively serially correlated as theory suggests, but the trades we observe are positively serially correlated. The empirical effect is to lower both π and α. One approach to dealing with this problem is to collapse a sequence of related trade reports to just one order, that is, to bunch related data. To align trades and orders, we define component trades of a broken-up order to be sequential trades at the same price on the same side of the market without any change in bid or ask quotes. Under this definition there is a one-to-one relation between an order and a trade. In effect, our approach is to treat a cluster of trades at the same price and unchanged quotes as a single order. This procedure overcorrects for the problem because it aggregates some independent orders, and consequently the empirical results from this dataset provide an upper bound on the adverse information component and the probability of price reversal.

Alternatively we could propose a model of order submission by investors which could be used to separate demand side serial correlation in order flow from microstructure-induced serial correlation. Our extreme assumption assigns all clustering to microstructure factors and consequently bunches together trades that are independent. Another approach would be to impose additional filters for bunching trades. For example, all sequential trades at the same price with no change in quotes within a 2-minute window may be deleted. Unfortunately, the effectiveness of such a procedure to identify broken-up trades is likely to be time varying and stock specific and is not implemented. In the end we rely on our simple procedure, recognizing that it provides an upper bound on α and π.

Table 6 presents the results based on the bunched dataset that collapses all sequential trades at the same price with no quote adjustments to just one trade. Panel A provides estimates for the extended model without size categories. The reversal probabilities are all greater than 0.5, averaging 0.87. The adverse information component is now positive, averaging 9.6% of the spread over all stocks, and is smaller than the inventory component, which averages 28.7% of the spread. The order processing component remains the largest single component, averaging 61.7% across all stocks.

We now turn to the question of how trade size in the extended model affects estimates of the components of the spread. The results are in panel B of Table 6. For this purpose, the bunching procedure used in panel A is retained, and the size category is determined by the sum of all the share volumes of trades bunched together. We present average values across the stocks in our sample.[13] The π coefficients exceed 0.5 and are of similar magnitude to those in panel

[13] The average is across 19 companies except for the large-to-large category for which one company is dropped due to a paucity of observations.

Table 6
Components of the spread: Estimates based on serial correlation in trade flows with bunching

Panel A

Company	No. of Obs.	α Coeff.	α Std. Err.	β Coeff.	β Std. Err.	π Coeff.	π Std. Err.
AXP	27,562	0.1647	0.0075	0.1386	0.0036	0.9740	0.0011
CHV	20,662	0.0384	0.0091	0.3771	0.0096	0.8111	0.0027
DD	32,871	0.0650	0.0051	0.3712	0.0051	0.8774	0.0019
DOW	32,360	0.0460	0.0056	0.3830	0.0058	0.8463	0.0021
EK	29,425	0.1039	0.0051	0.2010	0.0040	0.9282	0.0015
GE	44,496	0.1009	0.0050	0.3533	0.0047	0.8748	0.0015
GM	32,592	0.1727	0.0065	0.1751	0.0046	0.9087	0.0015
IBM	51,756	0.1101	0.0061	0.5802	0.0057	0.8367	0.0016
IP	23,658	−0.0160	0.0121	0.3800	0.0126	0.7213	0.0028
JNJ	55,996	0.1225	0.0052	0.2343	0.0045	0.8907	0.0013
KO	44,297	0.1231	0.0051	0.2091	0.0037	0.9298	0.0012
MMM	20,326	−0.0136	0.0121	0.4280	0.0132	0.7321	0.0030
MO	73,037	0.0798	0.0038	0.2046	0.0030	0.9168	0.0011
MOB	28,122	0.0645	0.0082	0.2772	0.0078	0.8178	0.0022
MRK	47,060	0.1036	0.0053	0.2894	0.0055	0.8012	0.0018
PG	35,039	0.0254	0.0060	0.3009	0.0064	0.8290	0.0020
S	20,272	0.1475	0.0088	0.2547	0.0070	0.8806	0.0021
T	50,721	0.2094	0.0053	0.0744	0.0019	0.9804	0.0007
XON	30,668	0.1737	0.0082	0.2111	0.0050	0.9248	0.0014
AVG	36,891	0.0959	0.0068	0.2865	0.0060	0.8675	0.0018

Panel B

Trade Size Category	Avg No. of Obs.	α Coeff.	α Std. Err.	β Coeff.	β Std. Err.	π Coeff.	π Std. Err.
Large to large	1364	0.0514	0.1091	0.6645	0.1087	0.6759	0.0125
Large to medium	2469	0.2199	0.0529	0.2310	0.0515	0.7004	0.0088
Large to small	1668	0.1055	0.0430	0.1076	0.0399	0.7936	0.0098
Medium to large	2293	0.0144	0.0394	0.7143	0.0386	0.8208	0.0085
Medium to medium	7359	0.1011	0.0189	0.3468	0.0175	0.8018	0.0043
Medium to small	6485	0.1141	0.0146	0.0919	0.0112	0.8807	0.0038
Small to large	1807	0.0210	0.0355	0.7371	0.0298	0.9114	0.0073
Small to medium	6314	0.0654	0.0155	0.3773	0.0128	0.9128	0.0036
Small to small	7199	0.1049	0.0128	0.0980	0.0082	0.9406	0.0029

