Quote Quality
in an Order Driven Market:
How Much Volatility is Information and How Much is Noise?*

David Veredas
ECARES, Université Libre de Bruxelles
CORE, Université catholique de Louvain
dveredas@ulb.ac.be

Roberto Pascual
Department of Business
Universidad de las Islas Baleares
rpascual@uib.es

Preliminary version:
May 2007
(Please do not quote)

* We are grateful to all the participants in seminars at ECARES, Université Libre de Bruxelles, and University of Luxembourg for their suggestions and comments. David Veredas acknowledges the financial aid of the Belgian National Bank. Roberto Pascual thanks *Sociedad de Bolsas* for providing the database, and acknowledges the financial support of the Spanish DGICYT project SEJ2004-07530-C04-04/ECON. Any errors are entirely our own.

Quote Quality
in an Order Driven Market:
How Much Volatility is Information and How Much is Noise?

ABSTRACT

We propose a new procedure to measure the quality of quotes. Quote quality is defined, first, as the proportion of total variance of quotes due to informational volatility and, second, as the proportion of the correlation between ask and bid quotes explained by the efficient price. These measures are obtained from the reduced form of a structural price formation model and applied to a sample of Spanish common stocks. With quotes sampled at 5-minute intervals, we show that microstructure frictions explain from 5% of the total variance for large caps to 50% for small caps. As quotes are sampled at lower frequencies, quote quality increases. We also report a positive cross-sectional relationship between liquidity and quote quality. Finally, we show that quote quality has a regular U-shaped intraday pattern.

Key words: Quote quality, frictions, microstructure noise, limit order book, price discovery.

JEL: G1

1. Introduction

The existence of trading frictions in financial markets is an essential notion in market microstructure research. Hasbrouck (1996) considers the distinction between permanent (informational) price changes, due to the market's assessment of news, and transitory price changes, attributable to price discreteness, dealer inventory control effects, transaction costs, asymmetric information, temporary order imbalances, the bid-ask bounce, monopoly rents, etc., as one of the two basic dichotomies in microstructure research, the other one being the distinction between trade-related and trade-unrelated information.

Stoll (2000) defines friction in terms of immediacy costs, that is, the price concession needed for an immediate transaction. In this manner, the bid-ask spread becomes the most natural measure of friction. The size of the spread depends on different types of costs carried by liquidity providers: inventory holding costs (Ho and Stoll, 1981, 1983), operative costs (Demsetz, 1968), and monopoly rents (Laux, 1995) constitute what Stoll (op. cit.) calls "real frictions", while adverse selection costs (e.g., Glosten and Milgrom, 1985) form what Stoll calls "informational" frictions. A remarkable research effort has been devoted to the measurement of these theoretical components.[1] Although the empirical findings are not easily reconciled, these studies have shown that both information flows and trading frictions matter in explaining the size of the spread, the intraday price dynamics, and the intraday price volatility.

In a related line of research, it has been shown that microstructure frictions are relevant in asset pricing. Amihud and Mendelson (1986, 1989), Brennan and Subrahmanyam (1996), Brennan, Chordia and Subrahmanyam (1998), and Easley,

---

[1] See, for example, Stoll (1989), Roll (1984), Glosten and Harris (1988), George, Kaul, and Nimalendran (1991), Lin, Sanger and Booth (1995), Huang and Stoll (1996), Madhavan, Richardson and Roomans (1997), and Huang and Stoll (1997).

Hvidkjaer, and O'Hara (2002) have evidenced that different friction measures, either real or informational, considered as proxies for illiquidity, are associated with higher expected returns.

Hasbrouck (2007) identifies trading frictions with the difference between the actual transaction prices and the underlying efficient or "true" price. When prices are sampled over long intervals, such as months, a simple random walk specification may be highly satisfactory in order to describe the price dynamics. Microstructure literature, however, focuses on price formation at very short horizons, such as minutes or seconds rather than months or days. When prices or quotes are sampled at such low frequencies, returns are affected by microstructure noise, inducing transient effects on prices and moving them away from the efficient price.

Particularly relevant for our study, Madhavan et al. (1997) use a structural model of price formation to decompose transaction price volatility into volatility arising from news shocks (trade-unrelated information) and volatility arising from market frictions such as price discreteness, asymmetric information, and real frictions. In the absence of frictions, either real or informational, the model reduces to a random walk process. They find that the fraction of variance attributable to market frictions varies from 54% at the open to 65% at the close for a sample of NYSE-listed stocks. Additionally, real frictions appear more important than informational frictions to explain price volatility. The contribution of real frictions varies from 22% at the open to 35% at the close, while the contribution of informational friction is 13.5% at the open and declines over the day. Though the structural model proposed by Madhavan et al. (op. cit.) is general enough to capture the essential features of trading and price discovery, the search for simplicity and tractability forces the imposition of unrealistic assumptions, such as the homoskedasticity of the innovations to both the efficient price and the transitory component of prices, or the lack of lagged effects on prices.

Hasbrouck (1993) proposes using a vector autoregressive (VAR) model for returns and signed trades to decompose security transaction prices into their efficient (random walk) and error (stationary) components. He also suggests using the standard deviation of the pricing error as a summary measure of market quality. Intuitively, this measure would capture how closely actual transaction prices track the efficient price. The advantage of the Hasbrouck approach resides in its generality; the bivariate VAR model could be seen as the reduced form of a wide range of microstructure models. The disadvantage is that only a lower bound for the true standard deviation of the market friction can be estimated. So as to recover the transitory component from the observed returns and trades, it is necessary to impose identification restrictions *a la* Beverige and Nelson (1981). Hasbrouck reports an average standard deviation of the pricing error of 0.33% of the stock price for a sample of NYSE-listed stocks, but it ranges from 0.55% for small caps to 0.153% for large caps. It is also larger at the beginning and, to a lesser extent, at the end of the trading session.

In this paper, we propose a new approach to measure quote quality that builds on Madhavan et al. (1997) and Hasbrouck (1993). As Hasbrouck, we understand quote quality as the difference between the actual quotes and the underlying efficient (random walk) price, but as Madhavan et al. we focus on decomposing the variance of ask and bid quotes into their information-related and friction-related components. We propose several new measures of quote quality: (a) the variance information share (VIS) is the fraction of the total quote variance which is information-related, and (b) the Quote Quality Measure (QQM) is the proportion of the correlation between ask and bid quotes explained by the efficient price. Intuitively, the closer VIS and QQM are to one, the higher the quality of ask and bid quotes.

This quality measures are computed using the reduced form of a structural co-integration model for ask and bid quotes proposed by Pascual and Veredas (2007). Since we are not aimed to separate real from informational frictions, we consider a model that is general enough to concur with the structural models previously suggested (e.g., Madhavan et al., 2007, Huang and Stoll, 1997) and, at the same time, to allow for more realistic assumptions in a time series analysis, such as heteroskedasticity and lagged effects in the price formation process. Moreover, instead of estimating the reduced form and, then, imposing identification restrictions, as Hasbrouck (1993), we proceed just on the opposite order. We estimate first the structural model using the Kalman Filter, and, afterwards, we obtain the reduced form and compute the quality measures.[2]

In this preliminary version of our paper, we apply our methodology to a sample of stocks traded in the electronic continuous platform of the Spanish Stock Exchange (SSE) in 2000. As far as we know, this is the first attempt to measure market quality in a “pure” order driven market. Using quotes sampled at 5-minute intervals, we show that microstructure frictions explain a significant fraction of total variance, ranging from 5% for large caps to 50% for small caps. As quotes are sampled as lower frequencies, however, quote quality increases. We also report a positive cross-sectional relationship between liquidity and quote quality. Finally, we show that quote quality has a regular

[2] This article is also related to the literature on daily integrated variance estimators. Estimators such as the realized volatility and the bi-power variation are biased and inconsistent when high frequency observations are contaminated by market frictions, and their bias increases unboundedly with the intraday frequency of sampling. This literature has focused on the search of techniques that provide estimators with better properties. Nonetheless, as far we know, no study has estimated the intraday variance due to the microstructure noise. Our methodology is a vehicle to perform such estimation.

U-shaped intraday pattern, that is, quote quality achieves minimum levels at the central intervals of the trading session and maximum levels at the begging and end of trading.

The remainder of this paper proceeds as follows. In section 2, we present the structural model for ask and bid quotes and derive its reduced form. We also introduce our measures of quote quality. In section 3, we provide a brief description of the most remarkable features of the continuous and electronic platform of the SSE; we present our database, and we provide some descriptive statistics on our sample of ordinary stocks. In section 4, we provide our preliminary findings. Finally, we conclude in section 5 and discuss our course of action to improve the contents of this preliminary version of this article.

2. Econometric model and quote quality measures

We measure quote quality by means of the reduced form of a structural co-integration model for ask and bid quotes introduced by Pascual and Veredas (2007). For expositional purposes, we provide next a brief discussion of this econometric model and we redirect the reader to the above mentioned paper for further details and comments.[3]

The model builds on the traditional microstructure price decomposition into an efficient or long run component and a transitory or noisy component (e.g. Hasbrouck, 2007). The latter reflects market-specific frictions attributable to the market organization and trading activity, such as inventory adjustments, adverse selection costs, temporary order imbalances, the bid-ask bounce, minimum price variations, etc. (see Stoll, 2000).

[3] The model we describe next is a simplified version of the model proposed by Pascual and Veredas (2007), which allows for correlation across innovations to the unobserved components of quotes and for dynamics in the volatility of the transitory components. Preliminary trials reveal that the reduced form of this more complex specification would be far more complex and cumbersome. To keep things tractable in this our preliminary approach to the problem, we will restrict ourselves to the most parsimonious model and we will leave the treatment of the most complex specification for posteriors versions.

