

Alvaro Escribano · Roberto Pascual

# Asymmetries in bid and ask responses to innovations in the trading process

Published online: 23 November 2005  
© Springer Verlag 2005

**Abstract** This paper proposes a new approach to jointly model the trading process and the revisions of market quotes. This method accommodates asymmetries in the dynamics of ask and bid quotes after trade related shocks. The empirical specification is a vector error correction (VEC) model for ask and bid quotes, with the spread as the co integrating vector, and with an endogenous trading process. This model extends the vector autoregressive (VAR) model introduced by Hasbrouck (Hasbrouck J (1991) Measuring the information content of stock trades. *J Finance* 46:179–207). We provide evidence against several symmetry assumptions, very familiar among microstructure models. We report asymmetric adjustments of ask and bid prices to trade related shocks, and asymmetric impacts of buyer and seller initiated trades. In general, buys are more informative than sells. The likelihood of symmetric quote responses increases with volatility. We show that our findings are robust across different model specifications, time frequencies, and trading periods. Moreover, we find similar asymmetries in markets with different microstructures.

**Keywords** Market microstructure · Bid and ask time series · VEC models · Adverse selection costs · Asymmetric dynamics

**JEL Classification** G1

---

This paper has benefited from the support of the Spanish DGICYT project #PB98 0030 and the European Project on VPM Improving Human Research Potential, HPRN CT 2002 00232. The authors are grateful for the comments received from an anonymous referee and from Mikel Tapia, Ignacio Peña, Winfried Pohlmeier and the attendants to the Econometrics Research Seminar at C.O.R.E., Université Catholique de Louvain, Belgium. We also appreciate the suggestions of participants at the CAF Market Microstructure and High Frequency Data in Finance Workshop, August 2001, Sønderborg (Denmark), and the European Financial Association Meeting, August 2001, Barcelona (Spain)

---

A. Escribano (✉)  
Departament of Economics, Universidad Carlos III de Madrid, C/Madrid 126, Getafe, 28903 Madrid, Spain  
E mail: alvaroe@eco.uc3m.es

R. Pascual  
Departamento de Economía de la Empresa, Universidad de las Islas Baleares, Madrid, Spain  
E mail: rpascual@uib.es

# 1 Introduction

In this paper, we propose a new econometric approach to jointly model the time series dynamics of the trading process and the revisions of ask and bid prices. We use this model to test the validity of certain symmetry assumptions very common among microstructure models. Namely, we test whether ask and bid quotes respond symmetrically to trade related shocks, and whether buyer initiated trades and seller initiated trades are equally informative. In essence, the procedure we propose generalizes Hasbrouck's (1991) vector autoregressive model for signed trades and changes in the quote midpoint by relaxing the implicit symmetry assumptions in his model.

The properties of the empirical model are derived from a structural dynamic model for ask and bid prices. In this model, ask and bid prices share a common long run component, the efficient price. The long term value of the stock varies due to buyer initiated shocks, seller initiated shocks, and trade unrelated shocks. The transitory components of ask and bid prices are characterized by two correlated and trade dependent stochastic processes, whose dynamics are allowed to differ. The trading process is endogenous. Buyer and seller initiated trades are generated by two idiosyncratic but mutually dependent stochastic processes. The generating processes of quotes and trades both depend on several exogenous variables that feature the trades and the market conditions.

We demonstrate that the empirical counterpart of this theoretical model is an extended vector error correction (VEC) model with four dependent variables: changes in the ask price, changes in the bid price, buyer initiated trades, and seller initiated trades. The bid ask spread is the error correction term. Our VEC model reverts to the Hasbrouck's (1991) bivariate VAR model when: (a) ask and bid responses to trade related shocks perfectly match; (b) the generating processes of buyer and seller initiated trades are equivalent, and (c) the trade sign only matters as far as the direction of the quote adjustments is concerned.

For robustness purposes, we implement the model using three different subsamples: the 11 most frequently traded NYSE listed stocks in 1996 and 2000, and the 11 most active stocks at the Spanish Stock Exchange (SSE) in 2000. With the two NYSE subsamples, we show that our main findings are not period specific. We also show that our findings are unaltered by the dramatic increase in trading activity and the progressive decrease in the minimum price variation experienced by the NYSE from 1996 to 2000. With the Spanish data, we show that our findings are not limited to the particular microstructure of the NYSE. We perform additional robustness test considering alternative specifications of the empirical model.

We find two main patterns characterizing the dynamics of market quotes. On the one hand, ask and bid quotes do not respond symmetrically after trade related shocks. They tend to be revised in the same direction, but not by the same amount. We show, however, that the likelihood of observing a symmetric response increases with volatility. On the other hand, ask and bid prices error correct after a trade, which causes the spread to revert towards the minimum. The speed of reversion is significantly non linear. The wider the bid ask spread, the quicker the response of ask and bid quotes. These patterns result in two simultaneous but opposite effects on the price dynamics: information induced positive cross serial correlation and liquidity induced negative cross serial correlation.

For the NYSE samples, we report that buyer initiated trades are more informative than seller initiated trades. Namely, we find that the average long term impact of a buyer initiated trade on the ask quote is larger than the average long term impact of a similar seller initiated trade on the bid quote. In the SSE, however, no statistical difference is found between the impact of buyer and seller initiated trades.

In general, our findings evidence that the dynamics between quotes and trades are more complex than suggested by classical microstructure models of quote formation. We also show that asymmetries are not exclusive of the NYSE, since they are also found in an electronic order driven market without market makers, the SSE. In addition, our findings demonstrate that there is an important loss of information in averaging the dynamics of ask and bid quotes through the quote midpoint instead of jointly modeling them.

The remainder of the paper proceeds as follows. In Section 2, we review the literature and motivate the model. In Section 3, we present the theoretical dynamic model and its empirical counterpart. In Section 4, we describe the data. In Section 5, we analyze in detail a representative stock: IBM. In Section 6, we perform several robustness tests. Finally, in Section 7, we conclude.

## 2 Motivation

A large part of market microstructure research builds on the notion that trades convey new information that updates the market's expectation about the long run value of the stock. This trade related information causes simultaneous revisions of market quotes (e.g., Hasbrouck 1996). Ask and bid quotes are usually modeled as the result of adding a premium and subtracting a discount to the efficient price (e.g., Glosten 1987). The magnitude of these perturbations depends on certain market frictions, such as price discreteness, and market making costs.<sup>1</sup>

For simplicity purposes, many classic theoretical models of price formation impose, in some degree, what we will call in this paper the “symmetry assumption”. First, the symmetry assumption implies that offer and demand quotes are posted symmetrically about the efficient price. Thus, in some models the transitory components of ask and bid prices are constant and equal sized (e.g., Roll 1984; Madhavan et al. 1997; Huang and Stoll 1997); in some other cases, their dynamics are characterized by the same stochastic process (e.g., Glosten and Harris 1988; Lin et al. 1995; Hasbrouck 1999b). Second, the symmetry assumption implies that ask and bid quotes respond identically after a trade related shock. Thus, a common premise in theoretical models is that ask and bid prices are simultaneously revised upward or downward, usually by the same amount, after a trade related shock (e.g., Glosten and Milgrom 1985; Stoll 1989). Finally, the symmetry assumption implies that whether a trade is buyer or seller initiated matters to determine the direction, but not the magnitude, of ask and bid updates (e.g., Easley and O'Hara 1992). That is, buys and sells are assumed to be equally informative.<sup>2</sup>

---

<sup>1</sup> See O'Hara (1995) for a review of the basics of this literature.

<sup>2</sup> The symmetry assumption is sometimes relaxed. For example, Easley and O'Hara (1987) allow different sequences of trades to have different price impacts.

Jang and Venkatesh (1991) reports that, in the NYSE, one step ahead revisions in ask and bid quotes right after trades do not generally support the theoretical prediction of symmetry. Moreover, quote revision patterns strongly depend on the level of the outstanding spread, with more symmetric adjustments as the spread augments. They argue the symmetry assumption may be violated in practice because certain theoretical premises are simply unrealistic. Thus, it is usually assumed that quotes can be adjusted in a continuous fashion, which is not possible because of the minimum price variations. In addition, it is usually supposed that posted quotes are always for the specialist own account. However, in most stock exchanges, as is the case of the NYSE, quotes reflect the interest of several traders that may be selectively offering one sided liquidity (see Madhavan and Sofianos 1998; Kavajecz 1999; Chung et al. 1999). Since different agents may be subject to different trading costs, the offer and demand components of the spread may vary asymmetrically about the efficient price.

Hasbrouck (1999a) points out that the usual theoretical premise of equal market making costs at the offer and demand sides of the market is reasonable if the same quote setter is active on both sides. However, even if the specialist would take all trades, it is not clear that she would adjust quotes simultaneously and by the same amount after a trade. Thus, if the trade signals the presence of informed traders, a natural response of the specialist would be to post a wider spread. Therefore, she would update ask and bid quotes asymmetrically. Moreover, the costs of ask and bid exposure might not necessarily be balanced. Thus, a specialist offering liquidity in times of an upward price pressure would suffer from higher exposure costs on the ask side than on the bid side of the market.

Biais et al. (1995) shows that asymmetries between ask and bid quotes are not an exclusivity of the NYSE. Using data on a pure order driven market, the Paris Bourse, they show that ask and bid one step ahead adjustments after trades are also asymmetric. Their findings hint that the ask (bid) quote may lead the adjustment of the bid (ask) price after an order to buy (sell). These authors conclude that “there is additional information in analyzing the dynamics of ask and bid prices jointly rather than averaging them through the quote midpoint” (pg. 1679).

There is also evidence suggesting that the impact of a buyer initiated trade may not be merely the reverse of the price impact of seller initiated trade. Empirical work on block trading (e.g., Holthausen et al. 1987; Griffiths et al. 2000; Koski and Michaely 2000) shows that orders to sell and orders to buy may have different permanent and transitory impacts on prices. Chan and Lakonishok (1993, 1995) use a broad range of trade sizes to evidence differences in the behavior of prices after institutional purchases and sales. Keim and Madhavan (1995) conclude that large buys take longer to execute than equivalent sells because traders perceive that price impacts of buys are greater than sells. Huang and Stoll (1996) report different realized spreads for buyer and seller initiated trades. Similarly, Lakonishok and Lee (2001) observe that the information content of insider’s activities come from purchases while insider selling appears to have no predictive ability.<sup>3</sup>

---

<sup>3</sup>These asymmetries might not be exclusive of large sized trades. Hasbrouck (1988, 1991), Barclay and Warner (1993), Kempf and Korn (1999), among others, evidence the relationship between trade size and price impact is increasing, but concave. Therefore, informed traders may concentrate their trades in medium sizes.

In all these papers, buyer initiated trades are usually found to be more informative than seller initiated trades. Price responses to buyer initiated versus seller initiated trades may be asymmetric for a variety of reasons. Firstly, short selling restrictions may prevent insiders from exploiting negative information (Kempf and Korn 1999). Secondly, since an investor typically does not hold the market portfolio, the choice of a particular stock to sell does not necessarily convey negative information. On the contrary, the choice of a particular stock to buy, out of the numerous possibilities on the market, is likely to convey favorable firm specific news (Chan and Lakonishok 1993). Finally, a sell order representing a small fraction of the initiator's known position may be considered as more liquidity motivated than a similar buy order from an investor without current holdings in the security (Keim and Madhavan 1995).<sup>4</sup>

Alternative empirical approaches have been proposed to jointly model the generating processes of quotes and trades. The most influential is probably due to Hasbrouck (1991).<sup>5</sup> He suggests the following vector autoregressive (VAR) specification,

$$\begin{aligned}\Delta q_t &= \sum_{i=1}^{\infty} a_i \Delta q_{t-i} + \sum_{i=0}^{\infty} b_i x_{t-i} + v_{1,t} \\ x_t &= \sum_{i=1}^{\infty} c_i \Delta q_{t-i} + \sum_{i=1}^{\infty} d_i x_{t-i} + v_{2,t} \end{aligned} \quad (2.1)$$

where  $\Delta q_t = (q_t - q_{t-1})$  represents the revision in the quote midpoint ( $q_t$ ) after a trade at  $t$  and  $x_t$  is a trade indicator that equals 1 for buyer initiated trades and -1 for seller initiated trades. The terms  $v_{1,t}$  and  $v_{2,t}$  are mutually and serially uncorrelated white noises that represent trade unrelated and trade related shocks respectively. This econometric approach covers the dynamics of many structural microstructure models as special cases (see Hasbrouck 1991, 1996).

Hasbrouck builds on a "weak symmetry assumption:" the quote midpoint must revert to the efficient price as the end of trading approaches. Therefore, ask and bid prices may not be symmetrically posted around the efficient price. However, since the quote dynamics are averaged through the quote midpoint ( $q_t$ ), the VAR model is not a valid framework to accommodate and evaluate possible asymmetries in the dynamics of ask and bid prices. Similarly, the trade dynamics are averaged through

---

<sup>4</sup> Other aspects of the trading process that may produce a lack of balance between the impact of buys and sells are transitory market conditions and the trade durations. Thus, a larger market pressure to sell than to buy might increase the expected impact of a seller initiated market order (e.g., Goldstein and Kavajecz 2004) versus a similar buyer initiated order. Easley et al. (1997) and Dufour and Engle (2000) show that trade durations (time between consecutive trades) partly explain the long term impact of trades, even when trade size is accounted for. If a market overreacts to bad news, it might produce shorter time durations after seller initiated trades than after buyer initiated trades, and therefore asymmetries in the responses of quotes between buys and sells.

