

Price effects of trading and components of the bid-ask spread on the Paris Bourse

Frank de Jong^{a,*}, Theo Nijman^a, Ailsa Röell^b

^a *Department of Econometrics, Center for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands*

^b *ECARE, Brussels, Belgium and Tilburg University, Tilburg, The Netherlands*

Abstract

In this paper we estimate price effects of trading on the Paris Bourse. We use several methods to decompose price effects into transitory and permanent parts. First, we use the Glosten (1994) model for an electronic order-driven market. In line with theoretical predictions, the price impact increases with trade size, and is estimated between 25% (for small transactions) and 60% (for large transactions) of the total spread. We then use a reduced form approach based on a multi-period Vector Auto Regression. The VAR estimates of the permanent price impact are between 40% and 115% of the spread, much larger than the estimates of Glosten's one-period model. We check the robustness of these results by less restrictive, direct estimates of long-run price effects and confirm the results of the VAR analysis. We separately analyze the price effects of off-exchange transactions. These appear to have no long run price impact at all. In all results, there is no reversal of the direction of trade, which suggests that inventory control is unlikely to be important on the Paris Bourse.

JEL classification: G14; C32

Keywords: Microstructure; Asymmetric information; VAR; Persistence; Robustness

1. Introduction

There is by now a large literature that analyzes transactions data for financial markets. Among the issues in this literature are the dynamic properties of

* Corresponding author. Fax: +31-13-4663280; E-mail: f.dejong@kub.nl.

transaction prices, and in particular the price effect of trading. Closely related to this work is the issue of estimating the components of the bid-ask spread. Much of the empirical literature in this area is based on theoretical models, although more reduced form approaches are also popular. In this paper, we analyze price effects of trading and the components of the bid-ask spread on a new data set from the Paris Bourse. The Bourse operates a fully automated order processing system (the CAC system) based on an electronic limit order book. This trading system gives very high quality limit order and transactions data.

We take several different approaches to estimating price effects of trading. The two main models that we use are the Glosten (1994) model and the VAR model pioneered in market microstructure work by Hasbrouck, 1991a, Hasbrouck, 1991b, Hasbrouck, 1993. The former model is closely linked to market microstructure theory, whereas the latter adopts a more agnostic, reduced form approach. We show that these models yield different price effect estimates. Moreover, we show that there is an important difference between the effects of trading on mid-quotes and on transaction prices.

Our paper is, to our knowledge, the first that empirically implements the Glosten (1994) model of liquidity provision via a public limit order book. Here each order that is initiated by an impatient trader can effectively be executed against a series of limit orders at different prices. We extend Glosten's model to incorporate the possibility that limit order providers may require compensation for order processing cost and/or oligopoly rents. We derive an explicit decomposition of bid-ask spreads faced by impatient traders into adverse selection and other costs incurred by the providers of liquidity.

A further methodological innovation in this paper is a test of the robustness of the VAR approach. We propose a direct estimate of expected price changes due to a transaction, and compare these estimates to the price effects implied by the VAR model. We show that most conclusions of the VAR model stand up against this test. The robust approach also allows us to analyze the price effects of off-exchange transactions separately.

The paper is organized as follows. In Section 2 we discuss the institutional structure of the Bourse and the data. In Section 3 we analyze the price effects of trading using structural models of transaction prices. In Section 4 we apply the reduced form methodology developed by Hasbrouck, 1991a, Hasbrouck, 1991b, Hasbrouck, 1993 to estimate the price effects of trading. In Section 5 we check the robustness of the results. In Section 6 we conclude.

2. Institutional background and data

In this section we briefly describe the trading structure on the Paris Bourse and the data. The Bourse is a continuous auction market that uses a centralized electronic system for displaying and processing orders, the Cotation Assistée en

Continu (CAC) system. This system, based on the Toronto Stock Exchange's CATS (Computer Assisted Trading System), was first implemented in Paris in 1986. Since then, trading in nearly all securities has been transferred from the floor of the exchange onto the CAC system. All the most actively traded French equities are traded on a monthly settlement basis in round lots of 5 to 100 shares set by the Société des Bourses Françaises (SBF) to reflect their unit price. The SBF itself acts as a clearing house for buyers and sellers, providing guarantees against counterparty default. There are no specialists or professional market makers. Instead, liquidity is provided by the public limit order book.

Every morning at 10 a.m. the trading day opens with a batch auction where all eligible orders are filled at a common market clearing price. Nowadays the batch auction is relatively unimportant, accounting for no more than 10 to 15% of trading volume. Its role is to establish an equilibrium price before continuous trading starts. Continuous trading takes place from 10 a.m. to 5 p.m.

In the continuous trading session there are two types of orders possible, limit orders and market orders. Limit orders specify the quantity to be bought or sold, a required price and a date for automatic withdrawal if not executed by then, unless the limit order is good till cancelled (“à révocation”). Limit orders cannot be issued at arbitrary prices because there is a minimum “tick” size of FF 0.1 for stock prices below FF 500, and FF 1 for higher prices. More than one limit order may be issued at the same price. Strict time priority applies in the execution of such orders.

Market orders only specify the quantity to be traded and are executed immediately “au prix du marché”, i.e., at the best price available. If the total quantity of the limit orders at this best price do not suffice to fill the whole market order, the remaining part of the market order is transformed into a limit order at the transaction price (for a detailed description of this system see Biais et al., 1995). Hence, market orders do not automatically walk up the limit order book, and do not always provide immediate execution of the whole order.¹

After the opening, traders linked up to the CAC system will see an onscreen display of the “market by price”. For both the bid side and the ask side of the market, the five best limit order prices are displayed together with the quantity of shares available at that price and the number of individual orders involved. The difference between the best bid and ask price is known as the “fourchette”. Brokers can scroll down to further pages of the screen to view limit orders available beyond the five best prices. In addition, some information concerning the

¹ A trader who wants to trade a certain quantity immediately can circumvent this mechanism by placing a *limit order* at a price that is very unfavorable to him. This limit order will then be executed against existing orders on the other side of the market that show a more favorable price. In the data, these show up as a series of transactions on the same side of the market with the same time stamp.

recent history of trading is given: time, price, quantity and buyer and seller identification codes for the five last transactions, the cumulative quantity and value of all transactions since the opening, and the price change from the previous day's close to the latest transaction.