Shocks to the long run component cause permanent price impacts because they are connected to the revelation of brand new, non-publicly known, information. Hence, a shock to the efficient price implies a revision in the conditional expectation about the terminal value of the stock. Since revisions in expectations are assumed to be unpredictable, the efficient price possesses a martingale property. This motivates empirical and theoretical representations where the efficient price follows a random walk process (see Hasbrouck, 2002). Therefore, the observed prices or quotes are assumed to consist on a random walk plus transitory disturbances. In our particular case, ask and bid quotes are decomposed into their theoretical unobservable components as follows,

$$\begin{aligned} a_t &= m_t + S_{at} + \beta \\ b_t &= m_t + S_{bt} - \beta \end{aligned}, \qquad [1]$$

where a_t and b_t stand for the best ask quote and the best bid quote, respectively. Since they share the same long run component, ask and bid quotes are co-integrated, with co-integration vector $(1,-1)$.[4] The $\beta \geq 0$ coefficient accounts for half the average bid-ask spread. There efficient price (m_t) is assumed to follow a random walk process,

$$m_t = m_{t-1} + \varepsilon_{m,t} \,. \qquad [2]$$

The observed quoted prices deviate from the efficient price in [2] by the stochastic transitory components S_{at} and S_{bt}, also unobserved. The dynamics of the transitory components model are given by the following bivariate first order autoregressive process,

$$\begin{aligned} S_{a,t} &= \phi_a S_{a,t-1} + \phi_{ab} S_{b,t-1} + \varepsilon_{a,t} \\ S_{b,t} &= \phi_{ba} S_{b,t-1} + \phi_b S_{b,t-1} + \varepsilon_{b,t} \end{aligned} \qquad [3]$$

[4] See Hasbrouck (1995), Engle and Patton (2004), and Escribano and Pascual (2006) for co-integration models for ask and bid quotes.

which allows for lagged causality between the transitory components.[5]

Finally, the innovations in [2-3] are assumed to be jointly normally distributed but uncorrelated,

$$\begin{pmatrix} \varepsilon_{a,t} \\ \varepsilon_{b,t} \\ \varepsilon_{m,t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{a,t}^2 & 0 & 0 \\ 0 & \sigma_{b,t}^2 & 0 \\ 0 & 0 & \sigma_{m,t}^2 \end{pmatrix} \right). \qquad [4]$$

The normality assumption, though censurable, permits keeping the model into a territory where the unobserved components can be estimated using the Kalman filter (see Harvey, 1992). Moreover, using Pseudo-Maximum Likelihood (PML), consistency is guaranteed (Gouriéroux, Monfort, and Trognon, 1984). Likewise, a diagonal variance-covariance matrix is a very strong assumption. Since microstructure theory has proven the existence of an adverse selection component in the bid-ask spread, the transitory disturbances could be correlated with the innovations to the efficient price.[6] A similar conclusion could be drawn from the literature on common factor models of price discovery (see Lehmann, 2002) and realized volatility (e.g., Bandi and Russell, 2005). Once more, these assumptions are imposed to maintain our model as tractable as possible and will be relaxed in more advanced versions of this paper.

[5] Pascual and Veredas (2007) allow for deterministic intraday patterns in the transitory components, recognizing the intraday regularities in the bid-ask spread widely reported in the literature (e.g., McInish and Wood, 1992). In this paper, we will take these regularities into account by analyzing quote quality at different time intervals of the trading session, but we do not explicitly model them in order to simplify the derivation of the reduced form of the structural model.

[6] Copeland and Galai (1983), Glosten and Milgrom (1985), and Easley and O'Hara (1987) for price-driven markets, and Glosten (1994), Foucault (1999), and Handa et al. (2003) for order-driven markets, discuss the role of adverse selection in price discovery. Glosten and Harris (1988), Stoll (1989), George, Kaul, and Nimalendran (1991), Easley, Kiefer, and O'Hara (1997), and Madhavan, Richardson, and Roomans (1997), among others, propose models for decomposing the bid-ask spread into its theoretical components, including the one due to adverse selection.

The reduced form of model [1-4] is a VMA(∞) model (see Appendix for derivation details),[7]

$$
\begin{aligned}
\Delta a_t &= \frac{\Delta\varepsilon_{at}}{\Phi^a(L)} + \varepsilon_{mt} + \phi_{ab}\frac{\Delta\varepsilon_{bt-1}}{(1-\phi_b L)\Phi^a(L)} \\
\Delta b_t &= \frac{\Delta\varepsilon_{bt}}{\Phi^b(L)} + \varepsilon_{mt} + \phi_{ba}\frac{\Delta\varepsilon_{at-1}}{(1-\phi_a L)\Phi^b(L)}
\end{aligned} \qquad [5]
$$

where $\Phi^a(L)$ and $\Phi^b(L)$ are infinite order stationary polynomials on the lag operator L ($L^k y_t = y_{t-k}$),

$$
\begin{aligned}
\Phi^a(L) &= 1-\phi_a L - \phi_{ab}\phi_{ba}(L^2 + \phi_b L^3 + \phi_b^2 L^4 + \phi_b^3 L^5 + \ldots) \\
\Phi^b(L) &= 1-\phi_b L - \phi_{ba}\phi_{ab}(L^2 + \phi_a L^3 + \phi_a^2 L^4 + \phi_a^3 L^5 + \ldots)
\end{aligned} \qquad [6]
$$

The stationarity of the autoregressive polynomials in [6] guarantees that their coefficients decay rapidly to zero. For example, for the ask polynomial, the j^{th} coefficient $\forall j>1$ is $\phi_{ab}\phi_{ba}\phi_b^{\,j-2}$. We will see in posterior sections that, in practical terms, we can cut off the polynomial at very short lags, say 3; adding further lags will not change our main findings.

Under the assumption of independence across innovations, model [5] implies that the conditional (unconditional) variances of the returns of ask and bid quotes are linear combinations of the conditional (unconditional) variances of the error terms,

$$
\begin{aligned}
Var(\Delta a_t) &= \sigma^2_{a,t}\omega^a_a + \sigma^2_{m,t} + \sigma^2_{b,t}\omega^a_b \\
Var(\Delta b_t) &= \sigma^2_{a,t}\omega^b_a + \sigma^2_{m,t} + \sigma^2_{b,t}\omega^b_b
\end{aligned} \qquad [7]
$$

where ω^j_i, for i and j = {a,b}, are a function of the autoregressive parameters.

[7] The infinite order of the VMA representation is due to the coefficients ϕ_{ab} and ϕ_{ba}. It they were zero, the model would be reduced to a VARMA(1,1).

We use an exponential form for the variances of the transitory components of quotes, where conditional variances are a function of the time of the day,

$$\sigma_{i,t}^2 = \exp\left(\sigma_{i,k}^2 + \sum_{\substack{j=1 \\ j\neq k}}^{J} \phi_j D_{t,j} \right),\ i = \{a,b\} \qquad [8]$$

where D_{tj}, for j ={1,…,J}, are dummy variables that take value 1 if the t[th] observation lies in a given day interval and zero otherwise, and $\sigma_{i,k}^2$ is the volatility during the reference interval k (13:00-15:00).[8] The exponential form guarantees the non-negativeness of the conditional variance, since some ϕ_j could be negative. Since observations are regularly spaced, the time of the day is deterministic, and by the law of large numbers, the unconditional variances, $\sigma_i^2 = E\left[\sigma_{i,t}^2\right]$ for i={a,b}, could be computed as,

$$\sigma_i^2 = \frac{1}{T}\sum_{t=1}^{T}\sigma_{i,t}^2 = \frac{1}{T}\exp\left(\sigma_{i,k}^2\right)\left[\sum_{\substack{j=1 \\ j\neq k}}^{J} N_j \exp\left(\phi_j^i\right) + N_k\right], \qquad [9]$$

where N_j , j = {1,…,J}, is the number of observations per intraday bin times the number of days, and k is the control interval.

As for the variance of the efficient price ($\sigma_{m,t}^2$), Pascual and Veredas (2007) follow Hasbrouck (1999) in using an EGARCH type of model which allows for deterministic and dynamic components,

$$\sigma_{m,t}^2 = \exp\left(\alpha_0 + \alpha_1 \ln\left(\sigma_{m,t-1}^2\right) + \alpha_2 \xi_{m,t-1} + \alpha_3\left[\left|\xi_{m,t-1}\right| - \sqrt{2/\pi}\right] + \sum_{\substack{j=1 \\ j\neq k}}^{J} \phi_j^m D_{t,j} \right) \qquad [10]$$

where $\xi_{m,t} = \varepsilon_{m,t} / \sigma_{m,t}$ is the standardized shock, α_0 is the intercept, α_1 is the persistence, α_2 is the leverage, and α_3 is the magnitude parameter. To compute the

[8] Pascual and Veredas (2007) also allow for dynamics in the volatility of the short run component. Preliminary estimations of more complex specifications suggest that by adding dynamics in the specification of the short run volatility quote quality decreases, meaning that the results presented could be understood as lower bounds to the contribution of microstructure noise to price adjustments.

unconditional variance of [10], we apply the law of iterated expectations and the law of large numbers to get that,

$$\sigma_m^2 = \frac{1}{T}\sum_{t=1}^{T}\sigma_{m,t}^2 \ . \qquad [11]$$

Variances of Δa_t and Δb_t can now be computed from [4] (see the Appendix for details),

$$Var(\Delta a_t) = \sigma_a^2\left(1+\sum_{j=1}^{\infty}\left(\theta_{j+1}^a - \theta_j^a\right)^2\right) + \sigma_m^2 + \phi_{ab}^2\sigma_b^2\left(1+\sum_{j=1}^{\infty}\left(\gamma_{j+1}^a - \gamma_j^a\right)^2\right), \qquad [12]$$

$$Var(\Delta b_t) = \sigma_b^2\left(1+\sum_{j=1}^{\infty}\left(\theta_{j+1}^b - \theta_j^b\right)^2\right) + \sigma_m^2 + \phi_{ba}^2\sigma_a^2\left(1+\sum_{j=1}^{\infty}\left(\gamma_{j+1}^b - \gamma_j^b\right)^2\right), \qquad [13]$$

with θ_j^i, for i = {a,b}, γ_j^a, and γ_j^b being the j^{th} element of, $\theta^i(L) = 1/\Phi^i(L)$, $\gamma^a(L) = \phi_{ab}/(1-\phi_b L)\Phi^a(L)$, and $\gamma^b(L) = \phi_{ba}/(1-\phi_a L)\Phi^b(L)$, respectively.

