

CROSS-LISTING, PRICE DISCOVERY AND THE INFORMATIVENESS OF THE TRADING PROCESS

Roberto Pascual, Bartolomé Pascual-Fuster and Francisco Climent*

WP-EC 2003-21

Correspondence to: Francisco Climent: Dpto. Economía Financiera, Universitat de València, Avendida de los naranjos, s/n, 46022 Valencia, (Spain), tel: 34-963828369, fax: 34-963828370, e-mail: F.Jose.Climent@uv.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Noviembre 2003

Depósito Legal: V-5026-2003

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

* R. Pascual y B. Pascual-Fuster: Dpto. Economía de la Empresa, Universidad de las Islas Baleares; F. Climent: Dpto. Economía Financiera, Universitat de València.

CROSS-LISTING, PRICE DISCOVERY AND THE INFORMATIVENESS OF THE TRADING PROCESS

Roberto Pascual, Bartolomé Pascual-Fuster and Francisco Climent

ABSTRACT

This paper analyzes the price discovery process of securities that trade at multiple markets with trading sessions that totally or partially overlap. Building on Hasbrouck (1995) information share approach, we introduce a methodology that distinguishes two sources of information asymmetries between markets: trade-related and trade-unrelated informative shocks. This approach determines how much of each market's relative contribution to the price discovery process during the overlapping period is attributable to its own trading activity. We provide empirical evidence on the contribution of the NYSE in the price discovery process of the Spanish cross-listed stocks during the daily two-hour overlapping interval.

Key words: Cross-listing, price discovery, trade shocks, ADRs.

RESUMEN

Este trabajo analiza el proceso de formación del precio de acciones negociadas en varios mercados con sesiones de negociación total o parcialmente solapadas. A partir del modelo propuesto por Hasbrouck (1995), introducimos una metodología que distingue dos fuentes de asimetría de información entre los mercados: la inferida a partir de la negociación de cada mercado e información ajena a la negociación. El objetivo es determinar en qué medida la contribución relativa de cada mercado al proceso de formación del precio durante el período de solapamiento es relevante y que parte de esta es atribuible a su propia actividad de negociación. Se obtiene evidencia empírica de la contribución del NYSE en el proceso de formación del precio de las acciones españolas durante el intervalo de dos horas de solapamiento diario.

PALABRAS CLAVE: Negociación en varios mercados, Formación del precio, Shocks de negociación, ADRs

1. Introduction

When an asset is traded at multiple markets, a crucial question naturally arises: which market does contribute more to the discovery of the efficient price. In the last years, alternative methodological approaches have been proposed to measure these relative contributions during intervals when the trading sessions of the different markets totally or partially overlap. Harris et al. (1995) use a vector error correction (VEC) model to study the adjustment mechanism of the NYSE and regional prices towards the common underlying efficient price. A significant error correction term in the NYSE price equation indicates that regional markets do contribute to the price discovery of NYSE-listed stocks. Hasbrouck (1995) proposes a common trend representation to model the NYSE and regional quotes. The fraction of the long-term variance (the variance of the common stochastic trend) that is explained by each market innovations defines its information share. Hasbrouck finds that the information share of the regional markets is relatively unimportant. Finally, Harris et al. (2002) uses the common factor estimation method proposed by Gonzalo and Granger (1995) to evaluate each market's proportion of the price discovery. In this methodology the long memory component of stock prices is characterized as a weighted average of the contemporaneous trade prices. The weights signify the incidence of trades that permanently move prices on each market. Their findings show changes in the location of price discovery over time. A recent special issue of the *Journal of Financial Markets* (Vol. 5, No.3) provides some discussion on the differences and similarities between these econometric methodologies.

Tse (2000) argues that different data sets (quotes versus trades) rather than models or periods matter when measuring relative contributions to price discovery. Thus, trades executed on the foreign/regional markets may provide information to domestic/home traders even if the economic meaning of the foreign/regional quotes is irrelevant. Therefore, a methodology intended to study the price discovery process of dually listed stocks should take into account at least both quotes and trades. Current methodologies cannot differentiate between alternative sources of information simply because the trading activity is not openly modeled. To the extent that a transaction conveys useful information, this should be reflected in the quotes of the exchange that is first aware of the transaction. Therefore, by decomposing the quote innovations into trade-related and trade-unrelated shocks we should be able to determine how much of each market's relative contribution is attributable to its trading activity. In addition, existing methodologies provide an incomplete characterization of a dominant-satellite

relationship because the trading process of a given market may be informative for the other markets even when the quotes are not informative at all. By modeling the trading processes of the different markets, we would allow for incremental information contained in trade shocks, beyond that reflected in each market's quotes.