The member firms of the Bourse (the "Sociétés de Bourse") key orders directly into the CAC system via a local terminal. All market participants can contribute to liquidity by putting limit orders on display. There is some scope for negotiated deals if the limit order book is insufficiently deep. A financial intermediary can negotiate a deal directly with a client at a price lying within the current fourchette, provided that the deal is immediately reported to the CAC system as a "cross order". For trades at prices outside the fourchette, the member firm acting as a principal is obliged to fill all central market limit orders displaying a better price than the negotiated price within five minutes.

Our data set is a transcription of all changes in the trading screen information for all shares on the CAC system for 44 trading days in the summer of 1991, starting May 25 and ending July 25. We have a complete record of the total limit order quantity at the five best prices on both the bid side and the ask side of the market and all transactions. Due to the automated trading system, the data are relatively clean. The time stamps indicate exactly the time of the transaction or quote change. Also, quote and trade information is in correct sequence, so that it is possible to find out exactly who initiated the trade, the buyer or the seller, by comparing the transaction price with the previous mid-quote, i.e., the average of best buy and sell limit order prices.

Concerning transactions, there is an indicator showing whether the transaction is a "cross" negotiated outside the CAC system. We also have broker identification codes for the buying and selling parties, which allow us to identify series of small transactions that were initiated by the same person as part of one large transaction. The transaction price per share for such transactions is defined as the quantity weighted average of the price of the small transactions that together make up the larger one.

In this paper, we use a sample of transactions in the shares of ten large firms, with 4000 to 11,000 observations per stock. Most transactions involve a limit order, so that it can be determined exactly who initiated the trade, the buyer or the seller. For the "crosses", which are negotiated outside the CAC system, we use the simple rule that transactions with a price above the mid-quote² are deemed to be buyer initiated, and below the mid-quote seller initiated. Transactions exactly at the mid-quote are not classified. The size of the transaction is normalized to the number of shares traded divided by the so-called Normal Market Size (NMS), which is a transaction size set by the authorities on London's SEAQ International market, and roughly corresponds to the median transaction size in that stock. For

² The mid-quote is defined as the average between the best bid and the best ask price available.

the Paris market, 1 NMS corresponds to the 99th percentile of the transaction size distribution, see De Jong et al. (1995).

French shares are also traded on other exchanges, especially London's SEAQ International. It would appear natural to analyze transactions data from both markets simultaneously. There are good reasons not to include London data into the analysis, however. Transactions in London are negotiated between traders and market makers by telephone and are not made public, so that other traders cannot see that they have taken place. The trading process in London is not very visible to outsiders; only the market makers' quotes are publicly observable, but these are often adjusted slowly and do not give a good indication of actual transaction prices. Therefore, we decided to analyze only the transactions from the Paris Bourse.

3. Estimating price effects by a structural market microstructure model

As described before, the CAC system operating on the Paris Bourse is close to an ideal electronic open limit order book system. Glosten (1994) presents a theoretical model of prices and price revisions due to trading in such a market. In this section, we develop an empirical implementation of Glosten's model and we estimate this model on the data from the Paris Bourse.

Our starting point is the original Glosten (1994) model where there are no explicit order processing costs. To simplify the exposition we discuss only buyer initiated transactions. As in Glosten (1994), let $R(q)$, the "revenue" function, denote the total payment made for his order by a buyer who initiates an order of size q , over and above the ex ante expected value of his order (that is the quantity q times y , the best public estimate of the stock before it is known that the order is forthcoming). The marginal price of a transaction of size q is determined by the following rule:

$$R'(q) = E_z[e(z) | z \geq q], \quad (1)$$

where $e(z)$ is the revision of the best public estimate of the security's value when it becomes known that a buyer-initiated order has arrived on the market, and that it is of size z . E_z denotes the expectation taken with respect to the transaction size distribution. We assume that this distribution is exponential, so that

$$F_z(q) = 1 - e^{-\frac{q}{\alpha}}. \quad (2)$$

The price revision describes the change in expectations of the true value of the stock due to a transaction of size q . For simplicity, we assume that this schedule is linear:

$$e(q) = e_0 + e_1 q. \quad (3)$$

Under these assumptions the marginal price schedule is

$$R'(q) = e_0 + e_1 E_z[z | z \geq q] = e_0 + e_1(q + \alpha), \quad (4)$$

where the latter equality follows from the properties of the exponential distribution. As an extension of Glosten's model, we introduce an order processing cost component in the marginal revenue schedule (Glosten, 1994). Let the order processing cost function be denoted by $C(q)$, then

$$R'(q) = C'(q) + e_0 + e_1(q + \alpha). \quad (5)$$

Our data do not concern marginal prices but rather average prices. Integrating (5) and dividing by q one obtains the average price schedule:

$$\frac{R(q)}{q} = \frac{C(q)}{q} + (e_0 + e_1\alpha) + \frac{1}{2}e_1q. \quad (6)$$

For simplicity, we assume that the average order processing cost is a linear function of q .³ Hence,

$$\frac{R(q)}{q} = c_0 + c_1q + (e_0 + e_1\alpha) + \frac{1}{2}e_1q \equiv R_0 + R_1q, \quad (7)$$

where $R_0 \equiv c_0 + e_0 + e_1\alpha$ and $R_1 \equiv c_1 + 1/2e_1$. In summary, in the Glosten model we have the following decomposition of the bid-ask spread:

$$\text{Adverse selection cost: } (e_0 + e_1\alpha) + \frac{1}{2}e_1q$$

$$\text{Order processing cost: } c_0 + c_1q$$

For the empirical implementation of this model we introduce the following notation:

p_t = logarithm of the transaction price (average price paid per share),

q_t = quantity (number of shares traded),

Q_t = "sign" of the transaction,⁴

y_t = expected value of the stock *before* the transaction,

ϵ_t = publicly observed change in the value of the stock,

where the time index counts transactions (i.e., the model will be specified in transaction time). The first equation of the empirical model states that the transaction price is equal to the expected value of the stock before the transaction, plus the average price premium, $R(q)/q$, given by Eq. (7). As in Madhavan et al. (1994) we add a random pricing error, u_t , that captures other influences on the

³ Note that this implies a quadratic cost function with zero intercept.