Similarly, the covariance between Δa_t and Δb_t is a weighted sum of the variances of the short run and long run unobserved components of quotes,

$$\begin{aligned} Cov(\Delta a_t, \Delta b_t) = {} & \phi_{ba}\sigma_a^2\left[\left(\theta_1^a - 1\right) + \sum_{j=2}^{\infty}\left(\theta_j^a - \theta_{j-1}^a\right)\left(\gamma_j^b - \gamma_{j-1}^b\right)\right] + \\ & +\phi_{ab}\sigma_b^2\left[\left(\theta_1^b - 1\right) + \sum_{j=2}^{\infty}\left(\theta_j^b - \theta_{j-1}^b\right)\left(\gamma_j^b - \gamma_{j-1}^b\right)\right] + \sigma_m^2 \end{aligned} . \qquad [14]$$

In order to measure quote quality, we propose different kinds of measures. The first ones, we call Variance Information Shares (VIS) are defined as the percentage of the total volatility of either quote that is due to information,

$$VIS^a = \frac{\sigma_m^2}{Var(\Delta a_t)}, \qquad [15]$$

$$VIS^b = \frac{\sigma_m^2}{Var(\Delta b_t)}. \quad [16]$$

The higher this measures, the higher the quote quality. The second measure that we call Quote Quality Measure (QQM) is based on the correlation between the returns of ask and bid quotes. In particular, QQM measures how much of the correlation between ask quote and bid quote's returns is due to the long run component. Using [14], the correlation is given by,

$$Corr(\Delta a_t, \Delta b_t) = \frac{\phi_{ba}\omega_a\sigma_a^2 + \phi_{ab}\omega_b\sigma_b^2 + \sigma_m^2}{\sqrt{Var(\Delta a_t)}\sqrt{Var(\Delta b_t)}}, \quad [17]$$

and the proportion due to the informational variance is,

$$QQM = \frac{\sigma_m^2}{\sqrt{Var(\Delta a_t)}\sqrt{Var(\Delta b_t)}}. \quad [18]$$

The closer QQM is to 1, the higher the quality of ask and bid quotes, because the closer their returns to the efficient returns.

3. Market background and data

Our database comprises high frequency limit order book and transaction data for 69 stocks listed in the Spanish Stock Exchange (SSE) from July to December 2000.[9] In 2005, the "Annual Report and Statistics" of the World Federation of Exchanges classifies the SSE as the 9th world largest stock exchange in terms of domestic market capitalization (US$960 billions), and the 7th in terms of total value of share trading (US$1566 billions). It is also the 4th stock exchange in Europe, right after the London Stock Exchange, Euronext, and the Deutsche Börse.

[9] We choose this particular semester to avoid modelling the incidence of an intraday circuit breaker mechanism, which combines stock specific price limits and rule-based short-lived volatility auctions, introduced in March 2001 (see Abad and Pascual, 2006, for details).

In the SSE, almost all the trading activity is handled by an electronic order-driven platform called SIBE (*Sistema de Interconexión Bursátil Español*). In the second semester of 2000, there were around 101 stocks negotiated in the continuous market that were never transferred to alternative trading mechanisms.[10] Liquidity provision in the SIBE exclusively depends on an open limit order book (LOB) since there are not designated market makers. During the trading process, market members have access to information about both the LOB and the trading activity through their vendor screens. LOB information includes up to the 20 best ask and bid quote levels, together with the displayed depth and number of orders at each level. Transaction information includes price, volume and market members' identities.[11] All this information is updated in real time every time a new order is submitted, modified or cancelled, or a new transaction is completed.

Continuous trading takes place between 9:00 a.m. to 5:30 p.m., preceded by a 30-minute opening auction and followed by a 5-minute closing auction. During the continuous session investor could submit limit orders, market orders, and market-to-limit orders.[12] Special conditions are allowed, such as hidden volume (iceberg orders), "fill or kill", and "minimum volume". Orders could also be modified and cancelled anytime. They are stored in the LOB following the usual price-time priority rule. Hidden volume losses time priority. A transaction occurs when a market, a market-to-limit, or a marketable limit order hits the opposite side of the book. Therefore, price

[10] Infrequently traded and highly illiquid stocks are negotiated in a parallel trading mechanism called Fixing, consisting on two consecutive and long-lasting double auctions. As trading and liquidity conditions improve, stocks could be transferred from this discrete trading venue to the continuous trading platform. Similarly, stocks in the continuous platform with declining trading and liquidity conditions are reassigned to the Fixing.
[11] In 2000, only the five best ask and bid quotes were available.
[12] Market orders in the SSE walk up or down the book till they are fully executed. Market-to-limit orders are restricted to the best price on the opposite side of the market. Limit orders which are not executed instantaneously get stored on the LOB until they find counterparty, they are cancelled, or they expire (by default, at the end of the day).

improvements, so common in US markets such as the NYSE, are not possible in the SSE.[13]

We consider in this study those stocks among all the stocks traded in the continuous market during our sample period with a median of daily transactions greater than 20. This filter eliminates 31 stocks. Among the 70 stocks remaining, we have the 35 most frequently traded and liquid, which are constituents of the official SSE index, the IBEX-35. One of these stocks, however, is excluded because of a merger. Table I provides average daily statistics about the remaining 69 stocks, the 34 index constituents (henceforth IDX stocks), and the 35 stocks not included in the IBEX-35 (henceforth NIDX stocks). We also provide separate statistics for the 5 largest IDX stocks to make obvious the uniqueness of these stocks, cross-listed in the US markets as ADRs, in the global mark of the SSE market. Trading activity is almost 9 times larger for IDX stocks than from NIDX stocks in terms of both number of trades and volume. Liquidity is also higher among IDX stocks, with smaller relative spreads (0.0035 versus 0.0088) and larger displayed book depth (8.182 shares versus 6.557 shares).[14] Limit order submissions are also 6.5 times larger for IDX than for NIDX stocks. Notice, however, that previous comparisons become more dramatic when the five largest stocks are compared with NIDX or even with the other IDX stocks.

[Table I]

Our database includes both book and transaction data. LOB files consist on all the updates of the five best ask and bid levels time stamped to the nearest hundredth of a second. The displayed depth and the number of limit orders accumulated at each level are also available. Transaction files provide information about the accumulated traded

---

[13] For further details on hidden volume in the SSE, see Pardo and Pascual (2006). For more detailed information on the microstructure of the SSE, please visit www.bolsasymercados.es/ing/home.htm.

[14] The average price (not reported) of NIDX stocks is also far smaller than for NIDX stocks, which explains why differences in displayed depth are not as striking as for other statistics.

volume (in shares) at each update of the first level of the LOB. By matching both files, we can identify all trades, classify them as buyer or seller initiated, and determine their size.[15]

4. Preliminary findings

In this section, we provide preliminary findings about quote quality for our sample of SSE stocks. These findings are preliminary because, so far, we have computed the reduced form of our structural model only for the subset of IDX stocks, and because these findings are based on the oversimplified structural model presented in section 2. In the next section, we will outline our research plan for improving the analysis reported in this version of the paper.

Table II provides a summary of the estimation, by the Kalman Filter, of the structural model in section 2 for the 69 stocks in our sample. Time series of quotes are constructed in 5-minute intervals.[16] Our findings do not remarkably differ from those reported by Pascual and Veredas (2007) using a more reduced set of stocks. We obtain that the median estimated spread (2β) is around 4.5 ticks (0.046€), and it is 3 times narrower for the 5 largest stocks (1.6 ticks). The cross-sectional correlation between $2\hat{\beta}$ and the true average bid-ask spread for the 69 stocks is 0.99. We report a high degree of persistence in the transitory components; the cross-sectional median of $\hat{\phi}_a$ ($\hat{\phi}_b$) is 0.848 (0.617). Persistency is also lower among the 5 largest stocks than for the rest of the sample. We also find $S_{a,t}$ and $S_{b,t}$ to cause each other, with $\hat{\phi}_{ab}$ ($\hat{\phi}_{ba}$) being positive. Hence, transitory components of ask and bid quotes tend to move together, though not

[15] Namely, we apply an algorithm originally developed by Pardo and Pascual (2006) to differentiate between trades, limit order submissions, cancellations, modifications, etc.

[16] We have also estimated the model in 1-minute and 10-minute periodicity. Results are available upon request from the authors.

necessarily in a symmetric manner; $\hat{\phi}_{ab}$ ($\hat{\phi}_{ba}$) are higher among the 5 largest stocks, suggesting more symmetry in quote adjustments for these stocks. In general, $\hat{\phi}_{ab}$ is close to $\hat{\phi}_{ba}$ except for the NIDX subsample, supporting Escribano and Pascual (2006) evidence that asymmetric dynamics are more common among infrequently traded stocks.

The cross-sectional median of transitory volatility during the control interval 13:00-15:00 ($\sigma^2_{a,k}, \sigma^2_{b,k}$) is lower for the 5 largest stocks, and it is not statistically different across the others stocks in the sample. This latter should not come as a surprise, since among the other 29 IDX stocks we find the most important high-tech and R&D firms traded in the SSE, whose stocks are among the most volatile. As for the long run volatility, the persistence parameter (α_1) is greater for the 5 largest stocks (α_1 = 0.975) than for the other stocks in the sample. If we assume that information comes in clusters and, therefore, informational volatility also progresses in clusters of high and low levels, our findings point toward information flow intensity being more stable among the largest stocks in the sample. Regarding the magnitude parameter (α_3), it is smaller among the 5 largest stocks (α_3 = 0.157) than for the other stocks, meaning that these stocks are less responsive to an informative shock of any given size. Finally, the leverage parameter (α_2) is smaller in magnitude than the other parameters in the EGARCH model, and it is the only parameter in the whole model that, for several stocks, it is not significant.

Figure 1 reports the cross-sectional median of the intraday regular patterns in transitory volatility given by the estimated parameters ϕ_j in [8]. Notice that, in the particular version of Pascual and Veredas' (2007) model we consider, transitory volatility is a purely deterministic process.[17] Transitory volatility shows a clear inversed J-pattern, with highest levels at the begging of the trading session and with a significant increase towards the end of the day, starting at the opening of the US markets. Notice also that this regular pattern is less prominent among the 5 largest IDX stocks.

[Figure 1]

Figure 2 displays the cross-sectional median of the intraday regular patterns in informational volatility given by the estimated parameters ϕ_j^m in [10]. Regular patterns in this case are very similar across subsamples. Informational volatility is very close to its level during the 13:00-15:00 interval almost all the session, except towards the end of trading. In moves up right after the opening of the US markets and achieves its regular maximum during the last half-hour of the session. A comparison of the magnitudes of the coefficients in Figure 1 and Figure 2 also reveals that the regular patterns found in transitory volatility are more extreme than in informational volatility.