In this paper, we propose a simplistic theoretical framework for two markets that simultaneously trade one stock. Each market forms its sequence of conditional expectations about the security's ultimate value drawing on the revisions of their available information set. We consider two possible sources of information asymmetries between markets. On the one hand, the presence of informed agents endowed with superior information about the true value of the stock. This information, we assume, is revealed through trading. Because informed agents must decide where to exploit their information advantage (e.g., Chowdhry and Nanda, 1991), trade-related information may cause transitory differences in the markets' expectations about the true value of the stock. On the second hand, public announcements, characterized as noisy signals (e.g., Harris and Raviv, 1993), may also cause information asymmetries. This trade-unrelated information is simultaneously exposed to all markets but they differ in their ability to process it (e.g., Kim and Verrecchia, 1994). A given public announcement provides the market with superior processing capacity with a temporary advantage over the other market. In this context, a pure satellite market has an uninformative trading process and is incapable of interpreting public announcements.

We show that the natural empirical counterpart of our theoretical framework is a VEC model that explicitly identifies the informative (unexpected) component of trading. Following Hasbrouck (1995), we use the common trend representation of this VEC model to measure each market's trading activity contribution to the long-term volatility of the stock (trade-related information shares). We provide empirical evidence on the contribution of the NYSE to the price formation of the Spanish stocks that were listed in the US market as ADRs during 2000. We center the attention on the daily two-hour overlapping interval between the NYSE and the Spanish Stock Exchange (SSE). For the most frequently traded NYSE-listed Spanish stock, we are unable to characterize the NYSE as a pure satellite market of the SSE. Nonetheless, we strongly accept the null hypothesis that the NYSE trading activity does not contribute to the price discovery of the Spanish ADRs. These findings are robust to alternative specifications of the empirical model.

The paper is organized as follows. In section 2, we present the theoretical framework. In section 3, we introduce and discuss the empirical model. In section 4, we define the information share measures. In section 5, we describe the data set. In section

6, we provide the estimation results and the information shares for the NYSE and the SSE and discuss some refinements of the empirical analysis. Finally, we conclude in section 7.

2. Motivation

Consider a stock that trades at two different markets, the home or domestic market (D) and an alternative/foreign market (F), with trading sessions that overlap during a given time interval. Without any loss of generality, suppose that F closes after D. Each market's information set updates because of trade-related and trade-unrelated information shocks. In this simplified world, we equate trade-unrelated shocks and public disclosures and trade-related shocks to private information (a strong assumption, as discussed in the next section). To start with, consider the case in which public disclosures are not noisy signals. The two markets adjust quoted prices at the same time and by the same amount after a trade-unrelated shock. Thus, we impose that the unique source of information asymmetries between markets is the trade-inferred information. Let m_t be the expected true value of the stock given the full information set at moment t . That is,

$$m_t = E[\psi_\tau | \phi_t], \quad [1]$$

where $E[.]$ is the conditional expectation, ψ_τ is the true value of the stock in a future reference moment τ (for example, the end of trading at F) and ϕ_t is the total information available, both common and uncommon to both markets, at moment t . We assume that this information is fully inferred from the time series of quotes and trades. The common information set (ϕ_t^C) includes the current and all previous trade-unrelated shocks and the whole history of trade-related shocks up to period $t-1$. Since informed agents must decide where to exploit their information advantage, the trade-related information might be revealed earlier either at market D (ϕ_t^D) or at market F (ϕ_t^F). Under this structure, we have that $\phi_t = \{\phi_t^C, \phi_t^D, \phi_t^F\}$. Given that, at some point in time, the information sets available for markets D and F may differ, their expectations about ψ_τ may also be different. Therefore, let the market i 's expectation at period t be,