⁴ The *sign* is defined as +1 if the transaction is initiated by the buyer (at the ask) and -1 if the transaction is initiated by the seller (at the bid).

transaction price, such as price discreteness and other factors that are not modeled. We assume that u_t is uncorrelated with the other variables in the price equation. Thus, the pricing equation becomes

$$p_t = y_t + (R_0 + R_1 q_t) Q_t + u_t. \quad (8)$$

The price revision can be modeled by the change in the expected value y_t , as follows:

$$y_{t+1} = y_t + (e_0 + e_1 q_t) Q_t + \epsilon_{t+1}, \quad (9)$$

where ϵ_{t+1} is publicly observed information that comes in between transaction t and $t + 1$, but is unrelated to the current transaction. Substitution of (8) into (9) yields the following equation for observed transaction price changes

$$\Delta p_{t+1} = \Delta(R_0 + R_1 q_{t+1}) Q_{t+1} + (e_0 + e_1 q_t) Q_t + e_{t+1}, \quad (10)$$

where $e_t = \epsilon_t + \Delta u_t$. The interpretation of this equation is simple: the coefficients of the “difference” variables are the intercept and slope of the average price, whereas the coefficients of the levels one period lagged are estimates of the intercept and slope of the price revision schedule.

The equation to be estimated is found by rewriting (10) slightly:

$$\Delta p_{t+1} = c + R_0 \Delta Q_{t+1} + R_1 \Delta(q_{t+1} Q_{t+1}) + e_0 Q_t + e_1 q_t Q_t + e_{t+1}. \quad (11)$$

A number of econometric issues concerning the estimation of this equation require special attention. Following Harris (1986) and Hasbrouck (1991b), who argue that observed covariance patterns in transaction returns are more consistent with transaction time than with calendar time, we assume that the relevant “clock” is transaction time. We include a constant term, c , in the model to capture the average return between transactions (i.e., a non-zero mean of ϵ_t). The variance of the errors is unspecified by the model. For several reasons, it is likely that the errors are heteroskedastic. For example, the variance may depend on the time of day, and the variance may depend on the trade size. If the pricing equation (Eq. (8)) is not exact, the regression error has an MA(1) serial correlation pattern. With this error structure, OLS gives consistent point estimates, but the usual standard error formula is incorrect. We compute heteroskedasticity and autocorrelation consistent (HAC) standard errors by the method of Newey and West (1987). The size of the transaction is censored at 2 times NMS, which corresponds roughly to the 99.5 percentile of the size distribution, so that approximately 0.5% of the transactions are affected. The reason for censoring is to mitigate the effect of very large trades on the estimates, see also Hausman et al. (1992). In the estimation, the dependent variable is the change in the logarithm of transaction prices, so that the parameter estimates can be interpreted as relative price effects. Overnight returns and opening prices are excluded from the sample. It is important to note that cross transactions are excluded from the analysis. These will be studied separately in

Table 1

Estimates of the Glosten (1994) model

$$\Delta p_t = c + R_0 \Delta Q_{t+1} + R_1 \Delta(Q_{t+1} q_{t+1}) + e_0 Q_t + e_1 Q_t q_t + e_t$$

$$c_0 = R_0 - e_0 - e_1 \alpha$$

$$c_1 = R_1 - 1/2 e_1$$

	R_0	R_1	c_0	c_1	e_0	e_1	α	ρ
Accor	10.21 (38.54)	1.50 (1.25)	7.31 (29.80)	-1.09 (1.09)	2.34 (10.56)	5.17 (4.29)	0.108	-0.23
Elf-Aquitaine	7.73 (49.35)	0.50 (1.11)	4.90 (33.81)	-1.13 (2.80)	2.36 (18.21)	3.25 (6.69)	0.144	-0.21
BSN	8.05 (71.86)	1.11 (1.39)	6.19 (54.71)	-1.38 (1.87)	1.67 (17.60)	4.89 (6.00)	0.040	-0.27
Carrefour	8.68 (46.33)	0.69 (1.56)	6.46 (36.18)	-0.73 (2.06)	1.81 (12.90)	2.84 (4.83)	0.144	-0.22
Axa-Midi	14.93 (29.72)	5.04 (2.51)	10.34 (19.81)	1.21 (0.79)	3.93 (10.16)	7.67 (3.07)	0.087	-0.13
Generale des Eaux	5.79 (33.46)	0.65 (1.85)	4.09 (31.33)	0.33 (1.46)	1.42 (5.90)	1.95 (4.02)	0.144	-0.08
l'Oreal	13.52 (50.86)	0.75 (0.69)	9.36 (35.81)	-1.56 (1.60)	3.87 (17.10)	4.65 (3.42)	0.063	-0.20
Ricard	15.14 (27.99)	0.24 (0.09)	10.28 (20.25)	-2.77 (1.48)	4.21 (9.76)	6.01 (2.07)	0.108	-0.23
Schneider	15.43 (36.76)	-0.68 (0.33)	10.66 (26.42)	-3.09 (1.88)	4.42 (13.98)	4.83 (2.16)	0.072	-0.22
UAP	18.91 (35.26)	-3.44 (2.76)	12.34 (25.11)	-3.57 (3.59)	6.52 (14.51)	0.27 (0.18)	0.180	-0.18

This table reports the estimated coefficients of the Glosten (1994) model, Eq. (11) with Δp in one-hundredths of a percent, and quantities in NMS.

The transaction size is truncated at 2 NMS; crosses are excluded from the sample.

α is estimated by the median of the transaction size distribution, divided by $\ln 2$.

ρ is the first order autocorrelation coefficient of the error term.

Newey–West t -statistics in parenthesis.

Section 5. Hence, the results in this section (and the next) concern the price dynamics of transactions executed entirely through the CAC limit order book.

In Table 1 the estimates of Eq. (11) are presented. The table presents estimates of R_i and e_i and the implied values of $c_i (i = 0, 1)$, from Eq. (7). The coefficient α is estimated by the median of the uncensored, non-cross transaction size distribution, divided by $\ln 2$.⁵ The interpretation of the coefficients is as follows. The bid-ask spread for a transaction of size q is twice $R_0 + R_1 q$. In the Glosten model, as given by Eq. (7), the order processing cost is equal to $c_0 + c_1 q$. The remainder, $(R_0 - c_0) + (R_1 - c_1)q$ is the adverse selection cost.⁶ Calculations using the

⁵ This follows from $F_z(q_{med}) = 1 - e^{-q_{med}/\alpha} = \frac{1}{2}$, where q_{med} is the median of q .

⁶ Substituting the expressions for R_0 and R_1 from Eq. (7), we find that the adverse selection component equals $(e_0 + e_1 \alpha) + 1/2 e_1 q$.

estimated coefficients in Table 1 show that the order processing cost for a small transaction ($q = 0$) accounts for 70% of the spread, whereas the remaining 30% is due to asymmetric information. For a transaction of the Normal Market Size ($q = 1$) the order processing cost accounts for about 55% of the spread, and the adverse selection component is 45%. These estimates are comparable to the estimates obtained in other research using US data, e.g., Stoll (1989) and Madhavan and Smidt (1991).