[Figure 2]

We now proceed with the computation of the quote quality measures introduced in section 2. At the moment, only the findings for the subsample of IDX stocks are available. Figure 3 reports the *VIS* measure in [15]-[16] for these stocks, computed over 5-minute intervals. Stocks are ordered by market capitalization. Remember that this measure captures how much of the total volatility of quotes is attributable to their common heteroskedastic random walk component. The patterns obtained using ask quotes (dark bars) are very close to those obtained using bid quotes (light bars). Even though the pattern is not smooth, Figure 2 shows a positive relationship between quote

[17] We split the SSE trading session in the following time intervals [9:00, 9:30), [9:30, 10:00), [10:00, 13:00), [13:00, 15:00), [15:00, 15:30), [15:30, 16:00), [16:00, 16:30), [16:30, 17:00), [17:00, 17:30]. With this splitting, we pretend to capture opening effects, the beginning of trading in the NYSE at 15:30 Spanish Time, the announcement on macroeconomics news (usually around 15:00), and any closing effect.

quality and market capitalization. The 5 largest IDX stocks (TEF, SCH, BBVA, REP, and ELE) have a median VIS^a (VIS^b) of 95% (92%), while the other IDX stocks have a median VIS^a (VIS^b) of 66% (70%). Two stocks with high *VIS* values, TRR and TPZ, experienced heavy declines in their market value during the trading period considered. In these cases, VIS indicates that these strong downward adjustments in prices merely reflected deep revisions in the fundamentals of these stocks. In the best case, TEF, microstructure noise explains less than 2% of the total variance; in the worst case, ALB, microstucture noise explains around 50% of the total variance of quotes.

[Figure 3]

Figure 4 provides the estimated QQM in [18] for the same subset of stocks, also ordered by market capitalization. QQM is greater among the 5 largest stocks (0.93 in median) than among the other stocks (0.68 in median). Indeed, Figures 3 and 2 are almost identical, suggesting that our two quote quality measures are close to each other. Indeed, there are several possible scenarios in which both measures will coincide. First, if volatility is only determined by the random walk innovations, $\sigma_a^2 = \sigma_b^2 = 0$, frictions will degenerate to zero; then, $VIS^a = VIS^b = QQM = 1$. Second, if ask and bid quotes move symmetrically so that both quotes are updated in the same direction and by the same amount after any shock, which would happen if $\varpi_a^a = \varpi_b^a$ and $\varpi_a^b = \varpi_b^b$ or, equivalently $\phi_a = \phi_b$ and $\phi_{ab} = \phi_{ba}$, then $VIS^a = VIS^b = QQM < 1$. Finally, when the dynamics of ask and bid quotes are symmetric in both mean and variance, that is, $\sigma_a^2 = \sigma_b^2$, $\varpi_a^a = \varpi_b^a$, and $\varpi_a^b = \varpi_b^b$, again $VIS^a = VIS^b = QQM < 1$.

[Figure 4]

Therefore, $\left|VIS^a - VIS^b\right|$ proxies for asymmetries in the dynamics of ask and bid quotes, which are due to differences in the process of ask and bid frictions, either in their mean or in their variance. The closer $\left|VIS^a - VIS^b\right|$ is to zero, the more symmetric the dynamics of ask and bid quotes would be, and the closer QQM is going to be to the VIS measure. Indeed, from [18] it is easy to see that if $\left|VIS^a - VIS^b\right| \neq 0$, $\min(VIS^a, VIS^b) < QQM < \max(VIS^a, VIS^b)$.

Figure 5 compares the QQM estimated using quotes sampled over 1-minute, 5-minute, and 10-minute time intervals. This figure shows that as we sample prices at progressively longer intervals, microstructure noise effect decreases, causing QQM to increase. We would expect differences to be more striking at lower frequencies, such as 30-minute intervals.

[Figure 5]

Next, we study to what extent stock features such as capitalization, trading activity, and liquidity help to explain the cross-sectional differences in our quality measures. So as to do that, we compute the following variables, averaged across days, for each stock: (a) the relative spread (RS) is the average bid-ask spread divided by the quote midpoint, both weighted by time; (b) the limit order book depth (D) is the accumulated euro value of the volume offered at the five best ask and bid levels of the limit order book, also weighted by time; (c) volume (V) is the number of shares transacted; trades (T) is the number of trades completed, and order flow (OF) is the number of limit orders, market orders, and cancellations submitted. Market capitalization (C) is measured at 12-31-2000.

Since these variables are highly cross-correlated, we perform a principal components analysis in order to extract the common factors. We two first principal components are,

$$PC_i^1 = (0.4246)C_i + (\text{-}0.3379)RS_i + (0.4197)D_i + (0.4241)V_i + \\ + (0.4239)T_i + (0.4121)OF_i$$

$$PC_i^2 = (0.1969)C_i + (0.9151)RS_i + (0.0589)D_i + (0.3313)V_i + \\ + (0.0876)T_i + (0.0563)OF_i$$

The first component has an eigenvalue of 5.056, explains 84.3% of the variability, and it is an average of all the variables considered, with the spread in negative sign and the others in positive sign. This factor could therefore be considered as a proxy for liquidity. The second component has an eigenvalue of 0.5, explains a marginal 8.34% of the variability, and it is basically determined by the relative spread.

Table III summarizes a regression of each of our quote quality measures in the principal components PC_i^1 and PC_i^2 above. We consider again three different frequencies to sample ask and bid quotes: 1, 5, and 10 minutes. The model has been estimated by Least Absolute Deviation (LAD) with iteratively re-weighted least squares. Standard errors are computed with 200 bootstrap draws. We also report the robust adjusted-R^2 of MacKean and Sievers (1987).

[Table III]

Table III shows that PC_i^1 is statistically significant at least at the 5% level in the 1-minute and 10-minute frequencies and at the 10% level for the 5-minute frequency. PC_i^2 is not statistically significant in any case. All the coefficients of PC_i^1 are positive, meaning that liquidity discriminates between stocks with higher and lower quote quality. More liquid stocks, measured in terms of higher activity, narrower spreads, higher depth, or higher market value, also have quotes with higher quality.

Finally, using the conditional variances in [7], it is possible to obtain quote quality measurements at different points in time during the trading session. In order to do that, we define the Conditional Quote Quality Measure (CQQM) as,

$$CQQM_t = \frac{\sigma^2_{m,t}}{\sqrt{Var(\Delta a_t)_t}\sqrt{Var(\Delta b_t)_t}} .$$

We compute the average across days of $CQQM_t$ at the different 5-minute intervals of the SSE trading session to study the regular patterns in quote quality. Figure 6 reports our findings for three IDX stocks in our sample, TEF, TRR, and ANA. The three stocks show a similar U-shaped pattern, with quote quality decreasing achieving its regular minimum at the intermediate intervals of the trading session. Information-related volatility contributes more to total volatility at the beginning and towards the end of trading.

5. Summary, conclusions, and future research agenda

We have proposed a methodology to measure the quality of ask and bid quotes. Quote quality is understood as how closely the actual quotes track the underlying efficient or true price of the stock. We assume this efficient price follows a random walk process, but with heteroskedatic innovations. We propose two different measures of quote quality: the Variance Information Share (VIS) is the fraction of total volatility that is information-related; the Quote Quality Measure (QQM) is the fraction of the correlation between ask and bid quotes attributable to the common efficient price. The closer both measures are to one, the higher the quality of quotes. We show that under the assumption of symmetric dynamics of ask and bid quotes, both measures are equivalent.

The abovementioned measures are estimated using the reduced form of a structural model of quote formation. In particular, in this preliminary version of our work, we use a simplified version of a structural co-integration model for ask and bid quotes proposed by Pascual and Veredas (2007). The model is estimated using the Kalman Filter. We show that its reduced form is a Vector Moving Average (VMA) model of infinite order.

We apply our methodology to the subset of 34 stocks traded in the SSE, an electronic order driven market, based on an open limit order book. We find that microstructure frictions explain a significant fraction of total variance, but ranging from 5% for the large caps in the sample to around 50% for the small caps. We also report a positive cross-sectional relationship between liquidity and quote quality using quotes sampled at different frequencies. However, as quotes are sampled as progressively lower frequencies, microstructure friction's effect in the returns decreases, and quote quality increases. Finally, we have also evidenced that quote quality has a regular U-shaped intraday pattern, achieving minimum levels at the central intervals of the trading session.

All the findings reported in this paper must be considered as preliminary. In our future research agenda, we are aimed to improve the analysis performed so far following the next guidelines:

(a) *Sample*: We expect to strengthen the cross-sectional evidence reported by applying our methodology to the 35 NIDX stocks in our sample.

(b) *Structural model*: As indicated in the paper, previous findings are based on a restricted version of the structural model suggested by Pascual and Veredas

(2007). In particular, we have assumed that transitory volatility is a deterministic process, while Pascual and Veredas (op. cit.) allow for dynamics in addition to regular patterns. We expect the fraction of quote volatility attributed to market frictions to increase as we relax our restriction. In addition, to facilitate the derivation of the reduced form model we have also avoided to model the deterministic intraday patterns in the bid-ask spread and, probably more important, we have imposed independence across innovations. We plan to relax these assumptions progressively and study their incidence in our measurement of quote quality.

(c) *Market quality*: So far we have applied our methodology to an order driven market, the Spanish Stock Exchange. Our methodology, however, could be useful to study the incidence of market regulation in the quality of quotes. Thus, we plan to construct matched samples of stocks traded in different markets, with differing market microstructures, such as NYSE and Nasdaq, and perform a comparative analysis of quote quality across markets.

(d) *Other methodologies*: Other methodologies to measure quote quality have been proposed in the literature. In particular, our methodology builds on Madhavan, Richardson, and Roomans (1997) and Hasbrouck (1993). In posterior version of this paper, we are intended to compare to what extent quote quality measurements differ across methodologies and study the sources of possible discrepancies.