<sup>5</sup> Other econometric approaches, parametric and semi parametric respectively, to model the relationship between the trading process and the price changes are Hausman et al. (1992), that used ordered probit models, and Kempf and Korn (1999) that employed a neural networks type model.

the trade indicator ( $x_t$ ). Therefore, the expected impact of a buyer initiated shock is exactly the reverse of a seller initiated shock.

So as to allow for asymmetric dynamics, next section modifies the structural model used by Hasbrouck (1991) to motivate his VAR model. The empirical counterpart of this most flexible structural model will be a VEC model for ask and bid revisions in response to buyer and seller initiated shocks, which generalizes Eq. (2.1). Given that ask and bid quotes have a common non stationary long run component, the efficient stock price, they must be co integrated time series (see Hasbrouck 1995). The VEC model is the most common efficient parameterization of vector autoregressive models with co integrated variables (e.g., Engle and Granger 1987). In this case, the co integration relationship is known a priori, which lets setting a very general parameterization of the model. We will also allow quote revisions after trade related shocks to follow non linear patterns due to trade features, such as size and durations, and market conditions, such as volatility and liquidity.

Our approach connects with other econometric applications in microstructure research. Hasbrouck (1995) uses a common trend representation to simultaneously model the quotes of both the NYSE and the regional markets. The model has an associated VEC representation. Hasbrouck is aimed to measure relative contributions to price discovery. Pascual et al. (2005) build on the model introduced in the next sections to incorporate the markets' trading processes into Hasbrouck's (1995) methodology. In this way, they are able to isolate, for each market, the trade related contribution from the trade unrelated contribution.

Hasbrouck (1999b) estimates an unobserved components model for the best market quotes. In this model, ask and bid quotes have a common random walk component, and the respective transitory terms are modeled as two unobserved and identical first order autoregressive processes. This model features discreteness, clustering, and stochastic volatility effects. However, it does not incorporate the trading process, and the transitory components of ask and bid quotes are assumed mutually independent. This analysis does not deal with asymmetric dynamics either. In a recent paper, Zhang et al. (2005) use Hasbrouck's (1999b) methodology to decompose the bid ask spread of a single stock (General Electric) into its ask and bid exposure costs constituents. They show that ask and bid components of the spread change asymmetrically about the efficient price. Our findings, based on a less sophisticated but more widespread econometric approach, are totally consistent.

Even closer to the purpose of this paper is the independent study by Engle and Patton (2004). These authors (henceforth, EP) also estimate an error correction model for ask and bid quotes using data on a large set of NYSE listed stocks. There are, however, remarkable differences between both empirical specifications. In EP's model, the trading process is exogenous. Thus, they model the dynamics of ask and bid quotes, but not the feedback from quotes to trades. As shown by Hasbrouck (1991), to accurately measure the informativeness of trades, we need to model not only how quotes evolve after the trade, but also how trading responds to the progressive adjustment in quotes. Since we are aimed to compare the information content of buyer and seller initiated trades, we propose a model that accommodates a broader set of dynamic interactions between quotes and trades than EP's model. Moreover, EP's model is quote driven, meaning that there is a new observation each time there is a change in quotes. Since we are interested in the price impact of trades, it makes sense to define our model in trade time. In this manner, we filter those quote changes that are not directly linked to the trading

process.<sup>6</sup> These technical disparities may explain the discrepant findings we will report later on regarding the information content of buys versus sells. Nevertheless, in our opinion, EP's paper and the present paper are complementary, since both evidence the relevance of modeling ask and bid prices jointly rather than averaging them through the quote midpoint.

### 3 The model

#### 3.1 A structural dynamic model of quote formation

In this subsection, we build on Hasbrouck (1991) to develop a dynamic model for ask and bid quotes. The model allows for asymmetric adjustment paths going after trades. Two main features differentiate our structural model from those previously found in the literature. First, ask and bid prices share a common long run component, the efficient price, which is updated due to trade related and trade unrelated informative shocks. The empirical evidence previously revised, however, suggests that buyer and seller initiated trades (henceforth, "buys" and "sells") may not be equally informative. To accommodate this empirical observation, we will distinguish between trade related shocks to buy and trade related shocks to sell. Second, quotes result from adding or subtracting a transitory component ( $w_t$ ), due to market frictions and market making costs, to the efficient price ( $m_t$ ). Since the evidence at hand points to asymmetric short term adjustments of the ask quote versus the bid quote, in our specification we will allow the short term components of ask and bid quotes to differ ( $\Delta w_t^a \neq \Delta w_t^b$ ).

We use the same notation as in Hasbrouck (1991). The model is defined in trade time. Thus, the subscript  $t$  denotes the  $t$  th trade in the chronological sequence of trades. Hereafter, the superscript  $a$  means "ask quote,"  $b$  means "bid quote,"  $B$  refers to buys, and  $S$  refers to sells.  $m_t$  is the efficient price after the  $t$  th trade, which could be either a buy ( $x_t^B$ ) or a sell ( $x_t^S$ ). Similarly,  $a_t$  and  $b_t$  are the quotes posted right after the  $t$  th trade. The adjustment in the posted quotes after the  $t$  th trade are  $\Delta a_t = a_t - a_{t-1}$  and  $\Delta b_t = b_t - b_{t-1}$ .

The efficient price follows the random walk process in Eq. (3.1). Three types of stochastic shocks update  $m_t$ : trade unrelated shocks ( $v_{1,t}$ ), and trade related shocks due to buyer initiated trades ( $v_{2,t}^B$ ) or seller initiated trades ( $v_{2,t}^S$ ). We let  $v_{1,t}$  be mutually and serially uncorrelated with  $v_{2,t}^B$  and  $v_{2,t}^S$ , while  $v_{2,t}^B$  and  $v_{2,t}^S$  are serially uncorrelated but, perhaps, mutually correlated. The parameters  $\lambda^B$  and  $\lambda^S$  measure the average amount of private information (adverse selection costs) conveyed by buys and sells, respectively. If  $\lambda^B = \lambda^S$  buy shocks would have the same information content than sell shocks.

$$m_t = m_{t-1} + \lambda^B v_{2,t}^B + \lambda^S v_{2,t}^S + v_{1,t}. \quad (3.1)$$

---

<sup>6</sup> Another minor difference is that, unlike EP's model, the properties of our empirical specification are derived from a theoretical framework presented in the next section. The empirical model we derive from the structural model is an "extended" VECM, that is, it incorporates lagged values of the error correction term. The EP's model is a standard VEC.

The generating processes of market quotes are given by Eqs. (3.2) and (3.3),

$$\begin{aligned} a_t &= m_t + w_t^a = \\ &= m_t + \alpha_m^a (a_{t-1} - m_t) + A_{x,t}(L)' x_t + \alpha_a^{EC} (a_{t-1} - b_{t-1}) + \varepsilon_t^a \end{aligned} \quad (3.2)$$

$$\begin{aligned} b_t &= m_t - w_t^b = \\ &= m_t + \alpha_m^b (m_{t-1} - b_{t-1}) + B_{x,t}(L)' x_t + \alpha_b^{EC} (a_{t-1} - b_{t-1}) + \varepsilon_t^b. \end{aligned} \quad (3.3)$$

In Eqs. (3.2)–(3.3),  $a_t$  and  $b_t$  are the result of adding a time varying stationary premium ( $w_t^a$ ) and subtracting a time varying stationary discount ( $w_t^b$ ), respectively, to the efficient price. These transitory components are time varying because we assume they are determined by the recent history of trades and quotes. We allow the magnitudes of these two components to differ. Therefore,  $a_t$  and  $b_t$  may not be symmetrically posted about  $m_t$ .<sup>7</sup> Moreover, we impose  $0 < \alpha_m^a < 1$  and  $0 < \alpha_m^b < 1$ , implying that, in the absence of trading,  $a_t$  and  $b_t$  revert to the efficient price. The noise terms  $\varepsilon_t^a$  and  $\varepsilon_t^b$  are idiosyncratic errors reflecting market frictions and model misspecifications. Finally, the vectors  $A_{x,t}(L)' = (A_{x,t}^B(L), A_{x,t}^S(L))$  and  $B_{x,t}(L)' = (B_{x,t}^B(L), B_{x,t}^S(L))$  are finite order polynomials in the lag operator  $L(L^k y_t = y_{t-k})$  with time varying components. These polynomials would capture the transitory effect of trades on quotes. The dynamic structure denotes that  $a_t$  and  $b_t$  adjustments to trade related shocks are progressive.<sup>8</sup>

The vector  $x_t' = (x_t^B, x_t^S)$  in Eqs. (3.2)–(3.3) includes the time series of buys and sells. The trading process is endogenous. The idiosyncratic, but mutually dependent, stochastic processes in Eqs. (3.4) and (3.5) generate buys and sells,

$$x_t^B = \mu^B (a_{t-1} - m_{t-1}) + \pi^B (a_{t-1} - b_{t-1}) + v_{2,t}^B \quad (3.4)$$

$$x_t^S = \mu^S (m_{t-1} - b_{t-1}) + \pi^S (a_{t-1} - b_{t-1}) + v_{2,t}^S. \quad (3.5)$$

In Eqs. (3.4)–(3.5), the likelihood of observing a new trade decreases with its specific exposure costs ( $\mu^B < 0$ ,  $\mu^S < 0$ ) and the costs of executing a round trip ( $\pi^B < 0$ ,  $\pi^S < 0$ ), defining downward sloping demand schedules. The terms  $v_{2,t}^B$  and  $v_{2,t}^S$  are the mutually correlated unexpected components of buys and sells, respectively.

The third element on the RHS of Eqs. (3.2)–(3.3) is decomposed in terms of buys and sells as follows,

$$\begin{aligned} A_{x,t}(L)' x_t &= A_x^B(L) f_a^B(MC_t, D_t) x_t^B + A_x^S(L) f_a^S(MC_t, D_t) x_t^S \\ B_{x,t}(L)' x_t &= B_x^B(L) f_b^B(MC_t, D_t) x_t^B + B_x^S(L) f_b^S(MC_t, D_t) x_t^S, \end{aligned}$$

where  $A_x^B(L)$ ,  $A_x^S(L)$ ,  $B_x^B(L)$ , and  $B_x^S(L)$  are finite time invariant order polynomials in the lag operator  $L$ , having all roots outside the unit circle.

<sup>7</sup> Hasbrouck (1999b) models the exposure costs for bid and ask quotes as two independent stochastic processes. In our case, the transitory components could be mutually correlated because of the common components.

<sup>8</sup> Stabilizing NYSE rules (see Hasbrouck et al. 1993) and heterogeneous priors among traders (e.g., Harris and Raviv 1993) may explain the lagged effects of trades.



The terms  $f_i^B(MC_t, D_t)$  and  $f_i^S(MC_t, D_t)$ ,  $i \in \{a, b\}$ , are functional forms of two vectors of variables. The first vector ( $MC_t$ ) includes exogenous variables that characterize the trade and the market environment. The second vector ( $D_t$ ) control for trading time regularities. The particular functional form considered is given in Eq. (3.6). We impose linearity for simplicity reasons. The price impact of a given trade is conditioned on these set of exogenous and deterministic variables, that we will specify latter on.

$$f_i^j(MC_t, D_t) = 1 + \sum_{k=1}^n \lambda_k^{i,j} MC_t^k + \sum_{h=1}^{n'} \gamma_h^{i,j} D_t^h, \quad i \in \{a, b\}, j \in \{B, S\} \quad (3.6)$$

From Eqs. (3.1) to (3.3),  $a_t$  and  $b_t$  are nonstationary, integrated of order one, processes. Nonstationarity comes from the common long run component ( $m_t$ ), implying that the time series  $a_t$  and  $b_t$  must be co integrated.<sup>9</sup> Our application has the unusual advantage that the co integration relationship has a known co integration vector (1, 1). The co integration relationship is, therefore,  $a_t - b_t$ , the bid ask spread (henceforth,  $s_t$ ).

An increase in  $s_t$  represents a departure from the long run equilibrium relationship between  $a_t$  and  $b_t$ . The error correction mechanism produces simultaneous revisions in both ask and bid quotes that correct such deviations. For this reason, we incorporate  $s_t$  into Eqs. (3.2)–(3.3) as a determinant of the transitory components of  $a_t$  and  $b_t$ . The coefficients  $\alpha_a^{EC}$  and  $\alpha_b^{EC}$  show how quickly do  $a_t$  and  $b_t$  revert to their common long run equilibrium value.