$$m_t^i = E[\psi_\tau | \phi_t^C, \phi_t^i], \quad i=\{D, F\}. \quad [2]$$

The market i 's expectation follows a random walk process,

$$m_t^i = m_t^* + \lambda_i w_t^i, \quad [3]$$

where m_t^* represents the expected true value of the stock when only the common information set is available, that is $m_t^* = E[\psi_\tau | \phi_t^C]$. The stochastic process w_t^i in [3], with $E[w_t^i] = 0$, $E[(w_t^i)^2] = \sigma_i^2$ and $E[w_t^i w_{t-j}^i] = 0 \quad \forall j \neq 0$, characterizes a trade-related innovation that updates the market i 's information set. We allow w_t^D and w_t^F being contemporaneously correlated, $E[w_t^D w_t^F] \neq 0$; this could be the case where the same informed investor or different traders endowed with the same privileged information trade simultaneously at different markets. For instance, Chowdhry and Nanda (1991) find that trade-related shocks across markets are correlated in high information periods. The parameter λ_i measures how much of a trade-related shock is informative.

The expectation based on the common information set (m_t^*) follows a random walk process,

$$m_t^* = m_{t-1} + w_t, \quad [4]$$

with w_t being a zero-mean uncorrelated stochastic process representing an innovation in the common information set ϕ_t^C , i.e. $\phi_t^C = \{\phi_{t-1}^C, w_t\}$, due to a public disclosure at period t . Notice that equation [4] implies that any specific information given away at period $t-1$, either at market D or at market F, becomes common information the next period ($\{\phi_{t-1}^D, \phi_{t-1}^F\} \subset \phi_t^C$). This imposes short-term convergence in expectations between both markets. Thus,

$$m_t^i = m_{t-1} + (\lambda_i w_t^i + w_t). \quad [5]$$

Consequently, the revisions in the expectations about the true value of the stock have a common component (w_t) and a market-specific component ($\lambda_i w_t^i$). For example, if $w_t^F = 0$ and $w_t^D \neq 0$ we have that $m_t^F = m_t^* = m_{t-1} + w_t$ and $m_t^D = m_t^*$. That is, D has a more precise expectation at period t than F. In this scenery, F would behave as a pure satellite market if all the trade-related information were disseminated through the D's trading activity (i.e. $\lambda_F = 0$). We impose the restriction that the trade-related shocks are uncorrelated with the shocks motivated by public announcements: $E[w_t^i w_{t-j}^i] = 0, \forall j$.

Quotes are the result of the firm demand and offer positions by liquidity providers. In this paper, we average quotes using the quote midpoint of the bid-ask spread, $q_t^i = (a_t^i + b_t^i)/2$, where a_t^i and b_t^i represent the market i 's best ask and bid quotes. These quotes incorporate all the information revealed up to period t . The change in quotes at t will be $\Delta q_t^i = q_t^i - q_{t-1}^i$. Let,

$$q_t^i = m_t^i + S_t^i, \quad [6]$$

where S_t^i satisfies that $E(S_t^i) = 0$, $E(S_t^i S_{t-k}^i) = \sigma_k \forall k \geq 0$ and $\forall i$. Therefore, S_t^i is a covariance-stationary or weak-stationary stochastic component of quoted prices. It captures transitory deviations between the quote midpoint and the efficient price. The size of $|S_t^F - S_t^D|$ depends on disparities in market-making costs, market frictions (like the tick size), and other specific features of the microstructures of both markets. Equation [6] implies that q_t^i is a non-stationary process since it depends on a long-run component (m_t^*) that is integrated of order 1. This non-stationary component is common to both markets. Therefore, there exists a linear combination of their quotes that does is stationary,

$$q_t^D - q_t^F = (\lambda_D w_t^D - \lambda_F w_t^F) + (S_t^D - S_t^F). \quad [7]$$

The difference between the quote midpoints is a stationary stochastic process, meaning that both prices are co-integrated with a theoretical co-integration vector $[1, -1]$. The co-integration condition is necessary to avoid profitable arbitrage opportunities.¹

Hereafter, we relax the assumption that public disclosures are not noisy signals. Hence, we allow for a second source of information asymmetries between markets: their respective ability to evaluate public signals. Public disclosures constitute imperfect information, in the sense that the valuable information is communicated with some distortion. A given market may have more capacity to identify useful information either because it more closely monitors the firm or because it has access to more complete information about the public signals. Therefore, markets differ in the quality of their