It is interesting to see how the interpretation of the coefficients of Eq. (11) would differ in the Glosten–Harris model, which presumes that competitive dealers quote a single price directly for an entire order (Glosten and Harris, 1988). In that model the adverse selection cost equals the price revision, $e_0 + e_1q$, and the order processing cost equals $(R_0 - e_0) + (R_1 - e_1)q$. Compared with the interpretation based on Glosten (1994), the Glosten–Harris model attributes a share to adverse selection that is greater by $1/2 e_1q - e_1\alpha$, which is negative if $q < 2\alpha$, and positive if $q > 2\alpha$. Thus, given that α is the mean trade size, a decomposition along the lines suggested by Glosten and Harris would underestimate (overestimate) adverse selection costs (order processing costs) for trade sizes under twice the mean, and vice versa for trade sizes above twice the mean. Using the estimated coefficients in Table 1, the Glosten–Harris model implies adverse selection costs ranging from 25% for $q = 0$ to 60% for $q = 1$. Hence, relative to the more appropriate Glosten (1994) model, the Glosten–Harris model gives a slight underestimate of the adverse selection component for small transactions and an overestimate of the asymmetric information component of about 15% for a large transaction.

To summarize the results of this section, we may conclude that the data support the predictions of Glosten's model for an electronic market with an open limit order book (Glosten, 1994). The adverse selection cost component of the bid-ask spread ranges from 30% for small transactions to 45% for large transactions. As predicted by the model, the price revisions after large transactions are substantially bigger than the revisions after small transactions.

4. Simultaneous analysis of prices and transactions

The structural model of the previous section is elegant, but has some disadvantages. First, the estimates assume that the model is correctly specified. For example, it is assumed that all asymmetric information is revealed immediately after the transaction so that there is only an immediate price effect of trading and no lagged effects. Second, it assumes that the pattern of trading is exogenous. If the trading pattern is not exogenous, the regression coefficients might be biased because some relevant lagged trade variables are omitted. The Vector Auto Regressive (VAR) model introduced into the market microstructure literature by Hasbrouck, 1991a, Hasbrouck, 1991b, Hasbrouck, 1993 explicitly takes these

considerations into account. In the VAR, the joint price and trade dynamics are modeled by the following system of equations:

$$\begin{pmatrix} 1 & -b_0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \Delta p_t \\ x_t \end{pmatrix} = \begin{pmatrix} a(L) & b(L) \\ c(L) & d(L) \end{pmatrix} \begin{pmatrix} \Delta p_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}, \quad V \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \Omega \end{pmatrix}, \quad (12)$$

where Δp_t denotes the price change and x_t is a vector of explanatory variables and $a(L)$ to $d(L)$ are polynomials in the lag operator. In our analysis, the vector of explanatory variables includes the trade sign (Q_t) and size ($z_t \equiv q_t Q_t$). The current vector of explanatory variables, x_t , is included in the equation for Δp_t , so that for identification the error terms e_{1t} and e_{2t} are supposed to be uncorrelated. This model allows a very general dependence of price changes and trade sign and size on the past, without the assumption that the pattern of trade is exogenous.⁷

4.1. Price effects of trading in the VAR model

Although some of the coefficients of the VAR are interesting in themselves, the effects of shocks on future returns and other variables are more interesting for the purposes of this paper. In particular, we are interested in the expected value of the stock price τ periods after a shock, given that the system initially is in the ‘‘steady state’’:⁸

$$pe_1(\tau) \equiv E(p_{t+\tau} - y_t | e_{1t} = 1, e_{2t} = 0, \Delta p_{t-1} = 0, \dots, x_{t-1} = 0, \dots), \quad (13a)$$

$$pe_2(\tau) \equiv E(p_{t+\tau} - y_t | e_{1t} = 0, e_{2t} = 1, \Delta p_{t-1} = 0, \dots, x_{t-1} = 0, \dots). \quad (13b)$$

Sims (1980) popularized the idea of computing such price effects from the *impulse responses* of the VAR model, which can be computed by inverting the VAR to the following Vector Moving Average (VMA) representation:

$$\begin{pmatrix} \Delta p_t \\ x_t \end{pmatrix} = \begin{pmatrix} \alpha(L) & \beta(L) \\ \gamma(L) & \delta(L) \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}. \quad (14)$$

To illustrate the usefulness of the VMA form, consider the equation for the price changes in more detail

$$\Delta p_t = \sum_{k=0}^{\infty} \alpha_k e_{1,t-k} + \sum_{k=0}^{\infty} \beta_k e_{2,t-k}. \quad (15)$$

⁷ Such an assumption was made by Glosten and Harris (1988), Hasbrouck (1988) and Stoll (1989).

⁸ This equation is for the case that x_t and e_{2t} are scalar. In the case of a multi-dimensional trade vector, the impulse responses need to be calculated from a VAR model with orthogonal innovations. In our work, this is achieved by adding the current sign Q_t as an explanatory variable in the equation for the size z_t to obtain orthogonal errors.

In words, the price differences are infinite sums of past innovations in the price equation (e_{1t}) and the transaction equation (e_{2t}). The effect of unit price and trade innovations on the price *change* k periods ahead are measured by α_k and β_k , respectively. Thus, the coefficients of the VMA are exactly the desired impulse responses. The effects of a unit shock on the price *level* τ periods ahead is measured by partial sums of the impulse responses:

$$pe_1(\tau) = \sum_{k=0}^{\tau} \alpha_k, \quad pe_2(\tau) = \sum_{k=0}^{\tau} \beta_k. \tag{16}$$

The long run effects of shocks are easily determined as the limits of the partial sums as $\tau \rightarrow \infty$:

$$pe_1(\infty) = \sum_{k=0}^{\infty} \alpha_k \equiv \alpha(1), \quad pe_2(\infty) = \sum_{k=0}^{\infty} \beta_k \equiv \beta(1). \tag{17}$$

Cochrane (1988) notes that this definition of the long run effects of innovations is unique and independent of any particular decomposition of the price process into permanent and transitory parts.