Appendix

(A) Derivation of the reduced form model

First, we substitute the transitory components in [3] into [1],

$$\begin{aligned} a_t &= \phi_a S_{at-1} + \phi_{ab} S_{bt-1} + \varepsilon_{at} + m_t + \beta \\ b_t &= \phi_{ba} S_{at-1} + \phi_b S_{bt-1} + \varepsilon_{bt} + m_t - \beta \end{aligned}.$$

By [2],

$$\begin{aligned} a_t &= \phi_a (a_{t-1} - m_{t-1} - \beta) + \phi_{ab} (b_{t-1} - m_{t-1} + \beta) + \varepsilon_{at} + m_t + \beta \\ b_t &= \phi_{ba} (a_{t-1} - m_{t-1} - \beta) + \phi_b (b_{t-1} - m_{t-1} + \beta) + \varepsilon_{bt} + m_t - \beta \end{aligned}. \quad \text{[A1]}$$

By lagging one period the second equation,

$$(b_{t-1} - m_{t-1} + \beta) = \frac{\phi_{ba}}{1-\phi_b L}(a_{t-2} - m_{t-2} - \beta) + \frac{1}{1-\phi_b L}\varepsilon_{bt-1}. \quad \text{[A2]}$$

Then, substitute [A2] in the ask equation in [A1],

$$(a_t - m_t - \beta)\left(1 - \phi_a L - \phi_{ab}\phi_{ba}\frac{L^2}{1-\phi_b L}\right) = \frac{\phi_{ab}}{1-\phi_b L}\varepsilon_{bt-1} + \varepsilon_{at}. \quad \text{[A3]}$$

We denote by $\Phi^a(L)$ the term in parenthesis in the left-hand side of [A3]. Diving both sides by $\Phi^a(L)$,

$$a_t = \frac{\phi_{ab}}{(1-\phi_b L)\Phi^a(L)}\varepsilon_{bt-1} + \frac{1}{\Phi^a(L)}\varepsilon_{at} + m_t + \beta. \quad \text{[A4]}$$

Using [2] and pre-multiplying both sides by $\Delta = (1-L)$,

$$\Delta a_t = \frac{\phi_{ab}}{(1-\phi_b L)\Phi^a(L)}\Delta\varepsilon_{bt-1} + \frac{1}{\Phi^a(L)}\Delta\varepsilon_{at} + \varepsilon_{mt}. \quad \text{[A5]}$$

Finally, follow the same steps to obtain the reduced form for the bid equation.

(B) Variance and covariance: $\theta^a(L)$, $\theta^b(L)$, $\gamma^a(L)$ and $\gamma^b(L)$

As previously defined,

$$\begin{aligned}\theta^a(L) &= \frac{1}{\Phi^a(L)},\ \gamma^a(L) = \frac{\phi_{ab}}{(1-\phi_b L)\Phi^a(L)} \\ \theta^b(L) &= \frac{1}{\Phi^b(L)},\ \gamma^b(L) = \frac{\phi_{ba}}{(1-\phi_a L)\Phi^b(L)}\end{aligned}.$$

From $\theta^a(L)$,

$$1 = (1-\phi_a L - \phi_{ab}\phi_{ba}L^2 - \phi_{ab}\phi_{ba}\phi_b L^3)(1+\theta_1^a L + \theta_2^a L^2 + \theta_3^a L^3 + \ldots).$$

Denote $A = \phi_{ab}\phi_{ba}$. After some simple algebra,

$$\begin{aligned}
\theta_1^a &= \phi_a \\
\theta_2^a &= \phi_a^2 + A \\
\theta_3^a &= \phi_a^3 + A(2\phi_a + \phi_b) \\
\theta_4^a &= \phi_a^4 + A(3\phi_a^2 + 2\phi_b\phi_a + A) \\
\theta_5^a &= \phi_a^5 + A(4\phi_a^3 + 3\phi_b\phi_a^2 + 2A\phi_b + 3A\phi_a) \\
\theta_6^a &= \phi_a^6 + A(5\phi_a^4 + 4\phi_b\phi_a^3 + 6A\phi_b\phi_a + 6A\phi_a^2 + A^2 + A\phi_b^2) \\
&\ldots
\end{aligned}.$$

Therefore,

$$\theta^a(L)\Delta\varepsilon_{at} = \varepsilon_{at} + \sum_{j=0}^{\infty}(\theta_{j+1}^a - \theta_j^a)L^{j+1}\varepsilon_{at},$$

with $\theta_0^a = 1$.

For $\gamma^a(L)$, we denote $B = \phi_a + \phi_b$ and $C = \phi_a\phi_b - A$.

$$(1-\phi_a L - \phi_{ab}\phi_{ba}L^2 - \phi_{ab}\phi_{ba}\phi_b L^3)(1-\phi_b L) = (1 - BL + CL^2 + A\phi_b^2 L^4).$$

Therefore,

$$1 = (1 - BL + CL^2 + A\phi_b^2 L^4)(1+\gamma_1^a L + \gamma_2^a L^2 + \gamma_3^a L^3 + \ldots).$$

After some algebra

$$
\begin{aligned}
&\gamma_1^a = B \\
&\gamma_2^a = B^2 - C \\
&\gamma_3^a = B^3 + 2BC \\
&\gamma_4^a = B^4 - 3B^2C + C^2 - A\phi_b^2 \\
&\gamma_5^a = B^5 - 4B^3C + 3BC^2 - 2AB\phi_b^2 \\
&\gamma_6^a = B^6 - 5B^4C + 6B^2C^2 - 3AB^2\phi_b^2 + 2AC\phi_b^2 - C^3 \\
&\ldots
\end{aligned}
.
$$

Therefore,

$$\gamma^a(L)\Delta\varepsilon_{bt-1} = \varepsilon_{bt-1} + \sum_{j=0}^{\infty}(\gamma_{j+1}^a - \gamma_j^a)L^{j+1}\varepsilon_{bt-1}$$

with $\theta_0^a = 1$. Following the previous steps, we get similar expressions for $\theta^b(L)$ and $\gamma^b(L)$.

References

Abad, David, and Roberto Pascual, 2006, Switching to a temporary call auction in times of high uncertainty, *Working Paper,* SSRN Electronic Library.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics*, 17, 223-249.

Amihud, Yakov, and Haim Mendelson, 1987, Trading mechanisms and stock returns: and empirical investigation, *Journal of Finance*, 42, 533-553.

Bandi, F. and J. R. Russell, 2006, Microstructure noise, realized variance, and optimal sam-pling, *Working Paper*, Graduate School of Business, The University of Chicago.

Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard, 2007, Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise, *Working Paper*, Nuffield College, Oxford.

Beverige, S., and C. Nelson, 1981, A new approach to the decomposition of economics time series into permanent and transitory components wit particular attention to the measurement of the business cycle, *Journal of Monetary Economics*, 7, 151-174.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: on the compensation for illiquidity in stock returns, *Journal of Financial Economics*, 41, 441-464.

Brennan, Michael J., Tarum Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected returns, *Journal of Financial Economics*, 49, 345-373.

Copeland, Thomas E., and Dan Galai, 1983, Information effects on the bid-ask spread, *Journal of Finance*, 38, 1457-1469.

Demsetz, Harold, 1968, The cost of transacting, *Quarterly Journal of Economics*, 82, 33-53.

Easley, David, Soeren Hvidkjaer, and Maureen O’Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance*, 57, 2185-2221.

Easley, David, Nicholas M. Kiefer, and Maureen O’Hara, 1997, One day in the life o a very common stock, *Review of Financial Studies*, 10, 805-835.

Easley, David, and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics*, 19, 69-90.

Engle, Robert F., and A. J. Patton, 2004, Impacts of trades in a error-correction model of quote prices, *Journal of Financial Markets*, 7, 1-25.

Escribano, Álvaro, and Roberto Pascual, 2006, Asymmetries in bid and ask responses to innovations in the trading process, *Empirical Economics*, 30, 913-946

Foucault, Thierry, 1999, Order flow composition and trading costs in a dynamic limit order market, *Journal of Financial Markets*, 99-134.

George, Thomas J., Gautam Kaul, and M. Nimalendran, 1991, Estimation of the bid-ask spread and its components: a new approach, *Review of Financial Studies*, 4, 623-656.

Glosten, Lawrence R., 1994, Is the electronic open limit order book inevitable?, *Journal of Finance*, 49, 1127-1161.

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

Glosten, Lawrence R., y Paul R. Milgrom, 1985, Bid, ask and transaction prices in specialist market with heterogeneously informed traders, *Journal of Financial Economics*, 14, 71-100.

Gouriéroux, C., Monfort, A. and A. Trognon, 1984, Pseudo Maximum Likelihood Methods: Theory, *Econometrica*, 52, 681-700.

Handa, Puneet, Robert A. Schwartz, and Ashish Tiwari, 2003, Quote setting and price formation in an order driven market, *Journal of Financial Markets*, 6, 461-489.

Hasbrouck, Joel, 1993, Assessing the quality of a security market: A new approach to transaction costs measurement, *Review of Financial Studies*, 6, 191-212.

Hasbrouck, Joel, 1995, One security, many markets: determining the contributions to price discovery, *Journal of Finance*, 50, 1175-1199.

Hasbrouck, Joel, 1996, Modeling microstructure time series, in G.S. Maddala, and C.R. Rao, Eds., *Handbook of Statistics 14: Statistical Methods in Finance*, Elsevier North Holland, Amsterdam.

Hasbrouck, Joel, 1999, The dynamics of discrete bid and ask quotes, *Journal of Finance*, 6, 2109-2142.

Hasbrouck, Joel, 2002, Stalking the efficient price in market microstructure specifications: an overview, *Journal of Financial Markets*, 5, 329-339.

Hasbrouck, Joel, 2007, *Empirical Market Microstructure*, Oxford University Press.

Huang, Roger D., and Hans R. Stoll, 1996,

Huang, Roger D., and Hans R. Stoll, 1997, The components of the bid-ask spread: a general approach, *Review of Financial Studies*, 10, 995-1034.

Ho, Thomas, and Hans R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics*, 31, 319-330.

Ho, Thomas, and Hans R. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance*¸ 38, 1053-1074.

Laux, Paul, 1995, Dealer market structure, outside competition, and the bid-ask spread, *Journal of Economic Dynamics and Control*, 19, 683-710.

Lehmann, Bruce N., 2002, Some desiderata for the measurement of price discovery across markets, *Journal of Financial Markets*, 5, 259-276.

Lin, Ji-Chai, Gary Sanger, and Geoffrey Booth, 1995, Trade size and components of the bid-ask spread, *Review of Financial Studies*, 8, 1153-1183.

Madhavan, Ananth, Matthew Richardson, and Mark Roomans, 1997, Why do security prices change? A transaction-level analysis of NYSE stocks, *Review of Financial Studies*, 10, 1035-1064.

Pardo, Angel, and Roberto Pascual, 2006, On the hidden side of liquidity, *Working Paper*, SSRN Electronic Library.

Pascual, Roberto, and David Veredas, 2007, Does the limit order book matters in explaining long run volatility?, *Working Paper,* SSRN Electronic Library.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance*, 39, 1127-1139.

Stoll, Hans R., 1989, Inferring the components of the bid-ask spread: theory and empirical tests, *Journal of Finance*, 19, 115-134.