¹ As an extreme case, consider that all public signals are first observed by the domestic market and transmitted to the foreign market with some lag. Let the market F's quote be given by $m_t^F = m_{t-1}^F + \lambda_F w_t^F$, which explicitly indicates that the market F's expectations at period t do not account for w_t . Under this specification quotes are still co-integrated. Furthermore, under the convenient assumption that $(S_t^F - S_{t-1}^D)$ is close to zero: $q_t^F \approx q_{t-1}^D + \lambda_F (w_t^F + w_{t-1}^F)$. Hence, also in this case, when $\lambda_F = 0$ the foreign market would be a pure satellite market for the cross-listed stock.

judgments. Kim and Verrecchia (1994) develop a model in which earning announcements provide information that may lead to different interpretations (see also Harris and Raviv, 1993). Some traders possess special capabilities that allow them to make (informed) judgments that are superior to the judgments of other traders. We just adapt Kim and Verrecchia's framework here. Consider that the public signal is,

$$\tilde{w}_t = w_t + \delta_t, \quad [8]$$

where δ_t is a stochastic process representing a distortion in the announcement, $E[\delta_t] = 0$, $E[\delta_t^2] = \sigma_\delta^2$ and $E[\delta_t \delta_{t-j}] = 0 \quad \forall j > 0$. We assume that δ_t and w_t are mutually uncorrelated. Simultaneously to the dissemination of \tilde{w}_t each market observes,

$$J_t^i = \delta_t + \zeta_t^i, \quad [9]$$

where ζ_t^i measures the information a market gleans about the random error by more closely studying the firm, its financial reports and businesses.² Let $E[\zeta_t^i] = 0$, $E[(\zeta_t^i)^2] = \sigma_{\zeta,i}^2$ and $E[\zeta_t^i \zeta_{t-j}^i] = 0 \quad \forall j > 0$ and $\forall i$. Again, we assume that δ_t and ζ_t^i are mutually uncorrelated. The quality of the markets' judgment depends on the precision of J_t^i . When $\sigma_{\zeta,i}^2 = 0 \quad \forall i$, this specification is equivalent to [1]-[5]: the two markets perfectly isolate the noise (δ_t) from the valuable information (w_t). Hence, both markets adjust their expectations simultaneously and there are not disparities in their interpretations. It follows that, *ceteris paribus*, the market F's behavior approximates that of a satellite market when $\sigma_{\zeta,D}^2 / \sigma_{\zeta,F}^2$ tends to zero. In this case, the judgments of the F agents are so imprecise that the update in their expectations (m_t^F) is unreliable and, hence, transitory. We assume that the realization w_t becomes public knowledge the period after the announcement. Therefore, public signals only provide a temporary advantage to the market with the most accurate information about the signal. In the period immediately after the signal, markets' expectations will converge if there are no additional trade-related or trade-unrelated shocks. The market i 's expectations are updated as follows,

$$m_t^i = m_t^* + \lambda_i w_t^i + (\tilde{w}_t - J_t^i). \quad [10]$$

² An alternative specification would be to assume that the public announcement provides each market with a different signal. In this case, a market may be provided with an inferior signal than the other.

Let ρ be the correlation coefficient between $(\bar{w}_t - J_t^D)$ and $(\bar{w}_t - J_t^F)$, such that $1 \geq \rho \geq 0$. If $\rho=1$ both markets observe the same information and have the same posterior beliefs. If $\rho=0$ the two markets will have posterior fully heterogeneous beliefs. The common-knowledge conditional expectation (m_t^*) is now given by $m_t^* = m_{t-1} = m_{t-1}^* + \lambda_D w_{t-1}^D + \lambda_F w_{t-1}^F + w_{t-1}$. Notice that if $\sigma_{\xi,i}^2 \neq 0$ the change in the expectation of market i has a distorting component that will be corrected the next period. The co-integration result between D and F quote midpoints still holds for this alternative scenario.

Thus far, only co-integration between quotes has been considered. However, other co-integration relationships can also emerge in this theoretical framework. Let the market i 's trading process be a function of both common and private information, such that