4.2. Decomposition in transitory and permanent components

For our analysis of price reactions to trades the partial sums of impulse responses provide all necessary information. However, it is also interesting to calculate explicitly the transitory component of the stock price. Given our assumption that the prices and transactions are generated by a bivariate VAR, the natural decomposition of the observed price, p_t , into a random walk component, μ_t , and stationary deviations around the random walk, s_t , is given by Beveridge and Nelson (1981).⁹ The decomposition is as follows:

$$\begin{aligned} p_t &= \mu_t + s_t, \\ \mu_t &= \mu_{t-1} + \sum_{k=0}^{\infty} \alpha_k e_{1t} + \sum_{k=0}^{\infty} \beta_k e_{2t}, \\ s_t &= \sum_{k=0}^{\infty} \tilde{\alpha}_k e_{1,t-k} + \sum_{k=0}^{\infty} \tilde{\beta}_k e_{2,t-k}, \\ \tilde{\alpha}_k &\equiv \sum_{j=k+1}^{\infty} \alpha_j, \quad \tilde{\beta}_k \equiv \sum_{j=k+1}^{\infty} \beta_j. \end{aligned} \tag{18}$$

⁹ Other decompositions of the stock price into permanent and transitory components are also possible. However, these all lead to vector ARMA models for the price-trade process, whereas we assume from the start a VAR model. Hasbrouck (1993) shows that the Beveridge–Nelson decomposition give a lower bound for the variance of the stationary price part among all possible decompositions.

This decomposition is achieved by subtracting the long run price effects of the innovations, given by Eq. (17), from Eq. (15). The natural economic interpretation of the random walk component μ_t is that it is the underlying equilibrium price of the stock, in which all public information is reflected.¹⁰ The stationary part, s_t , measures the deviations of the actual transaction price from the efficient price. Hasbrouck (1993) proposes to use the standard deviation of the stationary part, σ_s , as a summary measure of the quality of a security market. Intuitively, σ_s reflects how closely the transaction price tracks the efficient price on average. This ‘‘dynamic’’ measure of transaction costs can be seen as a generalization of Roll’s estimator (Roll, 1984). Under Roll’s special assumptions, σ_s is equal to half the realized bid-ask spread.

4.3. Econometric aspects

In actual empirical application of the VAR methodology, several econometric points deserve attention. Including the sign of the transaction in a simultaneous dynamic model creates some problems for estimation and computing dynamic effects. Because Q_t is a limited dependent variable that can only take the values -1 and $+1$, the first equation of the VAR cannot be a conditional expectation of Q_t for all values of $\Delta p_{t-i} \in \mathbf{R}$ if the coefficients of Δp_{t-i} are non-zero. However, for moderate values of Δp_{t-i} the linear equation may be a good approximation of the true conditional expectation, and the bias in OLS estimates is probably not too serious. Using Q_t as an explanatory variable in the equation for Δp_t causes no problems, because the errors of the return equation and the other equations of the VAR are uncorrelated, see Heckman (1978). Five lags¹¹ in the VAR are sufficient given the general absence of residual serial correlation in the estimated equations. Overnight returns and opening trades are excluded from the analysis. All reported standard errors are heteroskedasticity consistent estimates. Following the procedure in the previous section, we exclude all ‘‘cross’’ transactions from the analysis. The dynamics of the crosses will be studied in Section 5.

4.4. Empirical results

Table 2 reports two important quantities: the leading coefficients of the VAR and the sums of the moving average coefficients. The first provides an estimate of the half-spread, similar to the R_0 and R_1 coefficients of the Glosten model, (11). The estimates of the spread for small and large trades are very similar to the

¹⁰ Although equilibrium returns are probably correlated over longer horizons, see Conrad and Kaul (1989) and Lo and MacKinlay (1988), for the analysis of transactions data a good working hypothesis is that the efficient price changes are serially uncorrelated.

¹¹ This follows Hasbrouck (1991a).

Table 2
VAR results on transaction prices

Trade size:	Estimated bid-ask spread		Permanent price effect		Ratio of permanent effect to spread	
	Small	NMS	Small	NMS	Small	NMS
Accor	10.15	12.25	4.24	14.31	42	117
Elf-Aquitaine	7.83	8.53	3.95	11.45	50	134
BSN	8.13	10.04	2.07	13.25	25	132
Carrefour	8.92	9.55	3.35	8.89	38	93
Axa-Midi	15.57	21.28	5.91	21.58	38	101
Genarale des Eaux	5.58	6.84	2.34	6.78	40	99
l'Oreal	13.77	14.44	5.34	20.64	39	143
Ricard	15.09	15.96	6.39	15.90	42	100
Schneider	15.53	15.72	6.92	20.32	44	129
UAP	19.08	16.45	10.21	17.26	54	105
Average					41	115

This table shows some results of the VAR model (12) estimated on transaction prices, excluding all cross transactions.

The "spread" column shows the estimated realized half bid-ask spread, in units of one-hundredth of a percent.

The "permanent" column shows the estimated permanent price effect of a transaction, also in units of one-hundredth of a percent.

The "ratio" column reports the permanent price effect, expressed as a percentage of the estimated bid-ask spread.

A small transaction is a hypothetical transaction of size 0, whereas a large transaction is of Normal Market Size.

estimates obtained using the Glosten model.¹² The sum of the impulse responses provide an estimate of the permanent price effect. From the table it is clear that these estimates are much larger than the estimated price revisions in the Glosten model (cf. Table 1). The average ratio of permanent price effect to the half-spread ranges from 40% for small trades and 115% for large trades. These numbers are approximately twice as large as the estimates of the price revisions in Table 1. Thus, the one period model severely underestimates permanent price effects of trading.

In order to gauge how fast transaction prices adjust to new information, we plot the full impulse response function in Fig. 1. This figure graphs the cumulative impulse responses of transaction prices to trading, expressed as a fraction of the estimated spread, averaged over all ten series. For small transactions, the effect of

¹² The effects of a small transaction are equal to the impulse responses of a shock in the sign. The price effects of large transactions were computed by adding the impulse responses of the sign and the size, so these describe the price impact of a transaction of the Normal Market Size.

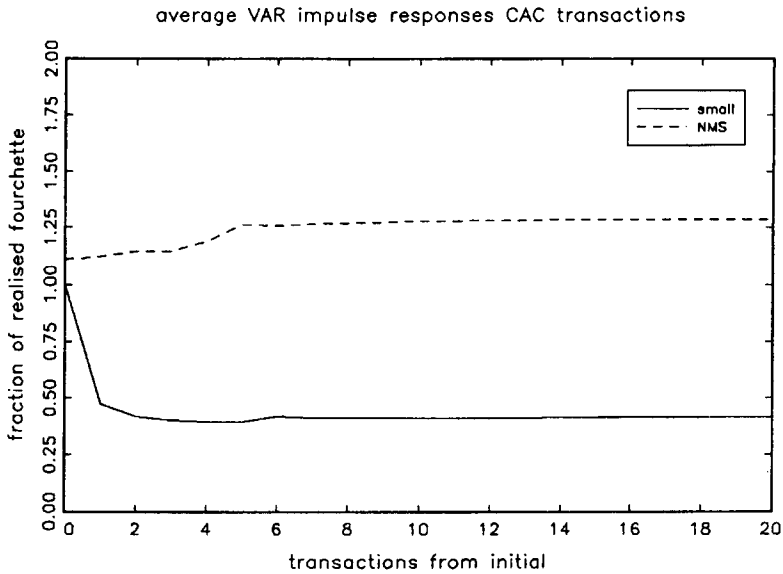


Fig. 1. Impulse responses of transaction prices.