The table presents the results of using the serial correlation in trade flows to estimate the components of the spread when trades are bunched. α is the estimated adverse selection component of the spread, β is the estimated inventory holding component of the spread, and π is the estimated probability of a trade reversal. Bunching collapses all sequential trades at the same price and the same quotes to one bunched trade. The trade size for the bunched trade is the sum of all the trades bunched together.
In Panel A, the model consists of Equations (21) and (26) and does not distinguish between trade size categories. AVG denotes the average statistics for all stocks.
In Panel B, the model consists of Equations (27) and (28) and is estimated for each trade size category separately and averages across companies are reported. Trade size categories refer to trade sizes at time $t-2$ and $t-1$ when changes in midpoint quotes occur over $t-1$ to t. A small trade size has 1000 shares or less, a medium trade size has greater than 1000 but less than 10,000 shares, and a large trade size has 10,000 or more shares.

A, but there is now variability across trade size sequences. For example, the probability of reversal, π, is clearly larger for a sequence of two small trades than for a sequence of two large trades. The adverse information component continues to be small, consistent with the findings in panel A. It varies from 1.44 to 21.99% depending on the sequence of trade sizes used in the estimation. The inventory components vary considerably by trade size sequence, ranging from a low of 9.19% for sequences of a medium trade followed by a small trade to a high of 73.71% for sequences of a small trade followed by a large trade.

Estimates of the basic model with trade size (Table 3) showed that quote adjustments depend on trade size, but that model could not determine how much of the effect was due to adverse information and how much was due to inventory. The results from the extended model with trade size in panel B of Table 6 indicate that a substantial portion of the trade size effect is due to inventory effects. For example, the estimate based on sequences of two large trades implies that the adverse information associated with the second trade in the sequence is 0.0514 of the spread, and the inventory effect is 0.6645. The order processing effect is therefore 0.2841. The estimate based on sequences of two small trades implies that the adverse information associated with the second trade in the sequence is 0.1049, and the inventory effect is 0.0980. The order processing effect is therefore 0.7971. While the estimate of adverse information is again small, the differences between these two cases (large-large versus small-small) in the inventory and order processing cost components are quite reasonable. Since there is a large fixed order processing cost component to a trade, a small trade will have larger order processing cost per share than a large trade. On the other hand, inventory costs rise with trade size.

The estimates of α are smaller for sequences of two trades that end in a large trade than for sequences that end in a medium or small trade. In other words, not only is the adverse information effect small overall, it tends to be smaller as a fraction of the spread for large trades than for small or medium trades. There are three possible explanations for this result. First, a large trade is often prenegotiated upstairs and may be preceded by transactions that convey information about the block. The fact that the quoted spread prior to a large trade is large, as shown in Table 7, suggests that this is the case. In other words, a large trade itself may not convey much information and may not have much subsequent impact on quotes beyond that contained in the existing spreads.[14] Second, large trades may be facilitated through the upstairs market where brokers can reputationally "certify" the trade. Third, a small adverse information cost component is still a large dollar amount in the case of a large trade.

[14] The causality could easily go the other way and wider spreads may attract large trades. For example, Madhavan and Cheng (1996) find that the probability of an upstairs trade is greater when spreads are wide because upstairs intermediation offers potentially lower costs.

Table 7
Average quoted half-spreads for alternative sequences of trade size

Trade size at $t-2$		Trade size at $t-1$ Large	Medium	Small
Large	S_{t-1}	0.0978	0.0971	0.0952
	S_{t-2}	0.0901	0.0904	0.0891
Medium	S_{t-1}	0.0909	0.0929	0.0896
	S_{t-2}	0.0886	0.0918	0.0890
Small	S_{t-1}	0.0832	0.0838	0.0825
	S_{t-2}	0.0830	0.0880	0.0874

The table presents averages of quoted half-spreads, S_t, at times $t-2$ and $t-1$ for alternative sequences of trade size. Trade size categories refer to trade sizes at time $t-2$ and $t-1$ when changes in midpoint quotes occur over $t-1$ to t. A small trade size has 1000 shares or less, a medium trade size has greater than 1000 but less than 10,000 shares, and a large trade size has 10,000 or more shares.

To facilitate interpretation of the estimates in Table 6, we simulate in Table 8 quote midpoint changes on the basis of the estimates in panel B of Table 6. The quote changes are stated as a fraction of the average half-spread at time $t-1$ and are calculated according to the estimated Equation (28), a notationally simplified version of which is

$$\frac{\Delta M_t}{\frac{1}{2}\bar{S}_{t-1}} = \hat{\alpha} Q_{t-1} - \hat{\alpha}(1-2\hat{\pi})\frac{\bar{S}_{t-2}}{\bar{S}_{t-1}} Q_{t-2} + \hat{\beta} Q_{t-1}, \tag{29}$$