Stoll, Hans, R., 2000, Friction, *Journal of Finance*, 55, 4, 1479-1514.

World Federation of Exchanges, 2005, *Annual Report and Statistics*, WFE.

TABLE I
Sample Statistics

This table provides daily average statistics for our sample of SSE-listed stocks. We provide statistics for the complete sample (69 stocks), and for the subsamples of index constituents (IDX) and non-index constituents (NIDX). Additionally, we split the IDX stocks into two groups, the 5 largest stocks in the market and the rest. Market capitalization is measured at the end of 2000, in euros. "Trades" is the number of trades per day. Trading volume is measured in number of shares. The relative spread and the displayed book depth (five best ask and bid quotes) are averaged weighting by time. "Order flow" is the number of limit orders submitted. "Volatility" is the daily high-low ratio of the quote midpoint. The table reports the median and inter-quartile range across days for each variable and sample or subsample.

			Daily statistics						
Stocks		Statistic	Capitalization* (12-31-00)	Trades	Volume** (shares)	Relative spread	Displayed Book depth	Order flow	Volatility
Sample		Median	1228.59	144.78	147.72	0.0054	6879.69	164.24	1.0263
(69 stocks)		IQR	(2138.58)	(362.31)	(405.99)	(0.0052)	(10102.31)	(317.87)	(0.0083)
IDX									
(34 stocks)									
	All	Median	2339.34	390.77	441.46	0.0035	8182.47	364.05	1.0286
		IQR	(4982.11)	(612.84)	(1163.01)	(0.0002)	(48330.20)	(1121.66)	(0.0053)
	5-largest	Median	50654.25	1990.89	7853.02	0.0012	59390.75	1165.66	1.0261
		IQR	(30697.46)	(1329.55)	(6901.59)	(0.0001)	(44492.20)	(595.69)	(0.0024)
	Others	Median	2037.50	335.42	358.15	0.0036	7323.76	307.32	1.0300
		IQR	(2050.26)	(271.05)	(621.09)	(0.0016)	(7957.49)	(270.50)	(0.0101)
NIDX									
(35 stocks)									
	All	Median	444.75	44.86	51.57	0.0088	6557.93	56.38	1.0233
		IQR	(528.24)	(52.15)	(58.69)	(0.0044)	(6050.74)	(36.69)	(0.0106)

* $x10^{-6}$, **$x10^{-3}$

TABLE II
Estimation of the Structural Model

This table summarizes the estimation of the co-integration state space model for ask and bid quotes. The parameters β, ϕ_a, ϕ_b correspond to half the average bid-ask spread, and the persistency of the transitory components of ask and bid quotes, as shown in [T1] and [T2], where $S_{a,t}$ and $S_{b,t}$ are the transitory component of ask and bid quotes, and m_t is the efficient price.

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} S_{at} \\ S_{bt} \\ m_t \end{pmatrix} \qquad \text{[T1]}$$

$$\begin{pmatrix} S_{a,t} \\ S_{b,t} \\ m_t \end{pmatrix} = \begin{pmatrix} \phi_a & \phi_{ab} & 0 \\ \phi_{ba} & \phi_b & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} S_{a,t-1} \\ S_{b,t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{a,t} \\ \varepsilon_{b,t} \\ \varepsilon_{m,t} \end{pmatrix} \qquad \text{[T2]}$$

The parameters σ_a^2 and σ_b^2 correspond to the average transitory volatility of ask and bid quotes, represented in equation [T3], and α_0, α_1, α_2, and α_3 are the parameters of the EGARGH model in [T4] that represents the volatility of the efficient price.

$$\sigma_{i,t}^2 = \exp\left(\sigma_{i,k}^2 + \sum_{\substack{j=1 \\ j \neq k}}^{J} \phi_j^i D_{t,j} \right), \; i = \{a, b\} \qquad \text{[T3]}$$

$$\sigma_{m,t}^2 = \exp\left(\alpha_0 + \alpha_1 \ln\left(\sigma_{m,t-1}^2\right) + \alpha_2 \xi_{m,t-1} + \alpha_3 \left[\left| \xi_{m,t-1} \right| - \sqrt{2/\pi} \right] + \sum_{\substack{j=1 \\ j \neq k}}^{J} \phi_j^m D_{t,j} \right) \qquad \text{[T4]}$$

The model is estimated with the Kalman filter. We provide cross sectional medians and inter-quartile ranges for the complete sample of SSE stocks, and for the subsamples of index constituents (IDX) and non-index constituents (NIDX). Additionally, we split the IDX stocks into two groups, the 5 largest stocks in the market and the rest.

TABLE II
Estimation of the Structural Model (Cont)

Stocks		Statistic	β	ϕ_a	ϕ_b	ϕ_{ab}	ϕ_{ba}	$\sigma_{a,k}^*$	$\sigma_{b,k}^*$	α_0	α_1	α_2	α_3
Sample		Median	0.0234	0.8478	0.6176	0.2694	0.1373	0.1767	0.1264	-0.3756	0.9465	0.0150	0.2472
(69 stocks)		IQR	(0.0276)	(0.1620)	(0.2438)	(0.2421)	(0.1756)	(0.3683)	(0.2640)	(0.2797)	(0.0348)	(0.0427)	(0.1268)
IDX													
(34 stocks)													
	All	Median	0.0207	0.7890	0.6246	0.2592	0.2174	0.1828	0.1553	-0.4000	0.9402	0.0093	0.3102
		IQR	(0.0231)	(0.1233)	(0.1585)	(0.1368)	(0.1371)	(0.4552)	(0.3468)	(0.3507)	(0.0548)	(0.0464)	(0.1337)
	5-largest	Median	0.0080	0.6069	0.4756	0.3782	0.4218	0.0295	0.0503	-0.1425	0.9759	-0.0145	0.1578
		IQR	(0.0032)	(0.0634)	(0.0666)	(0.0532)	(0.0528)	(0.0298)	(0.0580)	(0.2452)	(0.0329)	(0.0056)	(0.0215)
	Others	Median	0.0237	0.7955	0.6376	0.2464	0.2166	0.2482	0.2051	-0.4418	0.9257	0.0194	0.3173
		IQR	(0.0250)	(0.0906)	(0.1362)	(0.1171)	(0.0851)	(0.5119)	(0.3548)	(0.3750)	(0.0546)	(0.0405)	(0.1532)
NIDX													
(35 stocks)													
	All	Median	0.0283	0.9012	0.5839	0.3239	0.0561	0.1623	0.1200	-0.3562	0.9488	0.0183	0.2160
		IQR	(0.0299)	(0.0817)	(0.3300)	(0.7193)	(0.0773)	(0.1538)	(0.2371)	(0.2116)	(0.0244)	(0.0348)	(0.0849)

$* \; 10^{+3}$

TABLE III
Principal Components Analysis

This table summarizes a cross-sectional regression analysis where the dependent variable is a quote quality measure, either the Variance Information Share (VIS) of the ask quote or bid quote or the Quote Quality Measure (QQM). VIS measures how much of the volatility of either quote is information-related. QQM measures how much ask and bid quotes track the efficient price, which is assumed to be a random walk with heteroskedastic innovations. The data includes information for 34 index constituents in the Spanish Stock Exchange. The explanatory variables are the two first principal components of the following measures averaged across days for each stock: (a) the relative spread: the average bid-ask spread divided by the quote midpoint, both weighted by time; (b) the limit order book depth: the accumulated euro value of the volume offered at the five best ask and bid levels of the limit order book, also weighted by time; (c) volume: the number of shares transacted; trades: the number of trades completed; order flow: the number of limit orders, market orders, and cancellations submitted, and market capitalization is measured at 12-31-2000. The first principal component is a proxy for illiquidity. The second principal component is mainly determined by the relative spread,

$$PC_i^1 = (0.4246)C_i + (-0.3379)RS_i + (0.4197)D_i + (0.4241)V_i + (0.4239)T_i + (0.4121)OF_i$$
$$PC_i^2 = (0.1969)C_i + (0.9151)RS_i + (0.0589)D_i + (0.3313)V_i + (0.0876)T_i + (0.0563)OF_i$$

The model has been estimated by Least Absolute Deviation (LAD) with iteratively re-weighted least squares. Standard errors are computed with 200 bootstrap draws. We also report the robust adjusted-R2 of MacKean and Sievers (1987). We provide separated findings for quotes sampled at 1-minute, 5-minute, and 10-minute intervals.

Principal		1 minute			5 minutes			10 minutes		
Component		VIS[a]	VIS[b]	QQM	VIS[a]	VIS[b]	QQM	VIS[a]	VIS[b]	QQM
PC1	Coef.	0.0393	0.0403	0.0404	0.1080	0.1009	0.0996	0.0459	0.0509	0.0464
	Std. Dev.	0.0099	0.0092	0.0094	0.0809	0.0719	0.0752	0.0236	0.0155	0.0215
	t-Stat	3.9678	4.3984	4.3224	1.3358	1.4031	1.3254	1.9507	3.2792	2.1535
PC2	Coef.	-0.0301	-0.0271	-0.0300	-0.1653	-0.1549	-0.1809	-0.0292	-0.0123	-0.0251
	Std. Dev.	0.0225	0.0190	0.0239	0.3342	0.2884	0.3852	0.0550	0.0326	0.0466
	t-Stat	-1.3371	-1.4251	-1.2547	-0.4945	-0.5371	-0.4697	-0.5307	-0.3783	-0.5382
	R^2	0.4423	0.5203	0.4841	0.5672	0.5860	0.4963	0.4075	0.5737	0.5138

FIGURE 1
Intraday Regularities in Transitory Volatility

This figure reports estimated regular patterns in transitory volatility. Volatility is decomposed in its theoretical components using the co-integration state-space model given by equations [1] to [4] in the paper. The model is estimated with the Kalman filter. Transitory volatility is described by the deterministic process in [F1]. We split the trading session of the Spanish Stock Exchange into the following time intervals: [9:00, 9:30), [9:30, 10:00), [10:00, 13:00), [13:00, 15:00), [15:00, 15:30), [15:30, 16:00), [16:00, 16:30), [16:30, 17:00), [17:00, 17:30]. The figure reports the cross-sectional median of the each parameter expressed in deviations with respect to the control interval [13:00, 15:00). We provide our findings for the complete sample of SSE stocks, and for the subsamples of index constituents (IDX) and non-index constituents (NIDX). Additionally, we split the IDX stocks into two groups, the 5 largest stocks and the rest.