$$\Delta x_t^i = \gamma_0 + \gamma^i \Delta m_t^* + w_t^i. \quad [11]$$

where Δx_t^i is a stationary stochastic process representing the market i 's net trading activity in period t , γ_0 is a constant term, $\gamma^i \Delta m_t^*$ is the predictable component of Δx_t^i given ϕ_t^C and w_t^i is the unexpected component of trading. A positive value of Δx_t^i implies more buyer-initiated trading than seller-initiated trading during that period. This equation characterizes the new information inferred from the trading process (w_t^D, w_t^F) as the unpredictable component of the net trading given the common information set (e.g., Subrahmanyam, 1997, and Kempf and Korn, 1999). Hence, we are assuming that x_t^i is the summation of current and previous impacts of both trade-related and trade-unrelated informative shocks. It is straightforward to check that x_t^i (Δx_t^i) is a non-stationary (stationary) process. In addition, the trade-related shock could be decomposed into its informative component, $\lambda_i w_t^i$, and its uninformative one, $(1 - \lambda_i) w_t^i$. This last component may be explained by the rational response of investors to market makers' inventory control strategies or simply by noise trading (e.g., Hasbrouck, 1996). Since the trading processes x_t^F and x_t^D both depend on the same non-stationary long-run component, it would be possible to find a stationary linear combination. Indeed, in this simplified framework it would also be likely to find co-integration relationships between trades and quotes, all these meaning that the appropriate empirical counterpart to our framework is an error correction model. Next section develops our empirical specification.

3. The empirical model

The most common efficient parameterization of a vector of co-integrated variables is, from the Granger's Representation Theorem in Engle and Granger (1987), a VEC model. Equation [12] presents the standard error correction representation of market quotes,

$$\begin{aligned} \Delta q_t^D &= \alpha^D (q_{t-1}^D - \beta q_{t-1}^F) + \Phi_D^D(L) \Delta q_{t-1}^D + \Phi_F^D(L) \Delta q_{t-1}^F + u_t^D \\ \Delta q_t^F &= \alpha^F (q_{t-1}^D - \beta q_{t-1}^F) + \Phi_D^F(L) \Delta q_{t-1}^D + \Phi_F^F(L) \Delta q_{t-1}^F + u_t^F \end{aligned} \quad [12]$$

with $\Delta q_t^i = (q_t^i - q_{t-1}^i)$. The terms $\Phi_k^i(L)$, for k and $i=\{D, F\}$, are autoregressive polynomials in the lag operator L ($L^z y_t = y_{t-z}$) having all their roots outside the unit circle. The component $(q_{t-1}^D - \beta q_{t-1}^F)$ is the normalized error correction term. Presumably, β is equal to 1. The α^i term is the response of the market i to a divergence with the other market's quotes. If both α^D and α^F are statistically significant, this would evidence that we are facing a two-way price discovery process (see Harris et al., 1995). From the discussion in the previous section, the innovations $u' = (u_t^D, u_t^F)$ in [12] include both the trade-related and the trade-unrelated shocks. Explicitly, let u_t^D be,

$$u_t^D = \tilde{\theta}_D^D(L) w_t^D + \tilde{\theta}_F^D(L) w_t^F + \varepsilon_t^D, \quad [13]$$

where $\tilde{\theta}_k^i(L)$ are stationary lag polynomials. Equation [13] captures usual features of intra-daily data due to market frictions and specific trading rules. Because of market frictions, q_t^i may not instantaneously reproduce all the information trades release at period t .³ Thus, the unexpected component of trades may have lagged effects on the quote midpoint (e.g., Hasbrouck, 1991a, and Pascual et al., 2003). The vector of trade-unrelated shocks $\varepsilon' = (\varepsilon_t^D, \varepsilon_t^F)$ incorporates the information inferred from public signals, $(\tilde{w}_t - J_t^D)$ and $(\tilde{w}_t - J_t^F)$ respectively, but also idiosyncratic features of each market that we do not model explicitly, like the tick size. If both markets have similar

³ The NYSE rules state that the specialist should maintain a fair and orderly market. This includes the responsibility of stabilizing prices in their assigned stocks. The specialist ensures that trading in the stocks moves smoothly throughout the day (e.g. Hasbrouck et al., 1993). In the SSE there is no specialist or figure alike. However, the existence of hidden orders and stopped orders may also delay the full revelation of the information behind trades. Alternatively, traders with heterogeneous priors and private information may take some intervals of trading to have their expectational differences resolved (e.g., Kyle, 1985).