where $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\pi}$ are the estimated parameters of Equation (28) that are understood to be for a particular trade size sequence, where Q_{t-1} and Q_{t-2}, are the signs of the sequence of trades, and where $\bar{S}_{t-1}$ and $\bar{S}_{t-2}$ are the average quoted spreads for a particular trade size sequence. Table 7 provides the average values of $\bar{S}_{t-1}$ and $\bar{S}_{t-2}$ for the nine possible size sequences. In panel A of Table 8, we present the midpoint price changes over the period $t-1$ to t as a fraction of the average half-spread at time $t-1$ for the trade sign sequences, $Q_{t-2} = -1$, $Q_{t-1} = +1$ (bid to ask), $Q_{t-2} = +1$, $Q_{t-1} = +1$ (ask to ask), $Q_{t-2} = 0$, $Q_{t-1} = +1$ (mid to ask), and $Q_{t-2} = -1$, $Q_{t-1} = 0$ (bid to mid), and for all possible size sequences. In panel B of Table 8, we present the amount of the midpoint change that is due to adverse information, the first two terms on the right-hand side of Equation (29), and the amount due to inventory, the last term on the right-hand side of Equation (29).

Consider, for example, a large trade that occurs at the ask at time $t-1$. What is the reaction of the quote midpoint to this trade? If the prior trade at $t-2$ was large and was at the bid, the midpoint adjusts upward by 0.6992 of the half-spread (see the first case in panel

Table 8
Midpoint quote changes implied by estimates based on serial correlation in trade flows for bunched trades in Table 6

Panel A

Trade size at $t-2$	Bid at $t-2$ to ask at $t-1$			Ask at $t-2$ to ask at $t-1$			Mid at $t-2$ to ask at $t-1$			Bid at $t-2$ to mid at $t-1$		
	Trade size at $t-1$											
	Large	Medium	Small	Large	Medium	Small	Large	Medium	Small	Large	Medium	Small
Large	0.6992	0.3688	0.1551	0.7326	0.5330	0.2711	0.7159	0.4509	0.2131	−0.0167	−0.0821	−0.0580
Medium	0.7197	0.3876	0.1197	0.7377	0.5082	0.2923	0.7287	0.4479	0.2060	−0.0090	−0.0603	−0.0863
Small	0.7409	0.3860	0.1050	0.7753	0.4994	0.3008	0.7581	0.4427	0.2029	−0.0172	−0.0567	−0.0979

Panel B

Trade size at $t-2$		Bid at $t-2$ to ask at $t-1$			Ask at $t-2$ to ask at $t-1$			Mid at $t-2$ to ask at $t-1$			Bid at $t-2$ to mid at $t-1$		
		Trade size at $t-1$											
		Large	Medium	Small	Large	Medium	Small	Large	Medium	Small	Large	Medium	Small
Large	Info.	0.0347	0.1378	0.0475	0.0681	0.3020	0.1635	0.0514	0.2199	0.1055	−0.0167	−0.0821	−0.0580
	Inv.	0.6645	0.2310	0.1076	0.6645	0.2310	0.1076	0.6645	0.2310	0.1076	0.0000	0.0000	0.0000
Medium	Info.	0.0054	0.0408	0.0278	0.0234	0.1614	0.2004	0.0144	0.1011	0.1141	−0.0090	−0.0603	−0.0863
	Inv.	0.7143	0.3468	0.0919	0.7143	0.3468	0.0919	0.7143	0.3468	0.0919	0.0000	0.0000	0.0000
Small	Info.	0.0038	0.0087	0.0070	0.0382	0.1221	0.2028	0.0210	0.0654	0.1049	−0.0172	−0.0567	−0.0979
	Inv.	0.7371	0.3773	0.0980	0.7371	0.3773	0.0980	0.7371	0.3773	0.0980	0.0000	0.0000	0.0000

The calculations in the table use the average quoted half-spreads presented in Table 7. The estimates are based on the model with serial correlation in trade flows with bunched trades presented in Table 6, panel B. Trade size categories refer to trade sizes at time $t-2$ and $t-1$ when changes in midpoint quotes occur over $t-1$ to t. A small trade size has 1000 shares or less, a medium trade size has greater than 1000 but less than 10,000 shares, and a large trade size has 10,000 or more shares. Bunching collapses all sequential trades at the same price and the same quotes to one bunched trade. The trade size for the bunched trade is the sum of all the trades bunched together. The estimates not shown are the negative of those presented when bids and asks are replaced by asks and bids, respectively.

Panel A presents implied changes in the quote midpoint as a proporition of the quoted half-spread for a given buy-sell sequence (Q_{t-2} followed by Q_{t-1}) and alternative trade size categories (small, medium, and large).

Panel B decomposes the implied change into the amount due to information conveyed by the trade and the amount due to inventory effects.