### 3.2 The empirical model

The most common efficient parameterization of a vector autoregressive (VAR) model with co integrated variables is, from Granger's representation theorem in Engle and Granger (1987), a vector error correction (VEC) model. In the Appendix I, we give an explicit derivation of the VEC model in Eq. (3.7) from the structural model in the previous subsection,

$$\begin{pmatrix} 1 & 0 & A_{aB,t}^* & A_{aB,t}^* \\ 0 & 1 & A_{bB,t}^* & A_{bS,t}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ x_t^B \\ x_t^S \end{pmatrix} = \begin{pmatrix} \gamma_a^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A_t(L) \begin{pmatrix} \Delta a_t & 1 \\ \Delta b_t & 1 \\ x_t^B & 1 \\ x_t^S & 1 \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix}, \quad (3.7)$$

<sup>9</sup> Engle and Granger (1987), Stock and Watson (1988), Johansen (1991), and Escribano and Peña (1994), among others, provide formal derivations of this result.

with

$$A_t(L) = \begin{pmatrix} A_{aa}(L)A_{ab}(L)A_{aB,t}(L)A_{aS,t}(L) \\ A_{ba}(L)A_{bb}(L)A_{bB,t}(L)A_{bS,t}(L) \\ A_{Ba}(L)A_{Bb}(L)A_{BB,t}(L)A_{BS,t}(L) \\ A_{Sa}(L)A_{Sb}(L)A_{SB,t}(L)A_{SS,t}(L) \end{pmatrix}.$$

This model echoes the main features of the structural model in the previous subsection. First, the bid-ask spread  $s_t = a_t - b_t$  is the error correction term. Second, the matrix on the left-hand side of Eq. (3.7) reflects that the theoretical model is trade-driven. Thus, trades have a contemporaneous effect on ask and bid quotes. The reverse, however, is not true. Third, the matrix of autoregressive polynomials  $A_t(L)$  depicts the dynamical structure of the theoretical model. Moreover,  $A_{ij}(L)$ , for all  $i, j \in \{a, b, B, S\}$ , has its roots outside the unit circle. Thus, the influence of past quotes and trades decays with time. Finally, the polynomials  $A_{ij,t}(L)$  are time-varying because they depend on a set of exogenous variables ( $MC_t$ ) and trading-time dummies ( $D_t$ ). The following expression makes explicit the type of dependence,

$$A_{ij,t}(L)x_{t-1} = A_{ij}^B(L)f_{ij}^B(MC_{t-1}, D_{t-1})x_{t-1}^B + A_{ij}^S(L)f_{ij}^S(MC_{t-1}, D_{t-1})x_{t-1}^S.$$

The polynomials  $A_{ij,t}^B(L)$  and  $A_{ij,t}^S(L)$  have all the roots outside the unit circle. Finally,  $A_{ij,t}^* = A_{ij,t}(0)$ .

A salient feature of the VEC model Eq. (3.7) is the extra lags in the error correction term. This type of specification is called an extended vector error correction (EVEC) model. Arranz and Escribano (2000) show that extended error correction models are robust to the presence of structural breaks under partial co-breaking. Co-breaks represent those situations characterized by having breaks (level shifts, changes in trend etc.) occurring simultaneously in some variables, so that certain linear combinations of those variables have no breaks. The common long-run trend jointly with their discrete type of moves makes  $a_t$  and  $b_t$  the perfect example of co-integrated time series that are partially co-breaking. Thus, this property of the model is consistent with the properties of the time series of ask and bid prices. The error correction terms  $\gamma_a^{EC}(L)s_{t-1}$  and  $\gamma_b^{EC}(L)s_{t-1}$  should be such that  $\gamma_a^{EC}(1) - \gamma_b^{EC}(1) < 0$ , in order to impose the error correction characterization on the spread. Extended error correction parameterizations of VAR models with co-integrated variables could be formally justified using the Smith-MacMillan decomposition introduced by Engle and Yoo (1991).

The individual error terms  $u_t^i$  in Eq. (3.7)  $i = \{a, b, B, S\}$  are assumed to be serially uncorrelated random variables with zero mean and constant variance. We show in Appendix I that they cannot be treated as mutually uncorrelated since they have common components. Hence, the system of Eq. (3.7) is an example of seemingly unrelated regression equations, which can be efficiently estimated by SURE (see Zellner 1962). Estimating a system by SURE is equivalent to estimating it equation by equation by OLS when all equations have the same number of variables. In other case, all equations should be simultaneously estimated by SURE to get efficiency. Notice also that, under the restrictions imposed on the structural model in the previous section, Eq. (3.7) is exactly identified.

Next, we proceed with the estimation of the VEC model Eq. (3.7). In a preliminary step, however, we consider a base-line version of Eq. (3.7) where

$f_i^j(MC_i, D_i) = 1$  for all  $i$  and  $j$ . In this case, the matrix of autoregressive polynomials is time invariant,  $A_i(L) = A(L)$ , and the impact of trades on quotes is perfectly linear. We will show that this model suffices to illustrate the essentials of the dynamic relationship between trades and quotes. However, some other aspects of this relationship can only be captured by considering the more general case.

Following Hasbrouck (1991), we characterize the trading processes using indicator variables. Namely,  $x_t^B$  equals one for buys and zero otherwise, and  $x_t^S$  equals one for sells and zero otherwise. The discreteness of these variables, however, may introduce some problems in the estimation process.<sup>10</sup> To control for these potential problems, we also estimate Eq. (3.7) using the trade size ( $\tilde{x}_t^i$ ) to characterize each transaction. In particular, we define  $\tilde{x}_t^i = x_t^i \log(V_t)$ , where  $V_t$  is the size of the  $t$ th trade in shares.

## 4 Data

The database comprises high frequency data on trades and quotes from two markets with remarkably different microstructures: the NYSE and the SSE. The NYSE is a peculiar mixture of microstructure types. It combines an electronic limit order book, only partially transparent, with monopolist market makers, and an intensive trading activity at the floor market. The SSE, on the contrary, is a representative example of an electronic order driven venue. Liquidity provision depends exclusively on a fully transparent open limit order book. Twenty levels of the book are nowadays visible in real time for all market participants. There are no market makers, no floor trading, price improvements are not possible, and all the orders are submitted through vendor feeds, and stored or matched electronically.

We use data on two different markets to show that asymmetric dynamics between ask and bid quotes in response to trades are not exclusive of the NYSE. In addition, trades in the SSE always involve a market order (or equivalent), the initiating side, and one or more limit orders stored on the book. Therefore, trades are straightforwardly classified as either buyer or seller initiated by simply identifying the side of the book the market order hits. Thus, with the SSE data we do not bear the ambiguity and misclassification problems that appear when traditional trade direction algorithms, such as Lee and Ready (1991), are applied to NYSE data (see Ellis et al. 2000, and Odders White 2000). Finally, using Spanish data we do not have reporting delays neither in trades nor in quotes since the book and trade files are updated simultaneously and in real time. Therefore, we avoid the use ad hoc rules to match trades and quotes, like the classical “five second rule” applied to NYSE data.<sup>11</sup>

NYSE data is obtained from the TAQ database. We consider two different sample periods, January to March 1996 and 2000. Several details in the mi

<sup>10</sup> Our model is nonlinear and well behaved around the mean (nonlinear). Like standard linear probability models (LPM), only in the extremes it can give predictions out of the zero and one interval. The corresponding estimation theory for dynamic models with weakly dependent variables is covered in White (1994) and Wooldridge (1994). Park and Phillips (2000) extend it to cover nonlinear cointegration cases with limited dependent variables.

<sup>11</sup> Blume and Goldstein (1997) shows that the “five second rule” could not be generalized to all sample periods and markets. However, Odders White (2000) shows that this rule does not seem to explain much of the bias induced by the Lee and Ready’s (1991) algorithm.

crostructure of the NYSE changed from 1996 to 2000. Particularly interesting for the purposes of this paper is the progressive decrease of the minimum price variation or tick, from US\$1/8 in 1996 to 1 cent in 2000. Jang and Venkatesh (1991) remarks that symmetric responses of ask and bid quotes are impaired by the discreteness of quote changes. Therefore, a small tick should decrease the probability of observing asymmetric adjustments of ask versus bid quotes after trade related shocks. Trading activity in the NYSE has sharply increased since 1996. For example, from January to March 1996 IBM transacted 130,620 times; during the same interval in 2000 the number of trades was 234,766. For GE, the number of trades increased from 106,347 in 1996 to 350,795 in 2000. By considering these two NYSE subsamples, we have the opportunity to check whether microstructure and trading activity changes have influenced the dynamical relationship between trades and quotes.

The NYSE sample includes the 11 most frequently traded stocks in 1996 and 2000, respectively, excluding stocks that experienced splits. The complete set of stocks is listed in Appendix II. We consider trades from both the primary market (NYSE) and regional markets. However, we only keep NYSE quotes because the evidence suggests that regional quotes only follow with some delay those of the primary market (Blume and Goldstein 1997).<sup>12</sup> Trades not codified as “regular trades”, such as trades out of sequence or reported with error, have been discarded. Trades from the same market, with the same price, and with the same time stamp are treated as just one trade. All quote and trade registers prior to the opening and after the close are dropped. The overnight changes in quotes are treated as missing values. Quotes with bid ask spreads lower than or equal to zero or quoted depth equal to zero have also been eliminated. After these adjustments, around 3% of all trades have been eliminated. Finally, we follow Blume and Goldstein (1997) in deleting quoted spreads that exceed 20% of the quote midpoint, and quote updates that exceed 50% of the prior quote. Prices and quotes are coupled using the “five second rule” (Lee and Ready 1991). This rule assigns to each trade the first quote stamped at least five before the trade itself.

A trade is classified as buyer (seller) initiated when the transaction price is closer to the ask (bid) price than to the bid (ask) price. Trades with price equal to the quote midpoint are not classified. The trade indicator  $x_t^B$  ( $x_t^S$ ) equals one for buys (sells) and zero otherwise; for a midpoint trade, both indicators equal zero. A change in quotes, either  $\Delta a_t$  or  $\Delta b_t$ , is the difference between the quote prevailing right before the  $t$  th trade takes place and the quote prevailing right before the next trade in time.

The SSE database contains the 11 most frequently traded stocks in 2000, listed in Appendix II. Spanish data is supplied by the SSE Interconnection System (SIBE). We retrieve trades and quotes from July to September 2000 because data from January to March is not available. We apply the same filters as for the US

---

<sup>12</sup> Hasbrouck (1995) concludes that the contribution of the regional markets to the price discovery process of NYSE listed stocks is negligible. Harris et al. (1995), however, observe that both the NYSE and the regional markets error correct to deviations from each other, therefore suggesting that the regional quotes do are informative. Tse (2000) compares the methodologies used in these papers. He concludes that the discrepant findings are only due to the choice of quotes (Hasbrouck) versus trade prices (Harris et al.). Tse suggests that trades in the regional markets could contribute to price discovery even when quotes were non informative. This conclusion fundamentals our choice of discarding regional quotes while keeping regional trades.

data. Quote changes and trade indicators are computed analogously to the NYSE case. Since price improvements are impossible in the SSE, there are no transaction prices inside the bid ask spread. Therefore, we do not require trade direction algorithms to classify the trades as either buys or sells.

We rely on the theoretical and empirical research in market microstructure to determine the exogenous variables to be included in the vector  $MC_t$ . Easley and O'Hara (1987) formally show that large sized trades are more informative. Empirically, Hasbrouck (1991) and Barclay and Warner (1993), among others, show that this relationship is increasing but concave. In Easley and O'Hara's (1992) model, higher trading intensity signals new information. Consistently, Easley et al. (1997) and Dufour and Engle (2000) find that shorter trade durations are associated with larger price impacts. Subrahmanyam (1997) finds that trades in the regional markets are less informative than NYSE trades, arguably because they attract liquidity motivated traders (see Bessembinder and Kaufman 1997). Price instability means uncertainty about the true value of the stock (e.g., Bollerslev and Melvin 1994). Finally, a positive (negative) order imbalance between limit orders to sell and limit orders to buy may signal an overvalued (undervalued) stock (e.g., Huang and Stoll 1994).

The following variables are defined so as to capture the relationships detailed above. The trade size ( $V_t$ ) is measured in shares. Trade durations ( $T_t$ ) are computed as the time in seconds between two consecutive trades. A dummy variable ( $M_t$ ) identifies regional trades. Order imbalance ( $OI_t$ ) is computed as the difference between ask depth and bid depth. Finally, short term volatility ( $R_t$ ) is computed as the sum of the square changes of the quote midpoint  $\sum_{k=1}^z (\Delta q_k)^2$  in a 5 min interval before each trade.<sup>13</sup>

Finally, we construct eight trading time dummies for the NYSE session: one for trades during the first half hour of trading, five for each trading hour between 10:00 A.M. and 3:00 P.M. and, finally, two for the last trading hour, divided in two half hour intervals. Similarly, for the SSE session (9:00 A.M. to 5:30 P.M.), we construct nine dummy variables: one for the first half hour of trading, another one for the second half hour, six for each trading hour between 10:00 and 5:00 P.M. and, finally, one for the last half hour.

## 5 Estimation of the baseline model for IBM in 1996

In this section, we present the details of estimating a restricted version of model Eq. (3.7) described in Section 3 where  $f_i^j(MC_t, D_t) = 1$  for all  $i$  and  $j$ . We use data on a representative NYSE stock, IBM, in 1996. In the next section, we check whether the dynamic patterns about to be reported for IBM can be generalized to other stocks, other markets, other time periods, and other model specifications, including the unrestricted model Eq. (3.7).

We consider two alternative specifications of the model. The first one uses the trade sign indicators  $x_t^B$  and  $x_t^S$  to represent the trading process. The second one uses the trade size indicators  $\tilde{x}_t^B = x_t^B \log(V_t)$  and  $\tilde{x}_t^S = x_t^S \log(V_t)$ . The polynomials in the autoregressive matrix  $A(L)$  are all truncated at lag five, as in

<sup>13</sup> This variable is not defined for trades performed during the first 5 min of trading. In these cases, we treat volatility as missing.