$$\sigma_{i,t}^2 = \exp\left(\sigma_{i,k}^2 + \sum_{\substack{j=1 \\ j\neq k}}^{J} \phi_j^i D_{t,j}\right),\ i = \{a,b\} \qquad [F1]$$

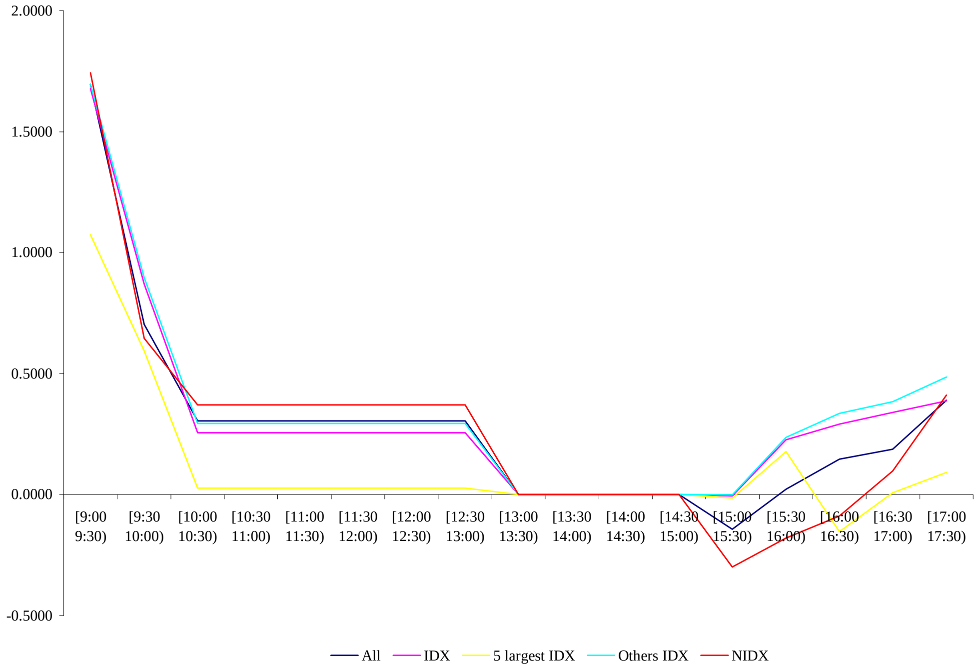


FIGURE 2
Intraday Regularities in Informational Volatility

This figure reports estimated regular patterns in informational (efficient) volatility. Volatility is decomposed in its theoretical components using the co-integration state-space model given by equations [1] to [4] in the paper. The model is estimated with the Kalman filter. Informational volatility is described by the EGARCH process in [F2]. We split the trading session of the Spanish Stock Exchange into the following time intervals: [9:00, 9:30), [9:30, 10:00), [10:00, 13:00), [13:00, 15:00), [15:00, 15:30), [15:30, 16:00), [16:00, 16:30), [16:30, 17:00), [17:00, 17:30]. The figure reports the cross-sectional median of the each parameter expressed in deviations with respect to the control interval [13:00, 15:00). We provide our findings for the complete sample of SSE stocks, and for the subsamples of index constituents (IDX) and non-index constituents (NIDX). Additionally, we split the IDX stocks into two groups, the 5 largest stocks and the rest.

$$\sigma_{m,t}^2 = \exp\left(\alpha_0 + \alpha_1 \ln\left(\sigma_{m,t-1}^2\right) + \alpha_2 \xi_{m,t-1} + \alpha_3 \left[\left|\xi_{m,t-1}\right| - \sqrt{2/\pi} \right] + \sum_{\substack{j=1 \\ j \neq k}}^{J} \phi_j^m D_{t,j} \right) \qquad [F2]$$

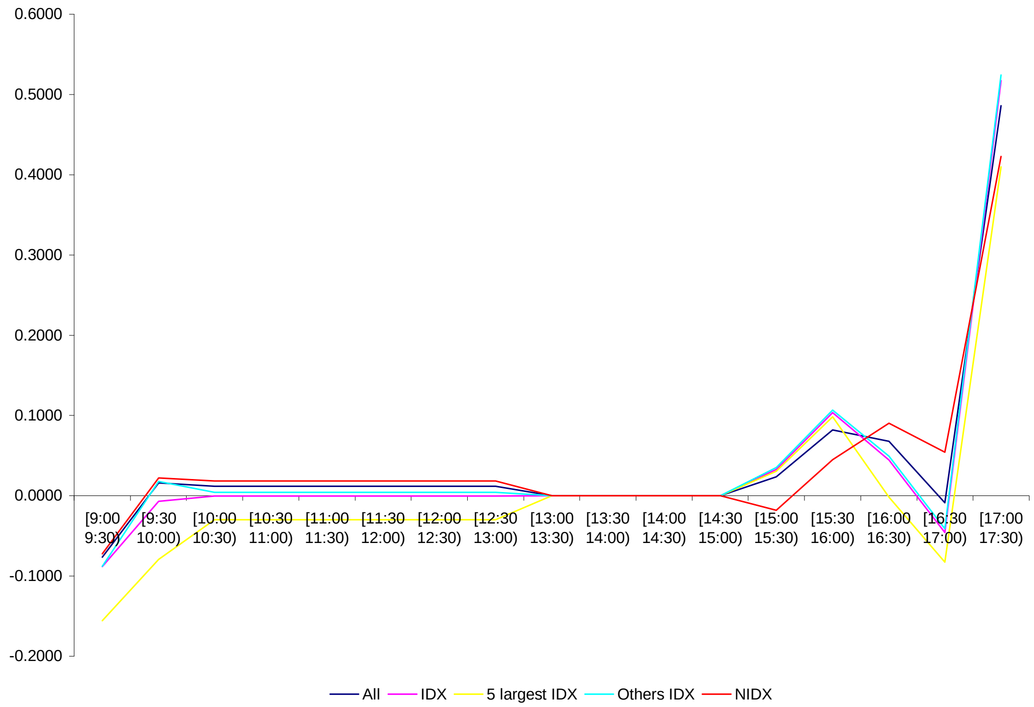


FIGURE 3
Variance Information Measure (VIS) for IDX Stocks

This figure reports the estimated quote quality measure Variance Information Share (VIS) for the subsample of index constituents (IDX) of the Spanish Stock Exchange. VIS is defined as the fraction of total variance of ask (blue bars) and bid (light bars) quotes which is explained by the long-run component of quotes, assumed to be a random walk component with heteroskedastic innovations. See equations [1]-[4] and [8]-[9] in the paper for details on the structural model. VIS is computed using the reduced form of the structural model. The closer VIS is to one, the higher the quality of ask and bid quotes.

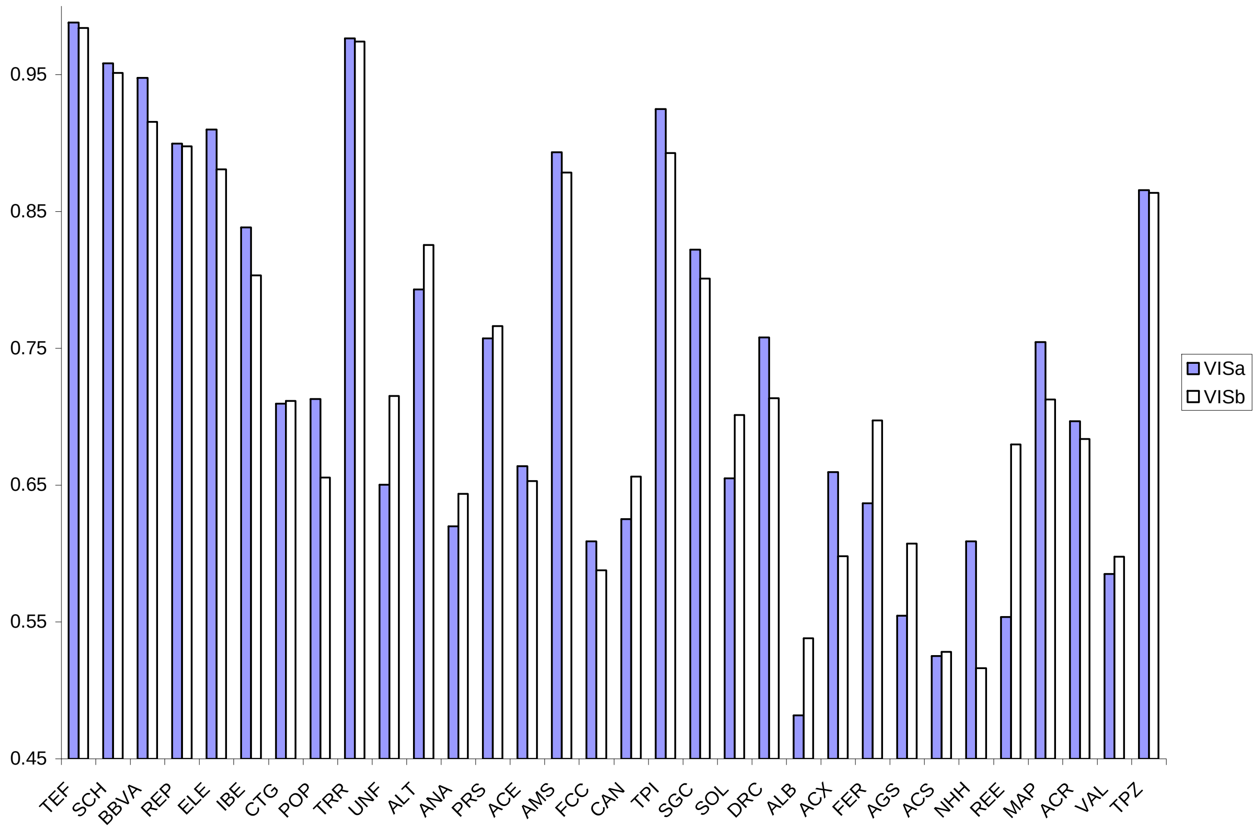


FIGURE 4
Quote Quality Measure (QQM) for IDX Stocks

This figure reports the estimated quote quality measure (QQM) for the subsample of index constituents (IDX) of the Spanish Stock Exchange. QQM is defined as the fraction of correlation between ask and bid quotes which is explained by the long-run (common) component of quotes, which is assumed to be a random walk component with heteroskedastic innovations. See equations [1]-[4] and [8]-[9] in the paper for details on the structural model. QQM is obtained from the reduced form of the structural model. The closer QQM is to one, the higher the quality of ask and bid quotes.