The solid line graphs the effect on transaction prices of a shock in the sign equation, and the dotted line graphs the effect of a shock in both the sign and the size equation.

a trade on subsequent transaction prices is virtually the same for all horizons. For large transactions, the price effect is slightly increasing over the 20 period horizon.

From this pattern, we conclude that new information is reflected virtually immediately in the transaction prices. Also, we can reject the presence of inventory control effects because there are no temporary price changes beyond the bid-ask spread. Moreover, at the most intuitive level, inventory control causes sign “reversals”, i.e., negative serial correlation in the direction of trade initiation. On the contrary, in our data there is strong positive serial correlation in the sign (about 0.3).¹³

Table 3 reports results of the decomposition of the transaction price in a stationary and random walk part. The standard deviation of the stationary part of the price also gives nearly the same estimate of the spread as the leading VAR coefficient. This is not very surprising given the immediate adjustment of prices to

¹³ Other evidence in the literature for the inventory control effect is at best weak. Madhavan and Smidt (1993) and Hasbrouck and Sofianos (1992) estimate inventory control models on samples of data from the NYSE. In particular, they test for mean reversion in the inventory levels of specialists. Both papers are only able to find mean reversion in the inventory levels if they allow for speculative shifts in the desired inventory level. Moreover, the estimated reversion to the desired level is very slow, and takes a number of days. Therefore, for analyzing intra-day price effects, inventory control is perhaps not so relevant.

Table 3
Variance decompositions

	σ_s	σ_w	$R_{w,x}^2$
Accor	10.31	8.11	0.42
Elf-Aquitaine	8.31	7.12	0.48
BSN	7.74	5.03	0.38
Carrefour	9.23	7.14	0.32
Axa-Midi	16.56	12.25	0.38
Generale des Eaux	6.25	4.94	0.40
l'Oreal	13.53	10.55	0.38
Ricard	15.14	11.40	0.37
Schneider	15.22	11.95	0.40
UAP	17.69	14.63	0.47

The σ_s column shows the variance of the stationary part of the transaction price (in units of one-hundredth of a percent).

The σ_w column shows the variance of the innovations in the efficient price.

The $R_{w,x}^2$ column shows the proportion of the variance of the efficient price explained by trading.

their new equilibrium values. Hasbrouck (1991b) proposes to use the proportion of the variance in the random walk part as a summary statistic for the informativeness of trades. In Table 3, this proportion is denoted by $R_{w,x}^2$. In line with the results in Hasbrouck (1991b), we find that between 30% and 40% of the variance of w_t is explained by trading. The remainder is attributable to public information that is unrelated to the trading process. This result has to be interpreted with some caution. Recall that the regression error e_{2t} includes the linearization error of the discrete sign Q_t , and is therefore a combination of innovations in the trade process and measurement errors. Hence, the variance of e_{2t} may be larger than the variance of information revealed by trading.

4.5. A digression: Analysis of mid-quotes

To compare our results with the work of Hasbrouck (1991a), we estimate the VAR using *mid-quotes* instead of transaction prices. Fig. 2 shows the average impulse responses of mid-quotes to shocks in the trading, expressed as a fraction of the estimated spread (cf. Fig. 1). In contrast to the immediate adjustment of transaction prices, the mid-quotes slowly adjust to the new long run level. This pattern is very similar to the pattern that Hasbrouck (1991a) reports for mid-quotes on the NYSE. The cause of the apparent conflict between the estimates on mid-quotes and transaction prices must be found in the strong positive serial correlation in trade sign and size. To see how serial correlation affects the price effects measures, suppose that the transaction price p_t is a markup on the mid-quote, y_t (cf. Eq. (8)):

$$p_t = y_t + \delta Q_t + u_t. \quad (19)$$

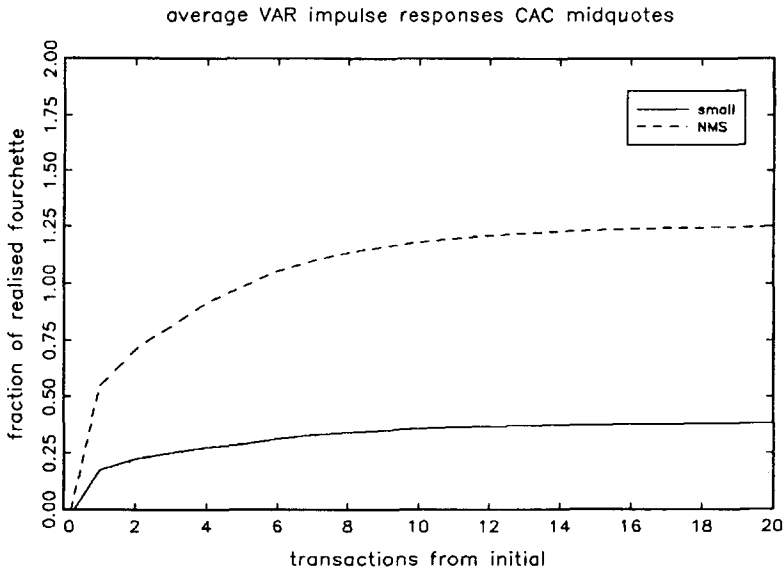


Fig. 2. Impulse responses of mid-quotes.

The solid line graphs the effect on mid-quotes of a shock in the sign equation, and the dotted line graphs the effect of a shock in both the sign and the size equation. The estimation results are obtained from the VAR model, described in Section 4.

In this simple model, the impulse response of $p_{t+\tau}$ to Q_t will be a factor $\delta \text{Cov}(Q_{t+\tau}, Q_t)$ larger than the impulse response of $y_{t+\tau}$. Apparently, the positive serial correlation in Q_t exactly cancels out the slow adjustment of the mid-quotes to their new equilibrium value. Because Q_t and $Q_{t+\tau}$ are almost uncorrelated for large τ , this model also explains why the estimates of the long run price effects are the same whether one uses mid-quotes or transaction prices. In our data, the first order autocorrelation in Q_t is high, about 0.3, and decays only slowly to zero. Hasbrouck (1991a) also reports estimates of serial correlation in the trade sign of the same order of magnitude.