Hasbrouck (1991). The system is estimated by SURE, using the Feasible Generalized Least Squares (FGLS) algorithm, described, for example, in Green (1997, pp. 674–688).

Preliminary tests indicate that the following null,

$$A(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sb}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix} \quad (5.1)$$

cannot be rejected. This null means that  $\Delta a_t$  ( $\Delta b_t$ ) depends of its own lags but not on  $\Delta b_t$  ( $\Delta a_t$ ) lags. This restriction prevents for multicollinearity problems. Additionally, buys (sells) do not depend on lagged values of  $\Delta b_t$  ( $\Delta a_t$ ).

Table 1 summarizes the estimation of the baseline version of Eq. (3.7) with the restrictions in Eq. (5.1). Panel A (B) reports the estimated coefficients and the residual correlation matrix for the model with trade sign (trade size) indicators. The noise terms ( $\tilde{u}_t^a, \tilde{u}_t^b$ ) in Eq. (3.7) are positively correlated (0.4362 in Panel A and 0.4324 in Panel B). This shows that the trade unrelated shocks tend to move ask and bid quotes in the same direction. The noise terms ( $\tilde{u}_t^S, \tilde{u}_t^B$ ) are negatively correlated (0.6804 in Panel A and 0.6038 in Panel B). This shows that unexpected increases in the buy pressure are usually coupled with unexpected decreases in the sell pressure. All these correlations are statistically significant at the 1% level. The remaining cross equation correlation coefficients are statistically equal to zero. Therefore, as Appendix I suggested, the coefficients of Eqs. (3.7) (5.1) cannot be efficiently estimated equation by equation.

The dynamics of ask and bid quotes after a trade are characterized by two simultaneous effects. First, ask and bid quotes error correct after a trade. The coefficients of the lagged bid ask spread,  $s$  in Table 1, reveal that deviations between quotes induce simultaneous corrections in ask and bid prices. This dynamic effect causes the current spread to mean revert as the ask decreases and the bid increases. Therefore, the model shows that changes in the spread are transient. This dynamic effect indicates that liquidity suppliers, either market makers or limit order traders, provide liquidity when it is valuable (see Biais et al. 1995).<sup>14</sup>

Second, the estimated coefficients for the trading process,  $x^B$  and  $x^S$  in Table 1, evidence that at least on dimension of the symmetry assumption is not satisfied at all. Ask and bid prices do not move symmetrically through time. This finding generalizes the one step ahead evidence in Jang and Venkatesh (1991). Table 1 Panel A reports that after a unitary buy shock, both the ask quote and the bid quote tend to increase. However, on average, the ask price is raised an accumulated US \$0.0247 five trade time periods later. The bid price is raised a remarkably lower US \$0.0038. Similarly, after a sell both quotes tend to be revised downwards. Nonetheless, the accumulated decrease in the ask price after five trade time intervals is US\$0.0010 while the bid price decreases a far larger US\$0.0190.

<sup>14</sup>The adjustment path that leads to the long run equilibrium between the ask price and the bid price is not necessarily linear. Following Escribano and Granger (1998), we have replaced the linear error correction term in Eq. (3.7) by a non linear one, a cubic polynomial on the contemporaneous spread,  $\eta_{i,1}^j s_t + \eta_{i,2}^j s_t^2 + \eta_{i,3}^j s_t^3$ . We find that all the coefficients are significant, indicating that the quote adjustment is faster the wider the quoted spread. However, we do not get too much improvement in terms of model adjustment.

**Table 1** The base line VEC model for IBM

Equation	Variable	Panel A: trade sign indicators	Panel B: trade size indicators
$\Delta a_t$	s	0.0537	0.0503
	$\Delta a$	0.0633	0.0285
	$x^B$	0.0247	0.0040
	$x^S$	0.0010	0.0004
$\Delta b_t$	s	0.0490	0.0541
	$\Delta b$	0.0940	0.0668
	$x^B$	0.0038	0.0007
	$x^S$	0.0190	0.0035
$x^B_t$	s	0.0934	0.6925
	$\Delta a$	2.7348	17.0895
	$x^B$	0.7042	0.6968
	$x^S$	0.1300	0.1265
$x^S_t$	s	0.1575	1.1918
	$\Delta b$	2.8268	18.0433
	$x^B$	0.1544	0.1267
	$x^S$	0.7422	0.7339
	Obs.	130,620	130,620
	R <sup>2</sup>		
	$\Delta a_t$	0.0726	0.0826
	$\Delta b_t$	0.0604	0.0722
	$x^B_t$	0.4586	0.4375
	$x^S_t$	0.5123	0.4905
Residual correlation matrix			
	$Cov(u_t^a, u_t^b)$	0.4362	0.4324
	$Cov(u_t^B, u_t^S)$	0.6804	0.6038
	$Cov(u_t^a, u_t^S)$	0.0000	0.0001
	$Cov(u_t^b, u_t^B)$	0.0003	0.0002

This table summarizes the estimation of the VEC model,

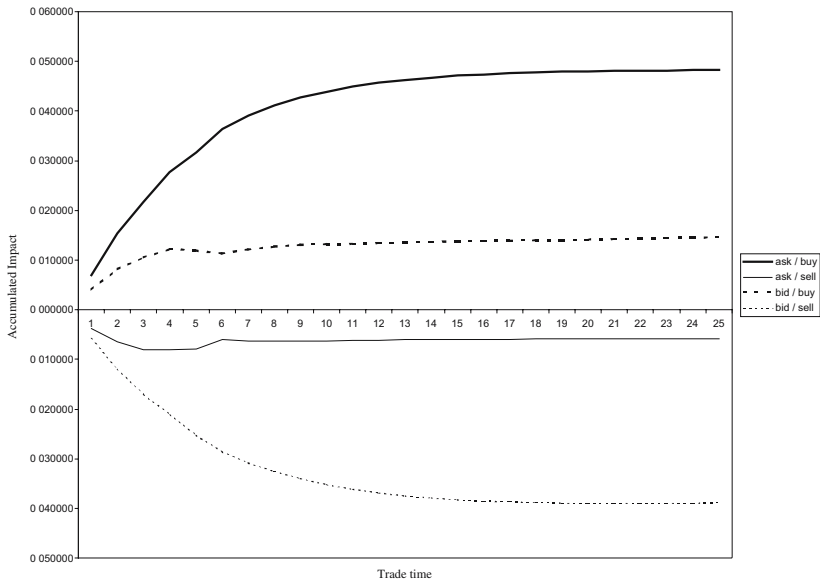
$$\begin{pmatrix} 1 & 0 & A_{aB}^* & A_{aB}^* \\ 0 & 1 & A_{bB}^* & A_{bS}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ x_t^B \\ x_t^S \end{pmatrix} = \begin{pmatrix} \gamma_{EC}^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A(L) \begin{pmatrix} \Delta a_{t-1} \\ \Delta b_{t-1} \\ x_{t-1}^B \\ x_{t-1}^S \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix} \text{ with the restrictions, } A(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sb}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix}$$

The model is defined in trade time and truncated at 5 lags. We use data for IBM from January to March 1996. The model is estimated by SURE. We report, for each variable, the sum of all lags whenever the coefficients are statistically significant at the 1% level. We also provide the R<sup>2</sup> for each equation in the system and information about the residual correlation matrix. Panel A uses trade sign indicators to characterize the trading process. The trade sign indicator  $x_t^B$  ( $x_t^S$ ) equals 1 for buys (sells) and zero otherwise. For midpoint trades both variables equal zero. The error correction term is the bid ask spread;  $\Delta a_t$  ( $\Delta b_t$ ) is the change in the ask (bid) quote between two consecutive trades. Panel B replaces the trade sign indicators by the following trade size indicators,  $V_t^a$  and  $V_t^b$ , where  $V_t^i$  is the trade size in number of shares

Figure 1 represents the impulse response function (IRF) of the model in Panel A derived by dynamic simulation. These curves represent the responses of ask and bid prices after both a unitary buyer initiated shock (increasing curves) and a

unitary seller initiated shock (decreasing curves). These trade related shocks occur after a steady state period characterized by constant quotes, no trades, and a null bid ask spread. The IRF measure the long run impact of a particular trade related shock on both quotes when the whole dynamic structure of the model is taken into account. Always on average terms, buys have a larger impact on the ask quote and sells have a larger impact on the bid quote. Statistical tests performed over the estimated VEC model corroborate that these differences are statistically significant. Briefly, quotes tend to be revised in the same direction but not by the same amount after a trade. This result suggests that ask (bid) quotes may lead the adjustment of the quoted prices after a buy (sell) shock.

Table 1 evidences two opposite and simultaneous effects associated with trade related shocks on the time series dynamics ask and bid quotes. First, trade related



**Fig. 1** Baseline VEC model for IBM: Impulse Response Function after an unexpected unitary trade related shock. This figure displays the Impulse Response Function (IRF) of ask and bid quotes to an unexpected unitary buyer initiated shock (increasing paths) and seller initiated shock (decreasing paths) according to the following VEC model, estimated using IBM data from January to March 1996,

$$\begin{pmatrix} 1 & 0 & A_{aB}^* & A_{aB}^* \\ 0 & 1 & A_{bB}^* & A_{bS}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ x_t^B \\ x_t^S \end{pmatrix} = \begin{pmatrix} \gamma_a^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A(L) \begin{pmatrix} \Delta a_{t-1} \\ \Delta b_{t-1} \\ x_{t-1}^B \\ x_{t-1}^S \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix} \quad \text{with the}$$

restrictions,  $A(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sb}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix}$  The model is defined in trade time

and truncated at 5 lags. We use data for IBM from January to March 1996. The initial shock is simulated after a steady state characterized by no trades, no changes in quotes, and a zero bid ask spread



shocks induce positive cross-serial correlation as both quotes tend to be adjusted in the same direction. Second, a trade-related shock tends to increase the bid ask spread as ask and bid quotes adjust asymmetrically. This effect sets in motion the error correction mechanism, which causes negative cross-serial correlation between quotes. The first one is an information-motivated effect. The second one is a liquidity-motivated effect. Our VEC model is able to identify and separate these two effects. A model that would summarize the quote dynamics through the quote midpoint, however, would confound them. Our model therefore supports Biais et al. (1995) intuition that there is additional information in analyzing the dynamics of ask and bid prices jointly rather than averaging them through the quote midpoint.

Finally, we use classical Wald tests (e.g., Davidson and MacKinnon 1993) on the coefficients of the two estimated VEC models for IBM to look for asymmetries in the average price impact of buys and sells. Namely, we compare the accumulated coefficients of  $x_t^B$  in the  $\Delta a_t$  equation with the accumulated coefficients of  $x_t^S$  in the  $\Delta b_t$  equation. In absolute terms, the impact of a buy shock on the ask price is larger than an equivalent impact of a sell shock on the bid price. In Panel A,  $\tilde{A}_{aB}(1) = 24.72$  vs.  $\tilde{A}_{bS}(1) = -18.95$  estimated coefficients multiplied by  $10^3$ . Similarly, the average response of the bid price to a buy shock,  $\tilde{A}_{bB}(1) = 3.822$ , is statistically larger than the response of the ask price to a sell shock,  $\tilde{A}_{aS}(1) = -0.551$ . The results for the model in Panel B are similar. Based on this test, we should conclude that on average, buyer-initiated shocks for IBM were more informative than similar seller-initiated shocks. Nonetheless, this test does not take into account the complete dynamics between trades and quotes captured in the VEC model (Eq. 3.7). In next sections, we will perform a more precise test based on IRFs to compare the information content of buys and sells.

The dynamics of the trading process are not the focus of this paper, but they show the patterns previously reported in other studies (e.g., Hasbrouck 1991). Particularly relevant is the strong positive autocorrelation in signed trades. Purchases are more likely followed by new purchases and sales are more likely followed by additional sales. Clusters of signed trades may be explained by traders successively reacting to new information, informed traders strategically splitting orders so as to ameliorate the price impact, imitative behavior among different traders, etc. Unfortunately, our model does not help in discerning the appropriate explanation.

Previous findings are unaltered when we control for intra-daily regularities by letting  $f_t^j(MC_t, D_t) = 1 + \sum_{h \neq A} \gamma_h^{i,j} D_t^h$ .

## 6 Robustness

In this section, we perform several robustness analyses to assess the degree of generality of the results obtained for IBM in the previous section. In the first subsection, we summarize the estimation of the baseline model for the remaining stocks in Appendix II. In the second subsection, we summarize the estimation of the unrestricted model Eq. (3.7). In the third subsection, we simulate the dynamics of the unrestricted model Eq. (3.7) to get a more precise understanding of the asymmetries evidenced in preceding subsections. Finally, in the fourth subsection we consider the effect of time aggregation.

## 6.1 Estimation of the baseline model for the complete sample

Table 2 summarizes the estimation of the baseline model for the three sets of stocks described in Appendix II. To be concise, we only report the results for the model with trade size indicators. Table 2 contains the average coefficients across the 11 stocks in each subsample. In addition, it includes the number of stocks for which the aggregated coefficients for a given variable in the model are significant and positive/negative.<sup>15</sup>

In general, the results are highly consistent with those previously reported for IBM'96: (a) Ask and bid quotes error correct, and (b) buys (sells) have a larger impact on the ask (bid) quote than sells (buys); thus, ask and bid quotes react asymmetrically to a trade related shock. Table 2 corroborates the existence of the simultaneous information related effects and liquidity related effects associated trade related shocks that were reported for IBM'96 in Section 5. Notice that these two effects coexist because of the asymmetric adjustments of bid and ask quotes after trades.