abilities to judge public announcements, then ε_t^D and ε_t^F should be highly correlated. Hence, we expect $E(\varepsilon_t^D \varepsilon_t^F) \neq 0$ due to a common factor. As to identify the components in [13], let us modify [12] by allowing the quote process to be,

$$\begin{cases} \Delta q_t^D = \alpha^D(q_{t-1}^D - \beta q_{t-1}^F) + \Phi_D^D(L)\Delta q_{t-1}^D + \Phi_F^D(L)\Delta q_{t-1}^F + \theta_D^D(L)\Delta x_t^D + \theta_F^D(L)\Delta x_t^F + \varepsilon_t^D \\ \Delta q_t^F = \alpha^F(q_{t-1}^D - \beta q_{t-1}^F) + \Phi_D^F(L)\Delta q_{t-1}^D + \Phi_F^F(L)\Delta q_{t-1}^F + \theta_D^F(L)\Delta x_t^D + \theta_F^F(L)\Delta x_t^F + \varepsilon_t^F \end{cases} \quad [14]$$

where Δx_t^i is the market i 's net trading activity at period t . All lag polynomials are stationary. At this point, let the generating process of Δx_t^D (Δx_t^F would be alike) be given by,

$$\Delta x_t^D = \Pi_{x,D}^D(L)\Delta x_{t-1}^D + \Pi_{x,F}^D(L)\Delta x_{t-1}^F + \Pi_{q,D}^D(L)\Delta q_{t-1}^D + \Pi_{q,F}^D(L)\Delta q_{t-1}^F + w_t^D, \quad [15]$$

with $\Pi_{h,k}^i(L)$, for $h=\{x, q\}$ and $k=\{D, F\}$, being lag polynomials with all roots outside the unit circle. By substituting recursively the trading processes into [14], it is straightforward to see that [14]-[15] is equivalent to [12]-[13]. Notice that in [15] the trading process does not depend on the contemporaneous change in market quotes. This is because quotes and trades are not determined simultaneously: the quote revisions follow the net trading activity. The model allows causality running from lagged quote revisions to trades but not contemporaneously. This causality structure is very common among theoretical adverse selection costs models (e.g., Huang and Stoll, 1997).

The model [14]-[15] is, however, incomplete. The theoretical framework in previous section allowed for multiple co-integration relationships. Indeed, in our four variable VEC model, we could find up to three linearly independent combinations of $y_t' = (q_t^D \ q_t^F \ x_t^D \ x_t^F)$ which ensure that these variables converge to their long-run steady-state. Presumably, one of these linear combinations must be the $(1 \ -1 \ 0 \ 0)$ vector discussed earlier. Therefore, we end with an empirical model with 4 equations, with $r \leq 3$ co-integrating vectors and with causality running from net trading to quotes,

$$A\Delta y_t = \alpha\beta'y_{t-1} + B(L)\Delta y_{t-1} + \xi_t, \quad [16]$$

where $\xi_t' = (\varepsilon_t^D \ \varepsilon_t^F \ w_t^D \ w_t^F)$, α is a $4 \times r$ matrix representing the speed of adjustment to disequilibrium, β is the $4 \times r$ matrix of long-run coefficients,

$$A = \begin{bmatrix} 1 & 0 & -\theta_{D,0}^D & -\theta_{F,0}^D \\ 0 & 1 & -\theta_{D,0}^F & -\theta_{F,0}^F \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B(L) = \begin{bmatrix} \Phi_D^D(L) & \Phi_F^D(L) & \theta_D^{*D}(L) & \theta_F^{*D}(L) \\ \Phi_D^F(L) & \Phi_F^F(L) & \theta_D^{*F}(L) & \theta_F^{*F}(L) \\ \Pi_{q,D}^D(L) & \Pi_{q,F}^D(L) & \Pi_{x,D}^D(L) & \Pi_{x,F}^D(L) \\ \Pi_{q,D}^F(L) & \Pi_{q,F}^F(L) & \Pi_{x,D}^F(L) & \Pi_{x,F}^F(L) \end{bmatrix},$$

with $\theta_k^{*i}(L) = (\theta_k^i(L) - \theta_{k,0}^i)L^{-1}$, k and $i = \{D, F\}$.