5. Robust measures of price effects

Measuring dynamic effects of trading by a VAR model is a very general approach, but nevertheless imposes strong restrictions on the pattern of impulse responses. Moreover, the estimated coefficients are dominated by the covariance structure on low order lags; the long run effects are essentially determined by extrapolating the short run pattern of correlations. Campbell and Deaton (1989) and Cochrane (1988) convincingly argue that small changes in the VAR specifica-

tion can lead to substantial changes in the estimates of long run effects. To check the robustness of the VAR results, we adopt a more direct approach to estimating price effects. Although this could be achieved by specifying a general non-parametric regression function, we restrict ourselves to a simple linear parameterization:¹⁴

$$pe_2(\tau) = E(p_{t+\tau} - y_t | Q_t, z_t) = \beta_0^\tau + \beta_1^\tau Q_t + \beta_2^\tau z_t. \quad (20)$$

The coefficients β_1^τ and β_2^τ measure precisely the price effects of current trade sign and size. For $\tau = 0$, estimates of the realized bid-ask spread are obtained.

This measure is not exactly equal to the impulse responses calculated before because we do not condition on past values of sign and size. This conditioning can be achieved by adding more lags to (20):

$$pe_2(\tau) = E(p_{t+\tau} - y_t | I_t) = \beta_0^\tau + \beta_1^\tau Q_t + \beta_2^\tau z_t + \sum_{i=1}^p [\gamma_i^\tau Q_{t-i} + \delta_i^\tau z_{t-i}], \quad (21)$$

where I_t denotes the information set consisting of all past and current trade sign and size. The coefficients can be estimated by simple linear regression. To increase efficiency we use overlapping observation intervals if $\tau > 1$.

Fig. 3 graphs the estimated price effects of trading obtained from regression model (21) with two additional lags for horizons up to 20 transactions. The figure shows the point estimates of β_1^τ and $\beta_1^\tau + \beta_2^\tau$, scaled by the estimated spread, and averaged over all ten series. These curves correspond to the price effects of a small and large transaction, respectively. The patterns are very close to the estimates obtained by the VAR.

So far, we analyzed only transactions executed through the CAC system. We excluded all cross transactions, which are negotiated off-exchange. We now extend the model to allow for the effects of cross trades. As the price effects of crosses are expected to be different from the effects of CAC trades, we use a multiplicative dummy variable d_t to separate cross from CAC transactions (d_t equals zero if the transaction is from the CAC system, and one if it is a cross). The regression model then becomes

$$p_{t+\tau} - y_t = \beta_0^\tau + \beta_1^\tau (1 - d_t) Q_t + \beta_2^\tau (1 - d_t) z_t + \gamma_1^\tau d_t Q_t + \gamma_2^\tau d_t z_t + \epsilon_{t+\tau}^\tau. \quad (22)$$

The coefficients γ_1^τ and γ_2^τ provide estimates of the price effects of cross transactions. This model is extended by including two lags of the explanatory variables in a way similar to (21).

¹⁴ Several authors, e.g., Holthausen et al. (1990), Keim and Madhavan (1994) and Chan and Lakonishok (1993), report that the price response to buyer and seller initiated transactions on the US stock market is asymmetric. We assume linearity and hence symmetry here to facilitate comparison of the results with the results of parametric models, where the asymmetry is not easily included.

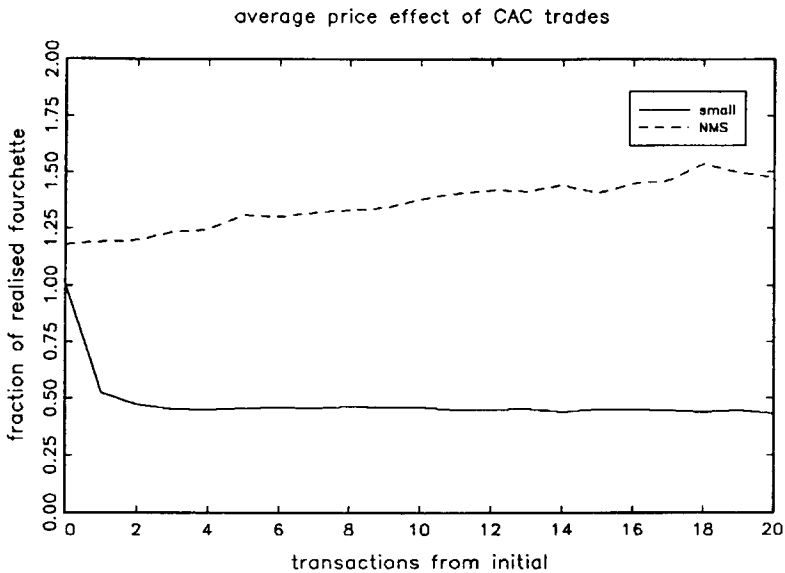


Fig. 3. Robust estimates of price effects of CAC transactions.

The solid line graphs the effect on mid-quotes of a shock in the sign, and the dotted line graphs the effect of a shock in both sign and size. The estimation results are obtained from regression Eq. (21) with two lags.

Fig. 4 reports the estimated price effects of the cross transactions. Although the initial effect ($\tau = 0$) is similar to the spread of the CAC transactions, the subsequent price effects are very different. The price effects of crosses are essentially zero at all horizons. Possibly cross transactions are informationless because an informed trader would not wish to incur the delays inherent in the process of searching for a counterparty, during which period his trading plans would be publicized. Moreover, counterparties would only be forthcoming if he could convince them that he was not trading on information.