As with IBM'96, we have simulated unitary buyer and seller initiated shocks on the estimated VEC models to obtain the IRF of ask and bid quotes. Table 3 reports the main findings. Namely, we test the null of equality of the absolute IRFs of (a) ask and bid quotes after a unitary buyer initiated shock; (b) ask and bid quotes after a unitary seller initiated shock, and (c) ask quote after a unitary buyer initiated shock and bid quote after a unitary seller initiated shock. The two first nulls are strongly rejected. As previously shown in Table 2, buys (sells) have a larger impact on the ask (bid) quotes. The third hypothesis, however, is only accepted for the NYSE data, at the 5% for the 1996 sample and at the 10% level for the 2000 sample. Therefore, Table 3 shows that buys are more informative than sells in the NYSE, but the same is not true for the SSE.<sup>16</sup>

## 6.2 Estimation of the unrestricted VEC model

We proceed now with the estimation of the unrestricted VEC model Eq. (3.7). We test whether the asymmetries evidenced with the base line model persist once we add additional structure to the model. We consider the model with trade size indicators. Thus, the function  $f_t^j(MC_t, D_t)$  includes trade durations ( $T_t$ ), the dummy for regional market trades ( $M_t$ ), the order imbalance ( $OI_t$ ), and the short term volatility ( $R_t$ ). These exogenous variables interact with the trade size indicators, introducing non linear patterns in the impact of trades on ask and bid quotes. We maintain the restrictions in Eq. (5.1), now applied to the time variant autoregressive matrix  $A_t(L)$ .

---

<sup>15</sup> More detailed results are available from the authors upon request.

<sup>16</sup> In order to gauge the importance of modelling the trading processes, we have performed the same exercise but considering only the dynamics captured by the quote equations. We find that, on average, for the NYSE'00 sample, the impact of a unitary buyer initiated shock is underestimated by a 65%. Similarly, the impact of a seller initiated shock is underestimated by a 59%. Similar percentages are found for the other subsamples. This evidences the relevance of considering the complete set of dynamical interactions and fee backs between trades and quotes.

Estimated coefficients are not reported because of space limitations, but they are consistent with theoretical predictions. They are also regular across markets.<sup>17</sup> We have already shown in previous sections that a larger trade size increases the price impact of trades. In addition, a buy (sell) of any size executed in a high volatile period, as measured by  $R_t$ , has a larger impact on the ask (bid) quote. The trading activity in the regional markets is less informative than in the NYSE; both buys and sells have a lower impact on quotes when they are worked through the regional venues. A positive order imbalance on the book, that is, more volume on the offer side than on the demand side, decreases (increases) the impact of an incoming buy (sell) on quotes. Finally, shorter durations increase the impact of buys (sells) on the ask (bid) quote, though this relationship is the weakest.

Next, we show that the asymmetries between ask and bid responses to trade related shocks evidenced with the baseline model persist with this more complex specification. As in the previous subsection, we use the estimated coefficients of the unrestricted VEC model to simulate the impact of unitary trade related shocks on ask and bid quotes. Also in this case, shocks occur after a steady state characterized by no trades, no changes in quotes, and a zero bid-ask spread. In this analysis, we are interested in the linear effect of a trade in quotes; hence, the exogenous variables are set equal to zero. We will investigate the consequences of altering the level of the exogenous variables in the next subsection.

Table 4 summarizes our findings. Compared with Tables 3, 4 not only corroborates the asymmetries observed with the baseline model, but it reinforces them since the statistical tests provide stronger support to the alternative hypothesis that NYSE buys are more informative than sells. This hypothesis is this time rejected at the 1% level for the NYSE'96 subsample and at the 5% level for the NYSE'00 subsample. For the SSE, however, the null of equal informativeness of buys and sells still cannot be rejected.

### 6.3 A closer look to the asymmetry assumption

In this subsection, we obtain the responses of ask and bid quotes to trade related shocks using model Eq. (3.7) when we let the level of the variables in  $MC_t$  to vary. The goal is to obtain additional insights on the asymmetries evidenced in previous subsections. We consider the model with the trade size indicators  $\tilde{x}_t^B$  and  $\tilde{x}_t^S$ .

We proceed as follows. As in previous simulation exercises, an unexpected trade happens after a steady state period with no prior trades, stable quotes, and zero spreads. For each exogenous variable, we compute the 25, 75, and 95% percentiles of its stock specific empirical distribution. These values define three different levels of the variable: small (S), medium (M), and large (L) respectively. We assume that each variable in  $MC_t$  follows a general probabilistic process, exogenous to the VEC model Eq. (3.7), that we approximate by an AR( $p$ ) model.<sup>18</sup> This model is estimated

<sup>17</sup> These results are available upon request from the authors.

<sup>18</sup> For the regional dummy, we simply compare the impact of a regional trade with the impact of a NYSE trade. The auto regressive order  $p$  is determined using likelihood ratio tests, starting with  $p=7$ .

**Table 2** The base-line VEC model for the entire sample

Equation	Variable	NYSE Sample 1996				NYSE Sample 2000				SSE Sample 2000			
		Coefficients average	Std	(+)	(-)	Coefficients average	Std	(+)	(-)	Coefficients average	Std	(+)	(-)
$\Delta a_t$	s	-0.0627	0.0234	0	11	-0.0401	0.0201	0	11	-0.0795	0.0298	0	11
	$\Delta a$	-0.0639	0.0568	2	9	0.0298	0.0475	9	2	-0.1935	0.1255	0	10
	$x^B$	0.0038	0.0015	11	0	0.0034	0.0022	11	0	0.0011	0.0010	11	0
$\Delta b_t$	$x^S$	-0.0004	0.0006	3	7	-0.0013	0.0012	0	11	-0.0004	0.0005	1	10
	s	0.0541	0.0226	11	0	0.0438	0.0187	11	0	0.1016	0.0140	11	0
	$\Delta b$	-0.0558	0.0629	2	9	0.0183	0.0412	8	3	-0.1834	0.0669	0	10
$x^B_t$	$x^B$	0.0002	0.0005	6	4	0.0006	0.0006	10	0	0.0004	0.0004	11	0
	s	-0.0040	0.0019	0	11	-0.0040	0.0027	0	11	-0.0012	0.0012	0	11
	$\Delta a$	0.3611	1.2205	8	2	2.0670	1.1449	11	0	-2.6797	9.3107	4	7
$x^S_t$	$\Delta a$	-34.9566	15.6377	0	11	-18.4364	13.8204	0	11	-79.9208	63.2079	0	11
	$x^B$	0.7402	0.0457	11	0	0.6917	0.0625	11	0	0.7300	0.0796	11	0
	$x^S$	0.1702	0.0297	11	0	0.2086	0.0410	11	0	0.2509	0.0574	11	0
$\Delta b_t$	s	1.2202	0.8637	10	1	1.8800	1.2852	11	0	15.5386	17.0832	10	1
	$\Delta b$	34.6731	16.7873	11	0	16.9249	14.1931	11	0	87.8538	70.3804	11	0
	$x^B$	0.1351	0.0429	11	0	0.1671	0.0360	11	0	0.2124	0.0511	11	0
$x^S$	0.6935	0.0653	11	0	0.6325	0.0454	11	0	0.7058	0.0575	11	0	

**Table 2** (continued)

Equation	Variable	NYSE Sample 1996		NYSE Sample 2000		SSE Sample 2000		
		Coefficients average	Std	(+) (-)	(-) (+)	Coefficients average	Std	(+) (-)
$R^2$	$\Delta a_t$	0.0826	0.0294		0.0537	0.0260	0.1444	0.0489
	$\Delta b_t$	0.0802	0.0299		0.0555	0.0263	0.1478	0.0493
	$x_t^B$	0.5157	0.0658		0.4860	0.0547	0.4845	0.1212
	$x_t^S$	0.4621	0.0798		0.3977	0.0298	0.5241	0.1128

This table summarizes the estimation of the VEC model,

$$\text{restrictions, } A(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sb}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix}$$

The model is defined in trade time and truncated at 5 lags. We use data on 11 NYSE-listed stocks from January to March 1996, 11 NYSE-listed stocks from January to March 2000, and 11 SSE stocks from July to September 2000. The model is estimated by SURE. We report, for each sample and variable, the cross-sectional average of the sum of all lags whenever the coefficients are statistically significant at the 1% level. We also provide the cross-sectional average  $R^2$  for each equation in the system. Finally, we include the number of stocks for which the coefficient of the corresponding variable is statistically positive/negative at the 1% level. The error-correction term is the bid-ask spread;  $\Delta a_t$  ( $\Delta b_t$ ) is the change in the ask (bid) quote between two consecutive trades. The trade-size indicators are  $\tilde{x}_t^B = x_t^B \log(V_t)$  and  $\tilde{x}_t^S = x_t^S \log(V_t)$ , where  $V_t$  is the trade size in number of shares, and  $x_t^B$  ( $x_t^S$ ) equals 1 for buys (sells) and zero otherwise. For midpoint trades both variables equal zero.

$$\begin{pmatrix} 1 & 0 & A_{aB}^* & A_{aS}^* \\ 0 & 1 & A_{bB}^* & A_{bS}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ \tilde{x}_t^B \\ \tilde{x}_t^S \end{pmatrix} = \begin{pmatrix} \gamma_{aC}^{EC}(L) \\ \gamma_{bC}^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} \begin{pmatrix} \Delta a_{t-1} \\ \Delta b_{t-1} \\ \tilde{x}_{t-1}^B \\ \tilde{x}_{t-1}^S \end{pmatrix} + A(L) s_{t-1} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix} \text{ with the}$$

**Table 3** Simulation of the base line VEC model for the entire sample

Absolute IRF×100				
Buy/ask vs. buy/bid		Ask	Bid	Ask bid t test
NYSE 1996	Mean	0.9749	0.1042	0.8707*
	Std.	(0.3889)	(0.0964)	(0.3241)
NYSE 2000	Mean	1.0570	0.2129	0.8441*
	Std.	(0.8160)	(0.2195)	(0.6900)
SSE 2000	Mean	0.3643	0.1239	0.2404*
	Std.	(0.4103)	(0.1644)	(0.2518)
Total	Mean	0.7987	0.1470	0.6517*
	Std.	(0.6375)	(0.1695)	(0.5374)
Sell/ask vs. sell/bid		Bid	Ask	Bid ask
NYSE 1996	Mean	0.8844	0.1024	0.7821*
	Std.	(0.3028)	(0.1102)	(0.2147)
NYSE 2000	Mean	1.0043	0.3447	0.6596*
	Std.	(0.7174)	(0.3187)	(0.4584)
SSE 2000	Mean	0.3021	0.1026	0.1994*
	Std.	(0.2695)	(0.1002)	(0.1705)
Total	Mean	0.7303	0.1832	0.5470*
	Std.	(0.5560)	(0.2283)	(0.3925)
Buy/ask vs. sell/bid		Buy/ask	Sell/bid	Difference
NYSE 1996	Mean	0.9749	0.8844	0.0905**
	Std.	(0.3889)	(0.3028)	(0.1585)
NYSE 2000	Mean	1.0570	1.0043	0.0528***
	Std.	(0.8160)	(0.7174)	(0.1193)
SSE 2000	Mean	0.3643	0.3021	0.0622
	Std.	(0.4103)	(0.2695)	(0.1719)
Total	Mean	0.8445	0.7935	0.0510**
	Std.	(0.6277)	(0.5513)	(0.1200)

This table reports statistical tests on the average impulse response functions (IRFs) of the VEC model

$$\begin{pmatrix} 1 & 0 & A_{aB} & A_{aB} \\ 0 & 1 & A_{bB} & A_{bS} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ \tilde{x}_t^B \\ \tilde{x}_t^S \end{pmatrix} \begin{pmatrix} \gamma_a^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A(L) \begin{pmatrix} \Delta a_t & 1 \\ \Delta b_t & 1 \\ \tilde{x}_t^B & 1 \\ \tilde{x}_t^S & 1 \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix}$$

with the restrictions,  $A(L) \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sa}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix}$

We compare (a) the impact of a unitary buyer initiated shock on the ask and bid quotes; (b) the impact of unitary seller initiated shock on the ask and bid quotes, and (c) the impact of a unitary buyer initiated shock on the ask quote with the impact of a unitary seller initiated shock on the bid quote. We use data on 11 NYSE listed stocks from January to March 1996, 11 NYSE listed stocks from January to March 2000, and 11 SSE stocks from July to September 2000. We report statistical tests of the null of equality of the absolute IRFs against the alternative of a positive difference. The error correction term is the bid ask spread.  $\Delta a_t$  ( $\Delta b_t$ ) is the change in the ask (bid) quote between two consecutive trades. The trade size indicators are  $\tilde{x}_t^B$   $x_t^B \log(V_t)$  and  $\tilde{x}_t^S$   $x_t^S \log(V_t)$ , where  $V_t$  is the trade size in number of shares, and  $x_t^B$  ( $x_t^S$ ) equals 1 for buys (sells) and zero otherwise. For midpoint trades both variables equal zero

\*Statistically greater than zero at the 1% level

\*\*Statistically greater than zero at the 5% level

\*\*\*Statistically greater than zero at the 10% level

by Generalized Least Squares (GLS), controlling for deterministic intraday patterns. The  $AR(p)$  models are used in the simulation exercise to generate the future values of each exogenous variable. Then, we compute the response of ask and bid quotes after a unitary trade related shock (either buyer or seller initiated) conditional on the level of one of the exogenous variables in  $MC_t$  while the others are kept equal to zero.