A). If the prior trade at the bid was medium, the upward adjustment is 0.7197, and if the prior trade at the bid was small, the upward adjustment is 0.7409. In other words, the quote adjustment depends primarily on the size of the observed trade, and less so on the size of the prior trade. According to the estimates of α and β in Table 6, the adverse information component of the upward adjustment of 0.6992 is 0.0347 and the inventory component is 0.6645 as shown in panel B. The adverse information component is less than the estimate of α in Table 6 of 0.0514 because a trade at the ask following a trade at the bid is not a big surprise. Such a sequence reflects the normal bid-ask bounce and consequently does not reflect a great deal of information. The pattern is similar, albeit the adverse information component is smaller, for a large trade preceded by a small trade at the bid. In that case, the upward adjustment of the midpoint, which is 0.7409, is composed of an adverse information component of 0.0038 and an inventory component of 0.7371. Again, the adverse information effect is small because the sequence of bid to ask is to be expected.

Consider now a medium trade at the ask preceded by a trade at the bid. The midpoint adjustment for a medium trade is about 0.37 to 0.39, considerably less than if the trade is large. However, the portion of the change ascribable to adverse information is substantially larger. For a medium trade at the ask preceded by a large trade at the bid, the adverse information component accounts for 0.1378 of the midpoint change and the inventory effect accounts for 0.2310 of the midpoint change.

A more pronounced adverse information effect is evident for a sequence of two trades at the ask, which is the second case in Table 8. This is evident in all size categories of trades at $t-1$ but is particularly evident for medium trades. The estimated adverse information effect is larger in a sequence of two trades at the ask because this sequence is a surprise, and consequently the trade conveys adverse information. The probability that the trader at time $t-1$ is informed is greater if the trade is counter to the usual bid-ask bounce as given by the probability π. The medium size trade may have the larger adverse information component because it is not prenegotiated and may therefore reflect the presence of an informed trader on the floor. The inventory component is independent of the prior trade.

The third case in Table 8 depicts midpoint adjustments after the trade sequence mid to ask. The midpoint adjustment following such a sequence is typically less than for the ask to ask sequence and greater than the bid to ask sequence. This difference reflects differences in the information content of the trades. A trade at the ask after a trade at the midpoint conveys more adverse information than a trade at the ask after a trade at the bid because the second sequence is more likely than the first. The values in Table 8 for this case are the same as the estimates of α and β in Table 6, since $Q_{t-2} = 0$.

The fourth case in Table 8 depicts midpoint adjustments after the trade sequence of bid to mid. There is little quote adjustment after a trade at the midpoint since such a trade has no identifiable inventory effect and has only a small adverse information effect. The small adverse information effect arises from the fact that after a trade at the bid, a trade at the midpoint is not expected. The fact that the trade was not at the ask conveys some information.

Summarizing, our purpose has been to develop a model that separates inventory and adverse information effects on the basis of the time-series characteristics of trades and quotes. We have shown that this is possible. We have also shown that the size of these two effects is a function of the sequence of trade sizes and trade signs. Our estimates imply that for our sample of large, actively traded stocks, the adverse information effect is small.

6. Three-Way Decomposition of the Spread Based on Buying and Selling Pressure

A second approach to distinguishing the adverse selection and inventory components of the spread makes use of the fact that inventory-induced quote changes result not only from inventory changes in the stock being examined but also from inventory changes in other stocks. This approach assumes that adverse information is specific to the stock but that inventory effects relate to the entire portfolios held by the suppliers of liquidity. Specifically, we model liquidity suppliers as taking a portfolio perspective in adjusting inventories of specific stocks.[15] A liquidity supplier who buys stock j at the bid will not only lower the bid and ask prices of stock j, but will also lower the bid and ask prices of other correlated stocks, the sale of which would hedge his position in stock j. Conversely, if other stocks are subject to buying pressure, the liquidity supplier may decide not to lower the bid and ask prices of stock j, for he wishes to encourage additional sales of stock j to hedge his purchases of other stocks. The portfolio approach allows for the possibility that the quotes in stock j are adjusted by a different amount than implied by the information content or inventory effect in stock j alone. In particular, selling (buying) pressure in other stocks will produce quote changes in the stock j as the liquidity suppliers attempt to keep their overall portfolios in balance.

[15] Ho and Stoll (1983) model the relation between quote changes in a specific stock and inventory changes in other stocks. They show that the quote adjustment in stock k in response to a trade in asset $*$ depends on $\text{cov}(R_k, R_*)/\sigma^2(R_*)$.

These kinds of adjustments were perhaps most obvious during the market crash of October 19, 1987, when relentless selling pressure in the absence of any specific news produced inventory-induced quote changes in specific stocks.

6.1 The extended model based on portfolio trading pressure

The basic model [Equation (5)] ignores the impact of general trading pressure since inventory adjustment modeled in Equation (2) is determined solely by own stock inventory. We now distinguish between transaction types of different stocks. Let superscript k designates security k, and replace Equation (2) with

$$M_t^k = V_t^k + \beta^k \frac{S^k}{2} \sum_{i=1}^{t-1} Q_i^*, \tag{30}$$

where Q_t^* is the aggregate buy-sell indicator variable. The aggregate buy-sell indicator variable is measured as

$$\begin{aligned} Q_t^* &= 1 \qquad \text{for} \quad \sum_{k=1}^{n} Q_t^k > 0 \\ &= -1 \qquad\qquad \sum_{k=1}^{n} Q_t^k < 0 \\ &= 0 \qquad\qquad\;\; \sum_{k=1}^{n} Q_t^k = 0, \end{aligned}$$

where n is the number of securities that liquidity suppliers span to get a feel for the market direction.