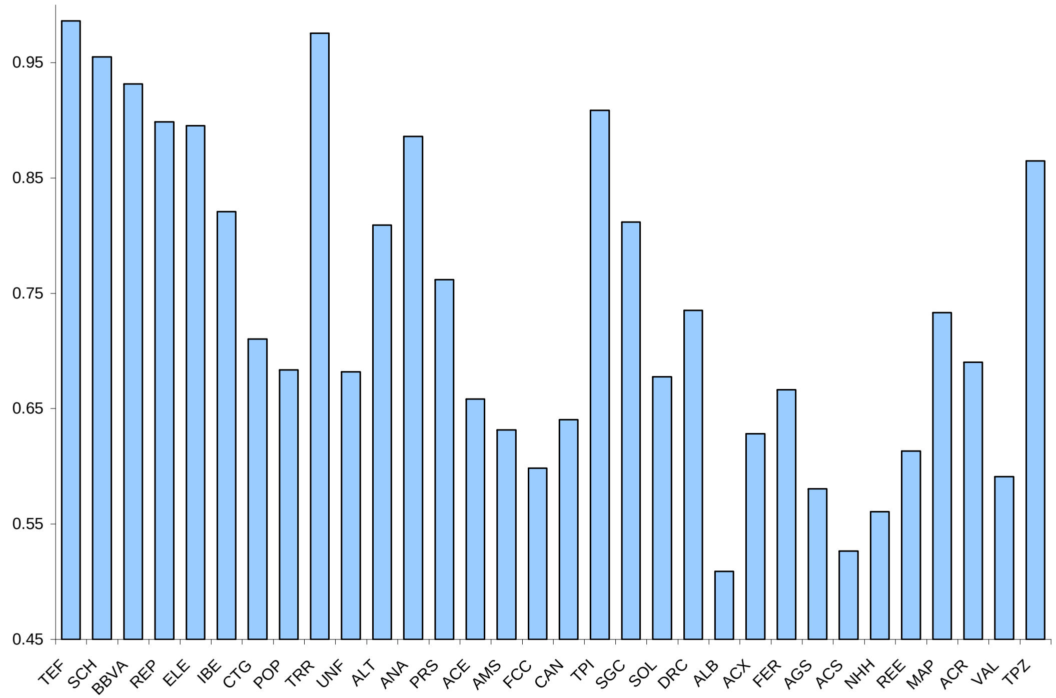


FIGURE 5
Quote Quality Measure (QQM) for IDX Stocks at Different Time Frequencies

This figure reports the estimated quote quality measure (QQM) for the subsample of index constituents (IDX) of the Spanish Stock Exchange at three different sample frequencies, 1 minute, 5 minutes, and 10 minutes. QQM is defined as the fraction of correlation between ask and bid quotes which is explained by the long-run (common) component of quotes, which is assumed to be a random walk component with heteroskedastic innovations. See equations [1]-[4] and [8]-[9] in the paper for details on the structural model. QQM is obtained from the reduced form of the structural model. The closer QQM is to one, the higher the quality of ask and bid quotes.

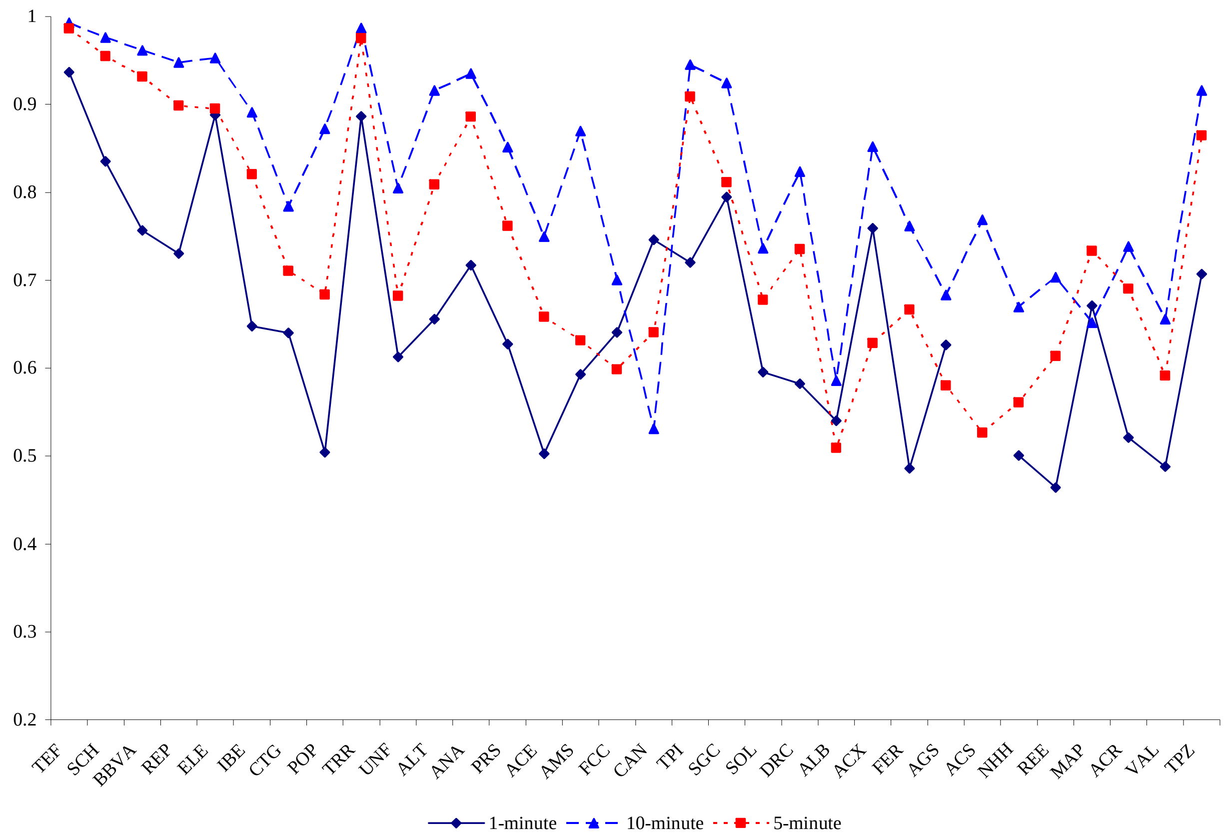
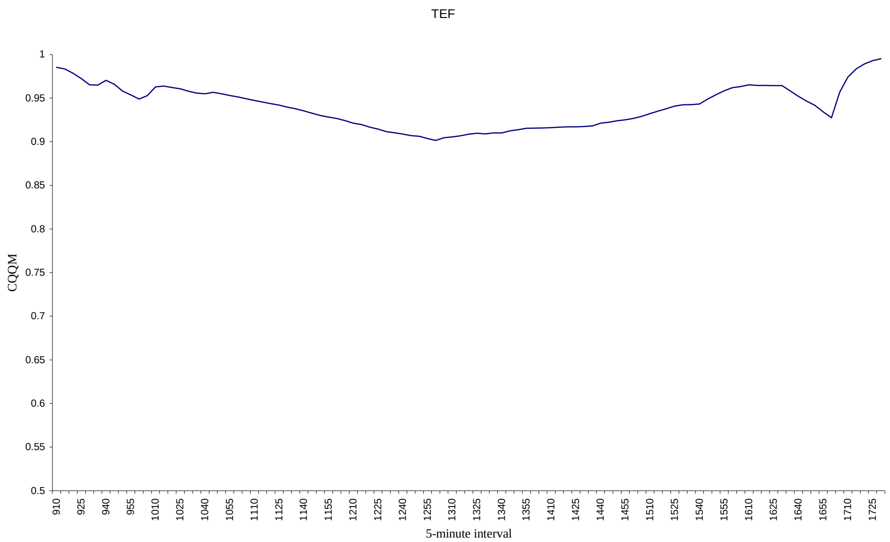


FIGURE 6
Intraday Conditional Quote Quality Measure (CQQM) for three IDX Stocks

This figure reports the estimated conditional quote quality measure ($CQQM_t$) for three index constituents (IDX) of the Spanish Stock Exchange, TEF, TRR and ANA, with a sample frequencies of 5 minutes. $CQQM_t$ is defined as the fraction of correlation between ask and bid quotes which is explained by the long-run (common) component of quotes, which is assumed to be a random walk component with heteroskedastic innovations, at a given time interval of the trading session. See equations [1]-[4] and [8]-[9] in the paper for details on the structural model. The $CQQM_t$ is obtained from the reduced form of the structural model. The closer $CQQM_t$ is to one, the higher the quality of ask and bid quotes. The figures report the average value of $CQQM_t$ for each stock across days at each five minute interval of the trading session.



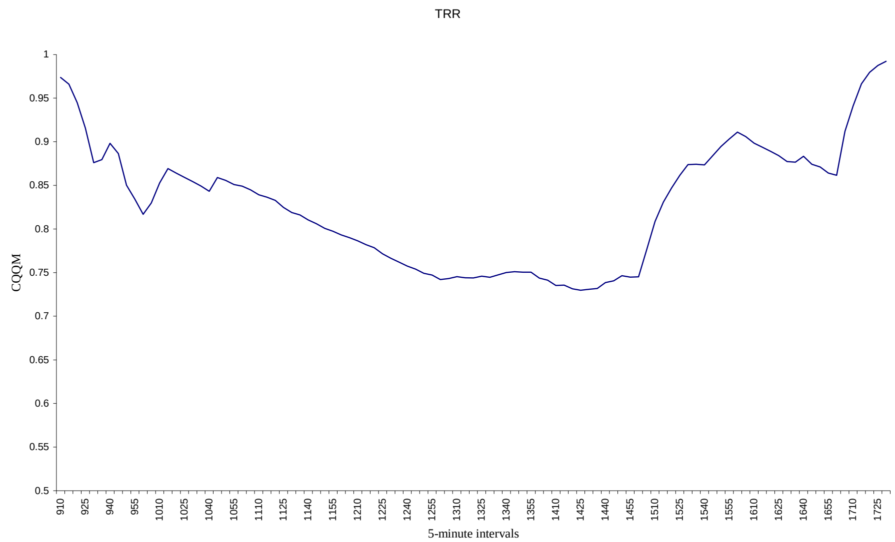


FIGURE 6 (Cont.)
Intraday Conditional Quote Quality Measure (CQQM) for three IDX Stocks