We compare the IRF after a unitary shock when all the exogenous variables are zero (Table 4) with the same IRF when the level of a given exogenous variable increases from zero to S, from zero to M, and from zero to L. Table 5 reports the relative change of the absolute value of the IRF 500 periods after the shock. We also provide the number of stocks in each subsample for which the impact is positive/negative and significantly different from zero.

Table 5 shows that asymmetries after regional trades are less important than after NYSE trades. As previously indicated, the positive (negative) impact of a buy (sell) shock on the ask (bid) quote is weaker when it comes from the regional markets. Moreover, bid (ask) quotes are not usually altered after an unexpected regional buy (sell). Therefore, after a regional trade ask and bid quotes do not move symmetrically either, but the asymmetry is less remarkable than after a NYSE trade.

Table 5 also evidences that the probability of observing asymmetric adjustments of ask and bid quotes decreases as volatility increases. For the NYSE'96 sample, the impact of a buyer initiated shock on the ask quote is 3.41% larger when volatility is high (L). The impact on the bid quote, however, is 50.84% larger. Similarly, the impact of a seller initiated shock on the bid quote is 5.38% larger in the more volatile scenario, but the impact on the ask quote is 42.44% larger. A similar finding is found for the SSE'00 sample. Therefore, the adjustments of ask and bid quotes after a trade related shock are more balanced in periods of high volatility. This finding would suggest that trades executed during volatile periods transmit more unambiguous signals, since they cause ask and bid quotes to be adjusted symmetrically more often than usual.

The results for trade durations and order imbalances are not conclusive.

## 6.4 Time aggregation

So far, we have shown that asymmetries exist when ask and bid quotes measured in trade time respond to trade related shocks. In this section, we study whether these asymmetries persist when we consider different time scales. In particular, we aggregate our time series of quotes and trades into 1 min and 5 min intervals. Thus, a change in quotes is now given by the difference between the final and the initial quote in each time interval. Similarly, the buyer (seller) initiated volume is the sum of the size of all buys (sells) executed during each time interval. Finally, the bid ask spread is given by the posted quotes at the end of each time interval. We estimate the baseline model Eq. (3.7) with the restrictions in Eq. (5.1) with  $t$  meaning either a 1 min or a 5 min interval. The model is truncated at three lags and estimated by SURE.

Table 6 provides our findings for the NYSE'00 sample. For the NYSE'96 and SSE'00 samples, results are similar and available upon request. Table 6 shows that the dynamics observed in trade time remain in these alternative scales. Ask and bid quotes error correct to deviations between them, causing the bid ask spread to

revert towards narrow levels. Ask (bid) quotes are also more sensible than bid quote to buys (sells) with aggregated data. In addition, buys are more informative than sells at the 5% level in the 1 min frequency and at the 1% level at the 5 min periodicity. This suggests that the asymmetry in ask and bid responses to trades is not just a high frequency phenomena.

## 7 Conclusions

This paper has introduced a new econometric approach to jointly model the time series dynamics of the trading process and the revisions of ask and bid prices. This model represents a generalization of the VAR model introduced by Hasbrouck (1991). We use this approach to check a very common theoretical assumption among microstructure models: the symmetry assumption. The symmetry assumption asserts that ask and bid quotes respond symmetrically to trades, that ask and bid quotes are posted symmetrically about the efficient price, and that buys and sells are equally informative.

Our model accommodates (not imposes) asymmetric responses of ask and bid prices to trade related shocks. It also captures asymmetric impacts of buyer and seller initiated trades. This is possible because it incorporates the co integration relationship between the ask price and the bid price, because buys and sells are generated by idiosyncratic but mutually dependent processes, and because these trading processes are endogenous. The properties of the empirical model are derived directly from a structural dynamic model for ask and bid prices. The model is estimated using data from two different markets, the NYSE and the SSE.

**Table 4** Simulation of the unrestricted VEC model

Absolute IRF×100				
Buy/ask vs. buy/bid		Ask	Bid	Ask bid t test
NYSE 1996	Mean	1.0279	0.1423	0.8855*
	Std.	(0.3546)	(0.1247)	(0.2854)
NYSE 2000	Mean	1.1401	0.4990	0.6411*
	Std.	(0.8348)	(0.3706)	(0.5025)
SSE 2000	Mean	0.2843	0.0549	0.2294*
	Std.	(0.2840)	(0.0832)	(0.2350)
Total	Mean	0.8174	0.2321	0.5853*
	Std.	(0.6565)	(0.2966)	(0.4441)
Sell/ask vs. sell/bid		Bid	Ask	Bid ask
NYSE 1996	Mean	0.9170	0.1741	1.0911*
	Std.	(0.3180)	(0.1302)	(0.4404)
NYSE 2000	Mean	0.9945	0.5595	1.5540*
	Std.	(0.6248)	(0.4194)	(1.0304)
SSE 2000	Mean	0.2412	0.0518	0.2930*
	Std.	(0.2087)	(0.0573)	(0.2575)
Total	Mean	0.7176	0.2618	0.9794*
	Std.	(0.5341)	(0.3310)	(0.8324)



**Table 4** (continued)

Absolute IRF×100				
Buy/Ask vs. Sell/Bid		Buy/Ask	Sell/Bid	Difference
NYSE 1996	Mean	1.0279	0.9170	0.1108*
	Std.	(0.3546)	(0.3180)	(0.1703)
NYSE 2000	Mean	1.1401	0.9945	0.1456**
	Std.	(0.8348)	(0.6248)	(0.2357)
SSE 2000	Mean	0.2843	0.2412	0.0431
	Std.	(0.2840)	(0.2087)	(0.1126)
Total	Mean	0.8174	0.7176	0.0998*
	Std.	(0.6565)	(0.5341)	(0.1796)

This table reports statistical tests on the impulse response functions (IRFs) of the VEC model,

$$\begin{pmatrix} 1 & 0 & A_{aB,t}^* & A_{aB,t}^* \\ 0 & 1 & A_{bB,t}^* & A_{bS,t}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ \tilde{x}_t^B \\ \tilde{x}_t^S \end{pmatrix} = \begin{pmatrix} \gamma_a^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A_t(L) \begin{pmatrix} \Delta a_{t-1} \\ \Delta b_{t-1} \\ \tilde{x}_{t-1}^B \\ \tilde{x}_{t-1}^S \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix} \text{ with}$$

the restrictions,  $A_t(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB,t}(L) & A_{aS,t}(L) \\ 0 & A_{bb}(L) & A_{bB,t}(L) & A_{bS,t}(L) \\ A_{Ba}(L) & 0 & A_{BB,t}(L) & A_{BS,t}(L) \\ 0 & A_{Sb}(L) & A_{SB,t}(L) & A_{SS,t}(L) \end{pmatrix}$

We compare (a) the impact of a unitary buyer initiated shock on the ask and bid quotes; (b) the impact of unitary seller initiated shock on the ask and bid quotes, and (c) the impact of a unitary buyer initiated shock on the ask quote with the impact of a unitary seller initiated shock on the bid quote. We use data on 11 NYSE listed stocks from January to March 1996, 11 NYSE listed stocks from January to March 2000, and 11 SSE stocks from July to September 2000. We report statistical tests of the null of equality of the absolute IRFs against the alternative of a positive difference. We keep all the polynomials in the autoregressive matrix constant during the simulation, even when they are time variant due to exogenous variables. The error correction term is the bid ask spread;  $\Delta a_t$  ( $\Delta b_t$ ) is the change in the ask (bid) quote between two consecutive trades;  $\tilde{x}_t^B$   $x_t^B \log(V_t)$ , and  $\tilde{x}_t^S$   $x_t^S \log(V_t)$ , where  $V_t$  is the trade size in number of shares, and  $x_t^B$  ( $x_t^S$ ) equals 1 for buys (sells) and zero otherwise. For midpoint trades both variables equal zero

\*Statistically greater than zero at the 1% level

\*\*Statistically greater than zero at the 5% level

\*\*\*Statistically greater than zero at the 10% level

The dynamics of ask and bid prices are characterized by two findings, robust across markets, trading periods, and model specifications. First, we show that these quotes do not follow symmetric patterns after trades. Ask and bid prices tend to be revised in the same direction but not by the same amount. Ask (bid) quotes are more sensible to buyer (seller) initiated shocks than bid (ask) quotes. We evidence, however, that the likelihood of symmetric responses increases with volatility. In addition, we show that asymmetries are less frequent in the NYSE after regional trades. In addition to the former information-motivated trade-related effect, we also observe a liquidity-motivated trade-related effect. Ask and bid quotes error correct to mutual deviations, which causes a strong mean reversion in the bid ask spread. These two findings produce simultaneous but opposite effects in the dynamics of ask and bid prices: information-induced positive cross-serial correlation and liquidity-induced negative cross-serial correlation.

**Table 5** Impulse-response functions conditional on the exogenous variables

Variable	Bid																			
	Ask								Bid											
	Buy shock				Sell shock				Buy shock				Sell shock							
	M	L	I	D	S	M	L	I	D	S	M	L	I	D	S	M	L	I	D	
NYSE 1996																				
$T_t$	-0.39	-1.71	-4.36	2	9	0.43	2.00	5.74	2	9	0.61	3.17	8.37	8	2	-0.29	-1.40	-3.61	1	10
$R_t$	0.18	1.35	3.41	10	1	3.98	21.88	50.84	8	2	2.53	16.98	42.44	9	1	0.29	1.97	5.38	9	2
$OI_t$	1.40	-2.53	-10.06	0	11	1.99	0.95	8.05	7	3	2.19	-0.80	-3.35	3	7	-1.29	2.31	9.95	11	0
$M_t$	-52.75	0	11	11	11	-93.26	0	3	3	3	-96.95	0	2	2	2	-49.66	0	11	11	11
NYSE 2000																				
$T_t$	-0.75	-2.10	-4.49	2	8	-0.47	-1.28	-2.96	8	2	-1.14	-3.18	-7.21	1	9	0.25	1.05	2.51	8	3
$R_t$	0.61	1.83	3.96	10	0	0.42	1.21	2.64	6	5	0.77	2.14	5.19	6	4	0.63	1.97	4.36	9	2
$OI_t$	0.47	-1.22	-5.06	1	10	-0.68	1.81	7.61	10	1	0.82	-2.19	-8.55	0	11	0.94	3.94	4.36	9	2
$M_t$	-47.77	0	11	11	11	-92.11	0	2	2	2	-96.01	0	1	1	1	-56.47	0	11	11	11
SSE 2000																				
$T_t$	-0.17	-1.16	-3.10	3	7	-3.75	-16.54	-5.76	1	8	0.45	2.67	7.98	7	3	-0.08	0.93	12.95	4	5
$R_t$	0.60	4.42	14.88	11	0	1.01	7.51	32.75	8	2	2.23	16.11	53.31	10	0	0.47	3.60	13.31	11	0
$OI_t$	0.65	-0.78	-3.49	0	11	-0.86	1.05	4.49	9	1	3.23	-3.42	-13.66	0	10	-0.28	0.33	1.41	9	2
$M_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

**Table 5** (continued)

Variable	Ask						Bid													
	Buy shock			Sell shock			Buy shock			Sell shock										
	$\Delta$ abs(IRF) (%)	Signif (1%)	$\Delta$ abs(IRF) (%)	Signif (1%)	$\Delta$ abs(IRF) (%)	Signif (1%)	$\Delta$ abs(IRF) (%)	Signif (1%)	$\Delta$ abs(IRF) (%)	Signif (1%)	$\Delta$ abs(IRF) (%)	Signif (1%)								
NYSE 1996	S	M	L	I	D	S	M	L	I	D	S	M	L	I	D					
All stocks																				
$T_t$	-0.44	-1.66	-4.00	7	24	-1.15	-4.76	-0.90	11	19	-0.03	0.89	3.05	16	14	-0.04	0.15	3.37	13	18
$R_t$	0.46	2.56	7.53	31	1	1.65	9.31	26.21	22	9	1.84	11.74	33.64	25	5	0.46	2.51	7.69	29	4
$OL_t$	0.84	-1.51	-6.20	1	32	0.12	1.29	6.75	26	5	3.80	-1.04	-4.68	3	28	-0.65	1.19	5.10	29	4
$M_t$	-49.80		0	22		-92.80		0	5		-96.64		0	3		-54.12		0	22	

This table reports the result of simulating a unitary trade-related shock using the VEC model, 
$$\begin{pmatrix} 1 & 0 & A_{aa,t} & A_{ab,t} & A_{ac,t} & A_{ad,t} & A_{ae,t} & A_{af,t} & A_{ag,t} & A_{ah,t} & A_{ai,t} & A_{aj,t} & A_{ak,t} & A_{al,t} & A_{am,t} & A_{an,t} & A_{ao,t} & A_{ap,t} & A_{aq,t} & A_{ar,t} & A_{as,t} & A_{at,t} & A_{au,t} & A_{av,t} & A_{aw,t} & A_{ax,t} & A_{ay,t} & A_{az,t} \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ \Delta c_t \\ \Delta d_t \\ \Delta e_t \\ \Delta f_t \\ \Delta g_t \\ \Delta h_t \\ \Delta i_t \\ \Delta j_t \\ \Delta k_t \\ \Delta l_t \\ \Delta m_t \\ \Delta n_t \\ \Delta o_t \\ \Delta p_t \\ \Delta q_t \\ \Delta r_t \\ \Delta s_t \\ \Delta t_t \\ \Delta u_t \\ \Delta v_t \\ \Delta w_t \\ \Delta x_t \\ \Delta y_t \\ \Delta z_t \end{pmatrix} = \begin{pmatrix} \gamma_{EC}^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B^{EC}(L) \\ \gamma_S^{EC}(L) \end{pmatrix} s_{t-1} + A_t(L)$$

with the restrictions,  $A_t(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{ab,t}(L) & A_{ac,t}(L) & A_{ad,t}(L) & A_{ae,t}(L) & A_{af,t}(L) & A_{ag,t}(L) & A_{ah,t}(L) & A_{ai,t}(L) & A_{aj,t}(L) & A_{ak,t}(L) & A_{al,t}(L) & A_{am,t}(L) & A_{an,t}(L) & A_{ao,t}(L) & A_{ap,t}(L) & A_{aq,t}(L) & A_{ar,t}(L) & A_{as,t}(L) & A_{at,t}(L) & A_{au,t}(L) & A_{av,t}(L) & A_{aw,t}(L) & A_{ax,t}(L) & A_{ay,t}(L) & A_{az,t}(L) \end{pmatrix}$  when we alter the level of the exogenous variables in the autoregressive matrix. The exogenous variables are: trade durations ( $T_t$ ), short-term volatility ( $R_t$ ), order imbalance ( $OL_t$ ), and a dummy variable that identifies regional market trades ( $M_t$ ). The error-correction term is the bid-ask spread. The endogenous variables are the change in the ask quote ( $\Delta a_t$ ) and the bid quote ( $\Delta b_t$ ), and trade-size indicators for buys ( $\hat{x}_t^B$ ) and for sells ( $\hat{x}_t^S$ ).