6. Conclusions

In this paper we analyzed the intra-day price effects of trading on the Paris Bourse. The estimates of the Glosten (1994) model imply that adverse selection costs account for 30–45% of the bid-ask spread, and order processing cost for the remainder. Special attention was paid to separating temporary from permanent price effects. One remarkable result of the analysis is the difference in estimates obtained from structural models and from an impulse response analysis. More specifically, the Glosten (1994) model estimates price revisions of between 25% and 60% of the bid-ask spread for small and large transactions, respectively. The estimates of permanent price effects based on a VAR model and robust impulse

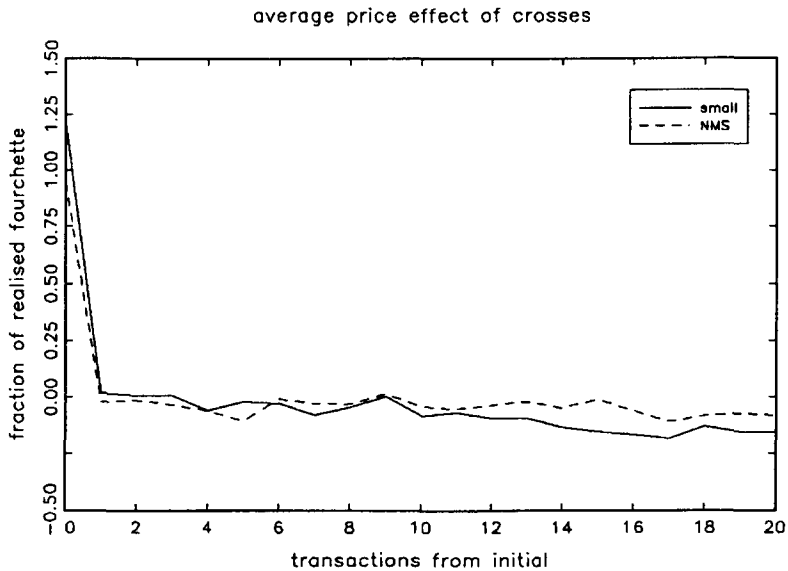


Fig. 4. Robust estimates of price effects of cross transactions.

The solid line graphs the effect on mid-quotes of a shock in the sign, and the dotted line graphs the effect of a shock in both sign and size. The estimation results are obtained from regression Eq. (22) with two lags.

response estimates are twice as large, ranging from 40% up to 115% of the bid-ask spread. Thus, the empirical results broadly confirm the predictions of the Glosten (1994) model for an electronic order driven market, but a one period empirical implementation of that model underestimates the price effects of trading.

The price effects of cross transactions are very different from the price effects of CAC transactions. Although the realized bid-ask spreads are similar to the spread for CAC transactions, the permanent price impact of crosses is virtually zero. The explanation must be that the off-exchange trading is not anonymous, and that asymmetric information plays less of a role in that market. The CAC system is anonymous and therefore potentially prone to adverse selection problems. This might explain why the CAC system seems to specialize as a retail market for small transactions, and large transactions are either crosses or are executed on London's SEAQ International.

7. For further reading

Easley and O'Hara (1987), George et al. (1991), and Glosten and Milgrom (1985).

Acknowledgements

We thank Gerard Pfann and two anonymous referees for their comments. Other useful comments on previous drafts were given by Bruno Biais, Ekkehart Boehmer and seminar participants at the universities of Amsterdam, Berlin (Humboldt), Mannheim, Tilburg, and Rotterdam. Of course, the usual disclaimer applies. The data were kindly provided by the French stock exchange.

References

- Beveridge and Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle". *Journal of Monetary Economics* 7, 151–174.
- Biais, Bruno, Pierre Hillion and Chester Spatt, 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance* 50, 1655–1689.
- Campbell, John and Angus Deaton, 1989, Why is consumption so smooth?, *Review of Economic Studies* 56, 357–374.
- Chan, Louis and Josef Lakonishok, 1993, Institutional trades and intraday stock price behavior, *Journal of Financial Economics* 33, 173–199.
- Cochrane, John, 1988, How big is the random walk in GNP?, *Journal of Political Economy* 96, 893–920.
- Conrad, Jennifer and Gautam Kaul, 1989, Mean reversion in short horizon expected returns, *Review of Financial Studies* 2, 225–240.
- De Jong, Frank, Theo Nijman and Ailsa Röell, 1995, A comparison of the cost of trading French shares on the Paris Bourse and on SEAQ International, *European Economic Review* 39, 1277–1301.
- Easley, David and Maureen O'Hara, 1987, Price, trade size and information in securities markets, *Journal of Financial Economics* 19, 69–90.
- George, Thomas, Gautam Kaul and M. Nimalendran, 1991, Estimation of the bid-ask spread and its components: A new approach, *Review of Financial Studies* 4, 623–656.
- Glosten, Lawrence, 1994, Is the electronic open limit order book inevitable?, *Journal of Finance* 49, 1127–1161.
- Glosten, Lawrence and Lawrence Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123–142.
- Glosten, Lawrence and Paul Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.
- Harris, Lawrence, 1986, A transaction data study of weekly and intra-daily patterns in stock returns, *Journal of Financial Economics* 16, 99–117.
- Hasbrouck, Joel, 1988, Traders, quotes, inventories and information, *Journal of Financial Economics* 22, 229–252.
- Hasbrouck, Joel, 1991a, Measuring the information content of stock trades, *Journal of Finance* 46, 179–207.
- Hasbrouck, Joel, 1991b, The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies* 4, 571–595.
- Hasbrouck, Joel, 1993, Assessing the quality of a security market: A new approach to transaction cost measurement, *Review of Financial Studies* 6, 191–212.
- Hasbrouck, Joel and George Sofianos, 1992, The trades of market makers: An analysis of NYSE specialists, mimeo.

- Hausman, Jerry, Andrew Lo and A. Craig MacKinlay, 1992, An ordered probit analysis of transaction stock prices, *Journal of Financial Economics* 31, 319–379.
- Heckman, James J., 1978, Dummy endogenous variables in a simultaneous equation system, *Econometrica* 46, 931–959.
- Holthausen, Robert, Richard Leftwich and David Mayers, 1990, Large-block transactions, the speed of response, and temporary and permanent stock-price effects, *Journal of Financial Economics* 26, 71–95.
- Keim, Donald and Ananth Madhavan, 1994, The upstairs market for large-block transactions: Analysis and measurement of price effects, Working paper (Wharton School, University of Pennsylvania).
- Lo, Andrew and A. Craig MacKinlay, 1988, Stock prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.
- Madhavan, Ananth, Matthew Richardson and Mark Roomans, 1994, Why do security prices change? A transaction level analysis of NYSE stocks, mimeo.
- Madhavan, Ananth and Seymour Smidt, 1991, A Bayesian model of intraday specialist pricing, *Journal of Financial Economics* 30, 99–134.
- Madhavan, Ananth and Seymour Smidt, 1993, An analysis of daily changes in specialist inventories and quotations, *Journal of Finance* 48, 1595–1628.
- Newey, Whitney and Kenneth West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 39, 1127–1139.
- Sims, Christopher, 1980, Macroeconomics and reality, *Econometrica* 48, 1–48.
- Stoll, Hans, 1989, Inferring the components of the bid-ask spread: Theory and empirical tests, *Journal of Finance* 44, 115–134.