With Equation (30), the model becomes

$$\Delta P_t^k = \frac{S^k}{2} \Delta Q_t^k + \alpha^k \frac{S^k}{2} Q_{t-1}^k + \beta^k \frac{S^k}{2} Q_{t-1}^* + e_t^k. \tag{31}$$

Equation (31) still confines the determinants of prices and quotes to trade indicator variables and produces Equation (5) as a special case when there are no spillover effects from other stocks. However, unlike Equation (5), all the components of the bid-ask spreads are separately identifiable.

6.2 Empirical results based on portfolio trading pressure

The extended model [Equation (31)] can be estimated using the GMM procedure described above with appropriate modifications to the orthogonality conditions. We illustrate the results of estimating Equation

(31) using our sample of 19 MMI securities.[16] For each stock in our sample, the aggregate trade indicator variable is calculated from the trade data using all the stocks in the sample. To estimate Equation (31), an alignment of trading time across securities is needed. We specify a 5-minute interval on our transactions data for the temporal alignment. The underlying data are unchanged except for the addition of a filter that confines all data to 9:30 A.M. to 4:00 P.M. EST. To obtain the 5-minute interval, we retain the latest transaction price and pair it with the latest quotes that are more than 5 seconds prior to the trade for each 5-minute interval. If security k does not trade within the 5-minute interval, Q_t^k is set to missing in Equation (31), but is set to zero in constructing our aggregate buy-sell indicator variable.

The results of estimating Equation (31) are displayed in Table 9, panel A. The results yield an order processing component that averages 68.9%. The adverse information component averages 21.5% and the inventory component averages 9.6%. The time alignment necessary to estimate Equation (31) in part solves the trade clustering problems that concerned us earlier. That is because we take only one trade from each 5-minute interval and eliminate clustered trades within the 5-minute interval. Nevertheless, our 5-minute sample contains some clustered trades that extend across 5-minute intervals. In panel B of Table 9, we provide for completeness, estimates of Equation (31) where all sequential 5-minute observations at the same price and the same quotes are collapsed to one observation. This is an overcorrection for clustering since we may collapse 5-minute observations where there were intervening price or quote changes.[17] Panel B of Table 9 shows, as in Table 6, that the clustering adjustment reduces the order processing component and increases the adverse selection component. In panel B, the order processing component averages 41.7%, the adverse selection component averages 46.2%, and the inventory component averages 12.1%.

The order processing component estimated by the cross-section approach and reported in panel A of Table 9 is of the same magnitude as the estimate from the time-series approach in Table 6, panel A (68.9% versus 61.8%). With regard to the other spread components, the cross-section approach yields a smaller inventory component and a larger adverse information component than the time-series approach in Table 6. This reflects the fact that we identify the inventory effect on the basis of a reaction to trades in a portfolio that is probably measured with error and consequently leads to an underestimate of

[16] We do not estimate the model with size categories because of the increased complexity in implementing it.

[17] The average number of observations per stock is over 16,000 for the estimates in panel A of Table 9 and over 11,000 for the estimates in panel B of Table 9.

Table 9
Traded spread and its components: Estimates based on market trading pressure