We use data on 11 NYSE-listed stocks from January to March 1996, 11 NYSE-listed stocks from January to March 2000, and 11 SSE stocks from July to September 2000. For each stock, we compute the response of ask and bid prices to both a buyer-initiated shock and a seller-initiated shock, 500 periods ahead. The impact depends on the level of several exogenous variables that feature the trade and the market conditions. Three levels of each variable are considered: “small” (S), “medium” (M), and “large” (L), obtained from the 25%, 75%, and 95% percentiles of the empirical distribution. For the  $M_t$  dummy, we compare the impact of a NYSE trade with the impact of a regional trade. This table reports the average relative change (in %) across stocks in the total impact of a unitary trade-related shock when the exogenous variable increases from zero to S, from zero to M, and from zero to L. We also provide the number of stocks in each subsample for which the absolute IRF significantly increases/decreases as we increase the level of the corresponding exogenous variable.

We also show that NYSE buyer initiated trades are more informative than seller initiated trades, both in 1996 and in 2000. This pattern persists even when we consider different model specifications, including different time periodicities. This finding, however, cannot be generalized to the SSE case.

This paper has also evidenced that market frictions like the minimum price variation are not enough to explain the violation of the symmetry assumption. We find similar asymmetric patterns in ask and bid responses in 1996, with a US\$1/8 tick, and in 2000, with a US\$0.01 tick. In addition, asymmetries are found in markets with very different microstructures, like the NYSE and the SSE. Since the specialist contribution is less essential in frequently traded NYSE stocks (Madhavan and Sofianos 1998), it sounds interesting to extend the analysis in this paper by considering a larger sample of NYSE and SSE stocks, stratified by trading frequency. This analysis would clarify the role that market makers play in explaining the asymmetric dynamics between ask and bid quotes. Moreover, our findings suggest that ask (bid) quotes may lead the price discovery process after buyer (seller) initiated trades. An interesting topic for future research would be evaluating the relative contribution of ask and bid quotes to price discovery, conditional on variables like recent market trends, accumulated net volume, order imbalances, and so on. Finally, the intriguing finding that buys are more informative

**Table 6** The base line VEC model for aggregated data

Equation	Variable	Panel A: 1 minute intervals				Panel B: 5 minute intervals			
		Coefficients average	Std.	(+)	( )	Coefficients average	Std.	(+)	( )
	$\Delta b_t$	0.2231	0.0752	0	11	0.3290	0.1177	0	11
	$\Delta a$	0.0633	0.0556	10	1	0.0297	0.0473	10	1
	$x^B$	0.0172	0.0089	11	0	0.1316	0.0886	11	0
	$x^S$	0.0133	0.0080	0	11	0.1135	0.0746	0	11
	$\Delta b_t$	0.2608	0.0710	11	0	0.4421	0.0870	11	0
	$\Delta b$	0.0593	0.0696	9	2	0.0326	0.0529	9	2
	$x^B$	0.0121	0.0074	11	0	0.1240	0.0846	11	0
	$x^S$	0.0173	0.0089	0	11	0.1241	0.0819	0	11
$x^B_t$	$s$	2.1055	2.3079	10	1	0.0361	0.4750	10	1
	$\Delta a$	1.2400	1.2389	0	11	0.0155	0.3778	0	11
	$x^B$	0.6748	0.1089	11	0	0.6472	0.1317	11	0
	$x^S$	0.2963	0.0879	11	0	0.3579	0.1307	11	0
$x^S_t$	$s$	1.9812	2.0402	10	1	0.1100	0.6099	10	1
	$\Delta b$	0.4601	1.8280	4	7	0.0033	0.1560	4	7
	$x^B$	0.3985	0.1223	11	0	0.4275	0.1433	11	0
	$x^S$	0.5396	0.1476	11	0	0.5601	0.1496	11	0
	Obs.	23,484	186			4,688	62		
$R_2$									
	$\Delta a_t$	0.1621	0.0400			0.1913	0.0725		
	$\Delta b_t$	0.1847	0.0387			0.2030	0.0730		
	$x^B_t$	0.9532	0.0222			0.9945	0.0015		
	$x^S_t$	0.9328	0.0382			0.9938	0.0015		

**Table 6** (continued)

Absolute IRF×100					
Buy/ask	buy/bid	Diference	p value t test	Diference	p value t test
Mean		1.7480	0.0017*	2.3780	0.0111**
Std.		(1.5230)		(2.9193)	
Sell/bid	sell/ask				
Mean		6.9870	0.0007*	49.1156	0.0024*
Std.		(2.7803)		(25.8798)	
Buy/ask	sell/bid				
Mean		1.9127	0.0105**	11.0851	0.0099*
Std.		(2.3180)		(13.2844)	

This table summarizes the estimation of the VEC model,

$$\begin{pmatrix} 1 & 0 & A_{aB}^* & A_{aB}^* \\ 0 & 1 & A_{bB}^* & A_{bS}^* \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta a_t \\ \Delta b_t \\ \tilde{x}_t^B \\ \tilde{x}_t^S \end{pmatrix} + \begin{pmatrix} \gamma_a^{EC}(L) \\ \gamma_b^{EC}(L) \\ \gamma_B(L) \\ \gamma_S(L) \end{pmatrix} s_{t-1} + A(L) \begin{pmatrix} \Delta a_{t-1} \\ \Delta b_{t-1} \\ \tilde{x}_{t-1}^B \\ \tilde{x}_{t-1}^S \end{pmatrix} + \begin{pmatrix} u_t^a \\ u_t^b \\ u_t^B \\ u_t^S \end{pmatrix} \text{ with the re}$$

strictions,  $A(L) = \begin{pmatrix} A_{aa}(L) & 0 & A_{aB}(L) & A_{aS}(L) \\ 0 & A_{bb}(L) & A_{bB}(L) & A_{bS}(L) \\ A_{Ba}(L) & 0 & A_{BB}(L) & A_{BS}(L) \\ 0 & A_{Sb}(L) & A_{SB}(L) & A_{SS}(L) \end{pmatrix}$

The model is defined in 1 minute (Panel A) and 5 minute intervals (Panel B), and truncated at 3 lags. We use data on 11 NYSE listed stocks from January to March 2000. The model is estimated by SURE. We report for each variable the cross sectional average of the sum of all lags whenever the coefficients are statistically significant at the 1% level. We also provide the cross sectional average  $R^2$  for each equation in the system. We include the number of stocks for which the coefficient of the corresponding variable is statistically positive/negative at the 1% level. Finally, we compare (a) the impact of a unitary buyer initiated shock on the ask and bid quotes; (b) the impact of unitary seller initiated shock on the ask and bid quotes, and (c) the impact of a unitary buyer initiated shock on the ask quote with the impact of a unitary seller initiated shock on the bid quote. We report the differences in the absolute impulse response functions (IRF) and the result of a t test on the null of equal IRFs against the alternative of a strictly positive difference. The error correction term is the bid ask spread. The endogenous variables are the change in the ask quote ( $\Delta a_t$ ) and the bid quote ( $\Delta b_t$ ) in each time interval, and trade size indicators for buys ( $\tilde{x}_t^B$ ) and for sells ( $\tilde{x}_t^S$ ) are computed as the accumulated volume of buyer initiated trades and seller initiated trades, respectively, in each time interval

than sells in the NYSE but not in the SSE suggests that microstructure differences may be playing a role. This is a possibility that deserves a more exhaustive analysis.

## 1 Appendix I

### 1.1 Derivation of the VEC model (3.7)

From Eq. (3.2)

$$[1 - \alpha_m^a L](a_t - m_t) = A_{x,t}(L)'x_t + \alpha_a^{EC}(a_{t-1} - b_{t-1}) + \varepsilon_t^a.$$

As  $0 < \alpha_m^a < 1$ ,  $\alpha(L) = [1 - \alpha_m^a L]$  is a stationary polynomial in  $L$ . Then,

$$(a_t - m_t) = \alpha(L) {}^1 A_{x,t}(L)' x_t + \alpha(L) {}^1 \alpha_a^{EC} s_{t-1} + \alpha(L) {}^1 \varepsilon_t^a. \quad (\text{A.1})$$

Let  $\Delta = (1 - L)$  be the first differencing operator. Pre multiplying in (A.1) by  $\Delta$ , and letting  $\alpha(L) {}^1 A_{x,t}(L) \Delta = \tilde{A}_{x,t}(L)$  and  $\alpha(L) {}^1 \alpha_a^{EC} \Delta = \tilde{\alpha}_a^{EC}(L)$  we obtain

$$\Delta a_t = \Delta m_t + \tilde{A}_{x,t}(L)' x_t + \tilde{\alpha}_a^{EC}(L) s_{t-1} + \theta(L) \varepsilon_t^a, \quad (\text{A.2})$$

where  $\theta(L) \varepsilon_t^a = (1 - L)(1 - \alpha_m^a L) \varepsilon_t^a$  which can be approximated by a moving average polynomial of finite order, say  $q$ ,  $\theta(L) \varepsilon_t^a \approx \tilde{\theta}(L) \varepsilon_t^a = \left(1 - \tilde{\theta}_1^a L - \tilde{\theta}_2^a L^2 - \dots - \tilde{\theta}_q^a L^q\right) \varepsilon_t^a$ . Similar expansions are made with  $\tilde{A}_{x,t}(L)$  and  $\tilde{\alpha}_a^{EC}(L)$ . Substituting Eq. (3.1) in (A.2) we have

$$\Delta a_t = \tilde{A}_{x,t}(L)' x_t + \tilde{\alpha}_a^{EC}(L) s_{t-1} + \xi_t^a. \quad (\text{A.3})$$

The error term  $\xi_t^a = \tilde{\theta}(L) \varepsilon_t^a + \lambda^B v_{2,t}^B + \lambda^S v_{2,t}^S + v_{1,t}$  has an invertible moving average (MA) representation. Inverting the MA or alternatively adding long enough dynamics of the regressors of (A.3),  $\Delta a_t$  and also  $\Delta b_t$  (since they are highly correlated), the moving average structure disappears. Therefore, Eq. (A.3) could parsimoniously be approximated by

$$\Delta a_t = \tilde{\alpha}_a^{EC}(L) s_{t-1} + A_{aa}(L) \Delta a_{t-1} + A_{ab}(L) \Delta b_{t-1} + \tilde{A}_{x,t}(L)' x_t + u_t^a. \quad (\text{A.4})$$

The errors are white noise,  $E(u_t^a) = 0$  and  $E(u_t^a, u_{t-k}^a) = 0 \forall k \neq 0$ , with the autoregressive polynomials  $A_{ij}(L)$  having all roots outside the unit circle. Let,

$$\tilde{A}_{x,t}(L)' x_t = A_{aB}^B(L) f_{aB}^B(MC_{t-1}, D_{t-1}) x_t^B + A_{aS}^S(L) f_{aS}^S(MC_{t-1}, D_{t-1}) x_t^S.$$

Equation (A.4) can now be written as

$$\begin{aligned} \Delta a_t &= \tilde{\alpha}_a^{EC}(L) s_{t-1} + A_{aa}(L) \Delta a_{t-1} + A_{ab}(L) \Delta b_{t-1} + A_{aB,t}(L) x_t^B \\ &\quad + A_{aS,t}(L) x_t^S + u_t^a, \end{aligned} \quad (\text{A.5})$$

which is the first equation of the system Eq. (3.7).

The corresponding equation for  $\Delta b_t$  is similarly obtained by repeating the previous steps for Eq. (3.3) obtaining the equivalent expression of Eq. (A.3) for  $b_t$

$$\Delta b_t = \tilde{B}_{x,t}(L)' x_t + \tilde{\alpha}_b^{EC}(L) s_{t-1} + \xi_t^b. \quad (\text{A.6})$$

Notice that  $\xi_t^a$  and  $\xi_t^b$  have a component in common ( $\lambda^B v_{2,t}^B + \lambda^S v_{2,t}^S + v_{1,t}$ ) and, therefore, they are mutually correlated. This correlation depends on the importance of the idiosyncratic components in each of the residuals. From the same arguments, we can obtain the equivalent model to (A.5) for  $\Delta b_t$ , with white noise errors

$$\begin{aligned} \Delta b_t &= \tilde{\alpha}_b^{EC}(L) s_{t-1} + A_{ba}(L) \Delta a_{t-1} + A_{bb}(L) \Delta b_{t-1} + A_{bB,t}(L) x_t^B \\ &\quad + A_{bS,t}(L) x_t^S + u_t^b. \end{aligned} \quad (\text{A.7})$$

As the errors  $u_t^a$  and  $u_t^b$  are mutually correlated and therefore efficient estimation requires at least a joint estimation of (A.5) and (A.7).

From Eq. (3.4) using (A.1) (A.2) we obtain,

$$x_t^B = \mu^B \tilde{B}_{x,t}(L)' x_{t-1} + (\mu^B \tilde{\alpha}_a^{EC}(L)L + \pi^B) s_{t-1} + \alpha(L)^{-1} \varepsilon_t^a + v_{2,t}^B = \varphi_{x,t}^B(L)' x_{t-1} + \varphi_s^B(L) s_{t-1} + \xi_t^B \quad (\text{A.8})$$

where the error term  $\xi_t^B = \alpha(L)^{-1} \varepsilon_t^a + v_{2,t}^B$  has an invertible moving average (MA) representation. As previously done, this moving average structure can be approximated by,

$$x_t^B = \alpha_B^{EC}(L) s_t + A_{Ba}(L) \Delta a_{t-1} + A_{Bb}(L) \Delta b_{t-1} + A_{BB,t}(L) x_{t-1}^B + A_{BS,t}(L) x_{t-1}^S + u_t^B \quad (\text{A.9})$$

where  $E(u_t^B) = 0$  and  $E(u_t^B, u_k^B) = 0 \forall k \neq 0$ .

The corresponding equation for  $x_t^S$  is similarly obtained by repeating the previous steps with Eq. (3.5). We first obtain,

$$x_t^S = \varphi_{x,t}^S(L)' x_{t-1} + \varphi_s^S(L) s_{t-1} + \xi_t^S \quad (\text{A.10})$$

where the error term  $\xi_t^S = \alpha(L)^{-1} \varepsilon_t^b + v_{2,t}^S$ . Following the argument stated right after Eq. (A.8), we get the last equation of the system (Eq. 3.7),

$$x_t^S = \alpha_S^{EC}(L) s_t + A_{Sa}(L) \Delta a_{t-1} + A_{Sb}(L) \Delta b_{t-1} + A_{SB,t}(L) x_{t-1}^B + A_{SS,t}(L) x_{t-1}^S + u_t^S \quad (\text{A.11})$$

Notice that Eqs. (A.9) and (A.11) have correlated errors if either  $v_{2,t}^S$  and  $v_{2,t}^B$  or  $\varepsilon_t^b$  and  $\varepsilon_t^a$  are correlated, which is a very likely event.

## 1 Appendix II

### 1.1 Sample

NYSE 1996 Stocks	Company	Observations (number of trades)
GE	General Electric Co	106,347
GT	Goodyear Tire Rubber Co	86,802
IBM	Int Business Machines Corp	130,620
JNJ	Jhonson & Johnson	64,607
KO	Coca Cola Co	72,620
MO	Phillip Morris Companies Inc	91,938
MRK	Merck & Co Inc	96,425
PG	Procter & Gamble Co	52,326
T	ATT Corp	87,882
TX	Texaco Inc	76,912
WMT	Wal Mart Stores Inc	102,660

NYSE 2000 Stocks		
AOL	America Online Inc	626,768
C	Citigroup Inc	260,149
EMC	EMC Corporation	282,196
GE	General Electric Co	350,795
IBM	Int Business Machines Corp	234,766
LU	Lucent Technologies Inc	705,948
MOT	Motorola Inc	195,067
NOK	Nokia Corp	221,005
NT	Nortel Networks Corp	239,094
PFE	Pfizer Inc	233,658
T	ATT Corp	236,662
SSE 2000 Stocks		
AMS	Amadeus Global Travel Distribution	49,824
BBVA	Banco Bilbao Vizcaya Argentaria	123,687
ELE	Endesa	80,844
IBE	Iberdrola	35,811
REP	Repsol YPF	90,213
SCH	Banco Santander Central Hispano	196,880
TEF	Telefónica	424,327
TPI	Telefónica Publicidad e Información	52,253
TPZ	Telepizza	36,880
TRR	Terra Networks	146,285
ZEL	Zeltia	88,339

---

## References

- Arranz MA, Escribano A (2000) Cointegration testing under permanent breaks: a robust extended error correction model. *Oxf Bull Econ Stat* 62:23–52
- Barclay MJ, Warner JB (1993) Stealth trading and volatility, which trades move prices? *J Financ Econ* 34:281–305
- Bessembinder H, Kaufman HM (1997) A comparison of trade execution costs for NYSE and Nasdaq listed stocks. *J Financ Econ* 46:293–319
- Biais B, Hillion P, Spatt C (1995) An empirical analysis of the limit order book and the order flow in the Paris Bourse. *J Finance* 50:1655–1689
- Blume ME, Goldstein MA (1997) Quotes, order flow, and price discovery. *J Finance* 52:221–244
- Bollerslev T, Melvin M (1994) Bid-ask spreads and volatility in the foreign exchange market. An empirical analysis. *J Int Econ* 36:355–372
- Chan LKC, Lakonishok J (1993) Institutional trades and intraday stock price behavior. *J Financ Econ* 33:173–199
- Chan LKC, Lakonishok J (1995) The behavior of stock prices around institutional trades. *J Finance* 50(4):1147–1174
- Chung KH, Van Ness BF, Van Ness RA (1999) Limit orders and the bid-ask spread. *J Financ Econ* 53:255–287
- Davidson R, MacKinnon JG (1993) *Estimation and Inference in Econometrics*. Oxford University Press
- Dufour A, Engle RF (2000) Time and the price impact of a trade. *J Finance* 55(6):2467–2498
- Easley D, O'Hara M (1987) Price, trade size, and information in securities markets. *J Financ Econ* 19:69–90
- Easley D, O'Hara M (1992) Time and the process of security price adjustment. *J Finance* 47(2):577–605
- Easley D, Kiefer NM, O'Hara M (1997) One day in the life of a very common stock. *Rev Financ Stud* 10:805–835
- Ellis K, Michaely R, O'Hara M (2000) The accuracy of trade classification rules: evidence from Nasdaq. *J Financ Quant Anal* 35:529–551
- Engle R, Granger C (1987) Co-integration and error correction: representation, estimation and testing. *Econometrica* 35:251–276



- Engle RF, Patton AJ (2004) Impacts of trades in a error correction model of quote prices. *J Financ Mark* 7:1 25
- Engle RF, Yoo BS (1991) Cointegrated economic time series: a survey with new results. In: Granger CWJ, Engle RF (eds) *Long run economic relations. Readings in cointegration*, Oxford University Press, pp 237 266
- Escribano A, Granger CWJ (1998) Investigating the relationships between gold and silver prices. *J Forecast* (17):81 107
- Escribano A, Peña D (1994) Cointegration and common factors. *J Time Ser Anal* 15:577 586
- Glosten LR (1987) Components of the bid ask spread and the statistical properties of transaction prices. *J Finance* 42:1293 1307
- Glosten LR, Harris LE (1988) Estimating the components of the bid/ask spread. *J Financ Econ* 21:123 142
- Glosten LR, Milgrom PR (1985) Bid, ask and transaction prices in specialist market with heterogeneously informed traders. *J Financ Econ* 14:71 100
- Goldstein MA, Kavajecz KA (2004) Trading strategies during circuit breakers and extreme market movements. *J Financ Mark* 7:301 333
- Green WH (1997) *Econometric analysis*. Prentice Hall, Upper Saddle River, NJ
- Griffiths MD, Smith BF, Alasdair D, Turnbull S, White RW (2000) The costs and determinants of order aggressiveness. *J Financ Econ* 56:65 88
- Harris M, Raviv A (1993) Differences of opinion make a horse race. *Rev Financ Stud* 6:473 506
- Harris FH, McNish TH, Shoesmith GL, Wood RA (1995) Co integration, error correction, and price discovery on internationally linked security markets. *J Financ Quant Anal* 30:563 579
- Hasbrouck J (1988) Trades, quotes and information. *J Financ Econ* 22:229 252
- Hasbrouck J (1991) Measuring the information content of stock trades. *J Finance* 46:179 207
- Hasbrouck J (1995) One security, many markets: determining the contributions to price discovery. *J Finance* 50(4):1175 1199
- Hasbrouck J (1996) Modeling market microstructure time series. In: Maddala GS, Rao CR (eds) *Handbook of statistics, vol 14. Statistical methods in finance*. Elsevier, North Holland, Amsterdam
- Hasbrouck J (1999a) Security bid/ask dynamics with discreteness and clustering: simple strategies for modeling and estimation. *J Financ Mark* 2:1 28
- Hasbrouck J (1999b) The dynamics of discrete bid and ask quotes. *J Finance* 54(6):2109 2142
- Hasbrouck J, Sofianos G, Sosebee D (1993) New York Stock Exchange systems and trading procedures, NYSE Working Paper #93 01
- Hausman JA, Lo AW, MacKinlay AC (1992) An ordered probit analysis of transaction costs prices. *J Financ Econ* 31:319 379
- Holthausen RW, Leftwich RW, Mayers D (1987) The effect of large block transactions on security prices. *J Financ Econ* 19:237 267
- Huang RD, Stoll HR (1994) Market microstructure and stock return predictions. *Rev Financ Stud* 7(1):179 213
- Huang RD, Stoll HR (1996) Dealer versus auction markets: a paired comparison of execution costs on NASDAQ and the NYSE. *J Financ Econ* 41:313 357
- Huang RD, Stoll HR (1997) The components of the bid ask spread: a general approach. *Rev Financ Stud* 10:995 1034
- Jang H, Venkatesh PC (1991) Consistency between predicted and actual bid ask quote revisions. *J Finance* 46:433 446
- Johansen S (1991) Estimation and hypothesis testing of co integration vectors in Gaussian vector autoregressive models. *Econometrica* 59:1551 1580
- Kavajecz KA (1999) A specialist's quoted depth and the limit order book. *J Finance* 54:747 771
- Keim DB, Madhavan A (1995) Anatomy of the trading process: empirical evidence on the behavior of institutional traders. *J Financ Econ* 37:371 398
- Kempf A, Korn O (1999) Market depth and order size. *J Financ Mark* 2:29 48
- Koski JL, Michaely R (2000) Prices, liquidity, and the information content of trades. *Rev Financ Stud* 13:659 696
- Lakonishok J, Lee I (2001) Are insider trades informative? *Rev Financ Stud* 14(1):79 111
- Lee CM, Ready MJ (1991) Inferring trade direction from intraday data. *J Finance* 46:733 746
- Lin JC, Sanger GC, Booth GG (1995) Trade size and components of the bid ask spread. *Rev Financ Stud* 8:1153 1183
- Madhavan A, Sofianos G (1998) An empirical analysis of NYSE specialist trading. *J Financ Econ* 48:189 210

- Madhavan A, Richardson M, Roomans M (1997) Why do security prices change? A transaction level analysis of NYSE stocks. *Rev Financ Stud* 10:1035 1064
- Odders White ER (2000) On the occurrence and consequences of inaccurate trade classification. *J Financ Mark* 3:259 286
- O'Hara M (1995) *Market microstructure theory*. Blackwell, Cambridge
- Park JY, Phillips PCB (2000) Nonstationary binary choice. *Econometrica* 68:1249 1280
- Pascual R, Pascual Fuster B, Climent F (2005) Cross listing, price discovery, and the informativeness of the trading process. *J Financ Mark*, forthcoming
- Roll R (1984) A simple implicit measure of the effective bid ask spread in a efficient market. *J Finance* 39:1127 1139
- Stock J, Watson M (1988) Testing for common trends. *J Am Stat Assoc* 83:1097 1107
- Stoll HR (1989) Inferring the components of the bid ask spread: theory and empirical tests. *J Finance* 19:115 134
- Subrahmanyam A (1997) Multi market trading and the informativeness of stock trades: an empirical intraday analysis. *J Econ Bus* 49:515 531
- Tse Y (2000) Further examination of price discovery on the NYSE and regional exchanges. *J Financ Res* 23:331 351
- White H (1994) *Estimation, inference and specification analysis*. Cambridge University Press, Cambridge
- Wooldridge JM (1994) Estimation and inference for dependent processes. In: Engle RF, McFadden DL (eds) *Handbook of econometrics*, vol IV. Elsevier, North Holland, Amsterdam
- Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J Am Stat Assoc* 57:348 368
- Zhang MY, Russell JR, Tsay RS (2005) Determinants of bid and ask quotes and implications for the cost of trading, Working Paper, University of Chicago