		Panel A			Panel B		
Company	Estimate	α	β	S	α	β	S
AXP	Coeff.	0.0581	0.0613	0.1120	0.3370	0.0904	0.1316
	Std. Error	0.0078	0.0069	0.0008	0.0305	0.0105	0.0025
CHV	Coeff.	0.2418	0.1043	0.1192	0.4991	0.1282	0.1408
	Std. Error	0.0157	0.0129	0.0017	0.0257	0.0162	0.0027
DD	Coeff.	0.3141	0.0733	0.1302	0.6349	0.0793	0.1615
	Std. Error	0.0117	0.0096	0.0013	0.0174	0.0113	0.0022
DOW	Coeff.	0.2935	0.0843	0.1371	0.5631	0.0971	0.1636
	Std. Error	0.0131	0.0111	0.0015	0.0195	0.0128	0.0024
EK	Coeff.	0.1266	0.0602	0.1227	0.4272	0.0788	0.1469
	Std. Error	0.0095	0.0077	0.0010	0.0239	0.0109	0.0023
GE	Coeff.	0.2835	0.0940	0.1185	0.5516	0.1074	0.1413
	Std. Error	0.0145	0.0119	0.0015	0.0221	0.0141	0.0023
GM	Coeff.	0.1143	0.1217	0.1048	0.3992	0.1723	0.1246
	Std. Error	0.0110	0.0099	0.0011	0.0279	0.0146	0.0023
IBM	Coeff.	0.5242	0.0715	0.1057	0.8001	0.0732	0.1305
	Std. Error	0.0197	0.0165	0.0019	0.0234	0.0174	0.0027
IP	Coeff.	0.3136	0.0731	0.1465	0.4860	0.0851	0.1639
	Std. Error	0.0184	0.0148	0.0024	0.0246	0.0173	0.0031
JNJ	Coeff.	0.1976	0.1170	0.1235	0.4123	0.1382	0.1409
	Std. Error	0.0171	0.0140	0.0017	0.0289	0.0171	0.0028
KO	Coeff.	0.0945	0.1010	0.1091	0.3068	0.1400	0.1236
	Std. Error	0.0132	0.0113	0.0012	0.0320	0.0156	0.0024
MMM	Coeff.	0.3081	0.1271	0.1361	0.5136	0.1527	0.1562
	Std. Error	0.0197	0.0163	0.0025	0.0270	0.0194	0.0034
MO	Coeff.	0.0985	0.1291	0.1094	0.2253	0.1659	0.1178
	Std. Error	0.0169	0.0136	0.0015	0.0321	0.0179	0.0024
MOB	Coeff.	0.1923	0.1057	0.1342	0.4013	0.1325	0.1528
	Std. Error	0.0131	0.0105	0.0015	0.0222	0.0134	0.0023
MRK	Coeff.	0.3113	0.1769	0.1254	0.5538	0.2146	0.1480
	Std. Error	0.0197	0.0157	0.0022	0.0280	0.0191	0.0032
PG	Coeff.	0.3113	0.0688	0.1505	0.5661	0.0784	0.1779
	Std. Error	0.0148	0.0122	0.0019	0.0212	0.0140	0.0029
S	Coeff.	0.1550	0.0824	0.1126	0.4945	0.1167	0.1394
	Std. Error	0.0109	0.0097	0.0012	0.0250	0.0139	0.0024
T	Coeff.	0.0332	0.0719	0.1084	0.2085	0.1145	0.1198
	Std. Error	0.0077	0.0068	0.0007	0.0340	0.0112	0.0023
XON	Coeff.	0.1154	0.1047	0.1039	0.3907	0.1379	0.1227
	Std. Error	0.0121	0.0104	0.0011	0.0303	0.0141	0.0024
AVG	Coeff.	0.2151	0.0962	0.1216	0.4616	0.1212	0.1423
	Std. Error	0.0140	0.0117	0.0015	0.0261	0.0148	0.0026

The model estimated is Equation (31). An index of market buying and selling pressure is used to estimate the traded spread and its components. S is the estimated traded spread, α is the estimated advserse selection component of the traded spread, and β is the estimated inventory holding component of the traded spread. The sample consists of 5-minute data. Panel A results are based on all the observations, whereas Panel B estimates are based on a data set that collasps all sequential trades observed at the same price and the same quotes to one bunched trade. The last category denoted as AVG presents the average statistics for all stocks.

the inventory effect. Undoubtedly our cross-section approach can be refined by choosing particular portfolios, such as that of a specialist, and by other methods.[18]

[18] For recent articles on the behavior of specialist inventories see Hasbrouck and Sofianos (1993), Madhavan and Smidt (1991, 1993), Madhavan and Sofianos (1994), and Sofianos (1995).

We also employed an econometric refinement that accounts for correlations across securities. Because all stocks respond to marketwide public information, the public information shocks embedded in e_t^k could be contemporaneously correlated across securities. Thus estimation of Equation (31) is more efficient with a simultaneous modeling of n securities. A joint model of Equation (31) can be expressed as

$$r_t = Sq_t + Aq_{t-1} + BQ_{t-1}^* + v_t, \tag{32}$$

where r_t is an $(n \times 1)$ vector with typical element ΔP_t^k, S is an $(n \times n)$ diagonal matrix with typical diagonal element $S^k/2$, q_t is an $(n \times 1)$ vector with typical element Q_t^k, A is an $(n \times n)$ diagonal matrix with typical element $(\alpha^k - 1)(S^k/2)$, B is an $(n \times n)$ diagonal matrix with typical element $\beta^k(S^k/2)$, and v_t is an $(n \times 1)$ vector with typical element e_t^k. Joint estimation of the 19 equations reduces the number of observations to 2532 when no exclusion for clustering is made. Despite the big reduction in sample size, these results support the same inferences made for panel A in Table 9 and are not reported. When data are reduced to account for clustering, joint estimation is not possible due to an inadequate number of observations.

7. Conclusions

A basic trade indicator model is constructed to unify existing spread models, which typically decompose the spread into just two components. Empirical estimates of the basic model for 19 large actively traded stocks in 1992 yield estimates of the sum of adverse information and inventory effects that are relatively small proportions of the spread. However, these proportions increase dramatically when the indicator model is modified to account for medium and large trades.

The basic model is extended in two ways to estimate three components of the spread — order processing, adverse information, and inventory — something not successfully accomplished heretofore. Earlier models have usually lumped inventory and order processing together. The models can also be used to provide estimates of the spread at which trades occur. The first extension relies on the fact that inventory effects induce negative serial correlation in orders and in quotes in addition to the serial correlation from the bid-ask bounce of prices. The second extension takes a cross-section approach and uses information on trading pressures in other stocks to infer the inventory component of the spread in a particular stock. We show that this approach can also be used to separate the three components of the spread. The basis for this approach is that quotes are adjusted in stock A in response to trades in other stocks in order to hedge inventory risk, but trades in other stocks convey little information about stock A.

Most of our empirical evidence is for the first extension. We estimate this extended model with 1992 trade and quote data for 19 large and actively traded stocks. Because trades are not matched one-for-one with orders, we estimate the model with and without adjusting for clustering of trades. The adjustment groups all trade at the same price and in the absence of quote changes into a single order. The adjustment yields an upper bound on the adverse information component of the spread because it increases the frequency with which quotes respond to trades. Under the adjusted results, the average order processing component of the traded spread is 61.8%, the average adverse information component is 9.6%, and the average inventory cost component is 28.7%.

The extended model is also estimated for different sequences of trade size that yield interesting new results. First, the quote adjustment to a trade is larger the larger the trade. Second, the quoted spread tends to be larger when large trades take place. This suggests that the large trade is anticipated through leakage in the upstairs market. Third, the adverse information component of the spread is smaller for large trades than it is for medium and small trades. It appears that large trades are prenegotiated in such a way that the trade price fully reflects the information conveyed by the trade. Fourth, the midpoint reaction and the spread components vary on the basis of prior trades. For example, the adverse information effect is larger for a medium trade at the ask if the prior trade was at the ask and if the prior trade was large.

Appendix A: The Glosten and Harris (1988) Model

Using our timing convention, Glosten and Harris (GH) model the "true" price as

$$V_{t+1} = V_t + \alpha \frac{S}{2} Q_t + e_t \tag{A1}$$

where V_{t+1}, the true price immediately after the trade at t, contains the public information e_t plus the information conveyed by the trade at t. In our corresponding Equation (1), the true price includes public information up to the point at which the true price is evaluated. Because the timing is immediate, the two timing conventions are equivalent. They model the price process in relation to the true price as

$$P_t = V_{t+1} + (1 - \alpha) \frac{S}{2} Q_t \tag{A2}$$

which corresponds to our Equation (4). Our equation contains an error term because we impose a constant spread assumption, whereas they do not.

Take first differences of Equation (A2) and combine it with Equation (A1) to obtain

$$\Delta P_t = \alpha \frac{S}{2} Q_t + (1-\alpha)\frac{S}{2}\Delta Q_t + e_t, \tag{A3}$$

which is analogous to GH's model [Equation (12)] given in the text. This can be rearranged to give

$$\Delta P_t = \frac{S}{2}\Delta Q_t + \alpha \frac{S}{2} Q_{t-1} + e_t, \tag{A4}$$

which is the same as Equation (5) under the assumption that $\beta = 0$.

Appendix B: Basic Model with Trade Size Categories

Denote a small trade as s, a medium trade as m, and a large trade as l. Next define

$$\begin{array}{lll} D_t^s & = Q_t & \text{if share volume at } t \leq 1000 \text{ shares} \\ & = \ 0 & \text{otherwise} \\ D_t^m & = Q_t & 1000 \text{ shares} < \text{if share volume at } t < 10,000 \text{ shares} \\ & = \ 0 & \text{otherwise} \\ D_t^l & = Q_t & \text{if share volume at } t \geq 10,000 \text{ shares} \\ & = \ 0 & \text{otherwise.} \end{array}$$

The corresponding equations to Equations (1) and (2) in the text are then, respectively,

$$V_t = V_{t-1} + \alpha^s \frac{S^s}{2} D_{t-1}^s + \alpha^m \frac{S^m}{2} D_{t-1}^m + \alpha^l \frac{S^l}{2} D_{t-1}^l + \varepsilon_t, \tag{B1}$$

$$M_t = V_t + \sum_j \left[\beta^j \frac{S^j}{2} \sum_{i=1}^{t-1} D_i^j \right], \tag{B2}$$

for $j \in \{s, m, l\}$ for the three size categories. The first difference of Equation (B2) is

$$\Delta M_t = \Delta V_t + \beta^s \frac{S^s}{2} D_{t-1}^s + \beta^m \frac{S^m}{2} D_{t-1}^m + \beta^l \frac{S^l}{2} D_{t-1}^l. \tag{B3}$$

Combining Equations (B1) and (B3) yields the counterpart to Equation (3),

$$\Delta M_t = (\alpha^s + \beta^s)\frac{S^s}{2}D^s_{t-1} + (\alpha^m + \beta^m)\frac{S^m}{2}D^m_{t-1} + (\alpha^l + \beta^l)\frac{S^l}{2}D^l_{t-1} + \varepsilon_t. \tag{B4}$$

The counterpart to Equation (4) is

$$P_t = M_t + \frac{S^s}{2}D^s_t + \frac{S^m}{2}D^m_t + \frac{S^l}{2}D^l_t + \eta_t. \tag{B5}$$

Combining Equations (B4) and (B5) then yields our basic model with different trade size categories:

$$\begin{aligned}\Delta P_t \; = \; & \frac{S^s}{2}D^s_t + (\lambda^s - 1)\frac{S^s}{2}D^s_{t-1} + \frac{S^m}{2}D^m_t + (\lambda^m - 1)\frac{S^m}{2}D^m_{t-1} \\ & + \frac{S^l}{2}D^l_t + (\lambda^l - 1)\frac{S^l}{2}D^l_{t-1} + e_t,\end{aligned} \tag{B6}$$

where $\lambda^j = \alpha^j + \beta^j$ for $j = s, m, l$.

Appendix C: Sample of Securities

Company Name	Symbol
American Express	AXP
Chevron	CHV
Du Pont	DD
Dow Chemical	DOW
Eastman Kodak	EK
General Electric	GE
General Motors	GM
IBM	IBM
International Paper	IP
Johnson & Johnson	JNJ
Coca Cola	KO
3M	MMM
Philip Morris	MO
Mobil	MOB
Merck	MRK
Procter & Gamble	PG
Sears Roebuck	S
AT&T	T
USX	X
Exxon	XON

References

Affleck-Graves, J., S. Hegde, and R. Miller, 1994, "Trading Mechanisms and the Components of the Bid-Ask Spread," *Journal of Finance*, 49, 1471–1488.

Amihud, Y., and H. Mendelson, 1980, "Dealership Market: Market Making with Inventory," *Journal of Financial Economics*, 8, 31–53.

Copeland, T. C., and D. Galai, 1983, "Information Effects on the Bid-Ask Spread," *Journal of Finance*, 38, 1457–1469.

Choi, J. Y., D. Salandro, and K. Shastri, 1988, "On the Estimation of Bid-Ask Spreads: Theory and Evidence," *Journal of Financial and Quantitative Analysis*, 23, 219–230.

Demsetz, H., 1968, "The Cost of Transacting," *Quarterly Journal of Economics*, 82, 33–53.

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

Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman, 1996, "Liquidity, Information, and Infrequently Traded Stocks," *Journal of Finance*, 51, 1405–1436.

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

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

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

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Harris, L., 1990, "Statistical Properties of the Roll Serial Covariance Bid/Ask Spread Estimator," *Journal of Finance*, 45, 579–590.

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

Hasbrouck, J., 1991, "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179–207.

Hasbrouck, J., and G. Sofianos, 1993, "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists," *Journal of Finance*, 48, 1565–1593.

Hausman, J., A. W. Lo, and A. C. MacKinlay, 1992, "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics*, 31, 319–380.

Ho, T., and H. R. Stoll, 1981, "Optimal Dealer Pricing under Transactions and Return Uncertainty," *Journal of Financial Economics*, 9, 47–73.

Ho, T., and H. R. Stoll, 1983, "The Dynamics of Dealer Markets under Competition," *Journal of Finance*, 38, 1053–1074.

Holthausen, R. W., R. W. Leftwich, and D. Mayers, 1987, "The Effect of Large Block Transactions on Security Prices: A Cross-Sectional Analysis," *Journal of Financial Economics*, 19, 237–267.

Huang, R. D., and H. R. Stoll, 1994, "Market Microstructure and Stock Return Predictions," *Review of Financial Studies*, 7, 179–213.

Huang, R. D., and H. R. Stoll, 1996a, "Competitive Trading of NYSE Listed Stocks: Measurement and Interpretation of Trading Costs," *Financial Markets, Institutions & Instruments*, 5, 1–55.

Huang, R. D., and H. R. Stoll, 1996b, "Dealer versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE," *Journal of Financial Economics*, 41, 313–357.

Jegadeesh, N., and S. Titman, 1995, "Short-Horizon Return Reversals and the Bid-Ask Spread," *Journal of Financial Intermediation*, 4, 116–132.

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

Jones, C., and M. Lipson, 1995, "Continuations, Reversals, and Adverse Selection on the Nasdaq and NYSE/AMEX," Financial Research Center Memorandum No. 152, Department of Economics, Princeton University.

Kraus, A., and H. R. Stoll, 1972, "Price Impacts of Block Trading on the New York Stock Exchange," *Journal of Finance*, 27, 569–588.

Lin, J.-C., G. Sanger, and G. G. Booth, 1995a, "External Information Costs and the Adverse Selection Problem: A Comparison of Nasdaq and NYSE Stocks," working paper, Louisiana State University.

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

Madhaven, A., and M. Cheng, 1996, "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets," *Review of Financial Studies*, 10, 175–203.

Madhavan, A., M. Richardson, and M. Roomans, 1996, "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks," working paper, University of Southern California.

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

Madhavan, A., and S. Smidt, 1993, "An Analysis of Changes in Specialist Quotes and Inventories," *Journal of Finance*, 48, 1595–1628.

Madhavan, A., and G. Sofianos, 1994, "Auction and Dealer Markets: An Empirical Analysis of NYSE Specialist Trading," Working Paper 94-1, New York Stock Exchange.

Neal, R., and S. Wheatley, 1994, "How Reliable are Adverse Selection Models of the Bid-Ask Spread?" working paper, Federal Research Bank of Kansas City.

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

Roll, R., 1984, "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127–1139.

Sofianos, G., 1995, "Specialist Gross Trading Revenues at the New York Stock Exchange," Working Paper 95-1, New York Stock Exchange.

Stoll, H. R., 1978, "The Supply of Dealer Services in Securities Markets," *Journal of Finance*, 33, 1133–1151.

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