In the theoretical framework proposed in section 2 there were two differentiated sources of information: private information and public disclosures. Trade-unrelated shocks were made equal to public disclosures and trade-related shocks were made equal to private information. We must be cautious, however, when adopting the residuals of model [16] as a proxy for such innovations. It is a widespread practice in microstructure research to link the innovations in the trading process with new information (e.g., Hasbrouck, 1988; Huang and Stoll, 1997). As indicated by Madhavan et al. (1997) and Kempf and Korn (1999), the revision of beliefs depends on the unexpected order flow because, by definition, the predictable order flow conveys no additional information. In model [16], we decompose the order flow in its expected and unexpected components under the assumption that the expected component only (linearly) depends on the past quote and trade history of both D and F markets –the same assumption made in the theoretical framework. Once we have identified the trade-related shocks, the trade-unrelated shocks are characterized as the trade-unrelated part of the quote revision innovation, as in Hasbrouck (1991a, 2002). These shocks are trade-unrelated because of the causality assumption in [16]. Therefore, we could identify all private information with (w_t^D, w_t^F) and all public information with $(\varepsilon_t^D, \varepsilon_t^F)$. In practice, however, this dichotomy is not so clear. On the one hand, in an order driven market quote revision innovations may reflect private information if informed traders submit limit orders instead of market orders. The usual claim is that informed traders prefer to submit market orders since their information is short lived and immediate execution is guaranteed (e.g., Angel, 1992, and Glosten, 1994). Nonetheless, Kaniel and Liu (2001) show that informed traders would prefer to submit limit orders when their information is long-lived or their valuation is close to the current market quotes. If the informed trader were patient, she would be more reluctant to submit large market orders since by doing so she would signal that the stock is mispriced. If she intends to execute an aggressive market order when her valuation is close to the current quotes, she would bear price risk. In a hybrid market, the quote revision innovations may reflect private information if either the market makers or the limit order traders possess superior information. Thus, Harris and Panchapagesan (2003) show that the main information advantage of the NYSE specialist is her privileged knowledge of the limit order book. On the other hand,

the assumption that trade innovations do not contain public information may be formalized as the requirement that public information is not useful in predicting the trade innovation. Under the conjecture that the relevant public information set is the history of quotes and trades, (w_t^D, w_t^F) in [16] should be unpredictable given ϕ_{t-1} . In practice, however, there are several reasons why this assumption may be violated. First, the relevant public information set might include other decision variables not explicitly modeled in [16]. Indeed, $(\varepsilon_t^D, \varepsilon_t^F)$ and (w_t^D, w_t^F) may indeed proxy for decision variables in the price discovery process or deterministic components that we have omitted. Second, Hasbrouck (1991a) claims that public information may affect posterior trading decisions when market imperfections or constraints impede quote revisions to fully reflect the public information, like price smoothing rules, verbal quotes in the crowd, stale limit orders etc. Indeed, limit orders in the crowd that are never disseminated off the floor could be considered as private rather than public information.

These remarks evidence that the line that separates public and private information is thinner in empirical treatments than in theoretical ones. Therefore, some information effects could be erroneously attributed to public or private sources. Additionally, since D and F may have different microstructures, the relative transparency also plays a role. For the less transparent market, the thin line between public and private information will be even more imprecise. As a result, model [16] should be more precise in separating trade-inferred information and information gathered from other sources than in separating public and private information. Since our main concern is to determine how much of a change in the efficient price is related to the trading activity of each market, model [16] is a correct tool. Nonetheless, we must keep in mind that the estimated innovations will always be contingent on the linear autoregressive structure imposed in model [16].⁴

⁴ There are other issues to mention. First, section 2 predicts $E(\varepsilon_t^D \varepsilon_t^F) \neq 0$ due to noisy public disclosures. Therefore, the empirical model is a system of seemingly unrelated equations that could be efficiently estimated by SURE (Zellner, 1962). In empirical applications, the correlation between the ε_t^i terms is expected to increase with the time interval used to define [16]. Consequently, if the time interval is short enough, the benefits of using SURE are mitigated. Second, Chiang and Lin (1999) show that the existence of a minimum price variation and the bid-ask bounce may induce a bias in the estimation of a VEC model using high-frequency transaction price data. However, the authors also show that by using quote midpoints this bias is considerably reduced.

4. The information content of trades

Hasbrouck (2002) confronts the information share approach in Hasbrouck (1995) with the permanent-transitory approach in Harris et al. (2002). He shows that in the case of a two markets model with private and public information, similar to the one presented in this paper, the information share approach is more reliable. The bound generated by the information share approach contains (up to estimation error) the true value. This cannot be said for the permanent-transitory approach. Therefore, we slightly modify the information share approach in Hasbrouck (1995) to differentiate between the information share that corresponds to each market's trade-related and trade-unrelated information.

Hasbrouck (1991b) defines the information content of the trading process as the variance of the expected impact of a trade innovation on the efficient price. In our case,

$$Var(E[\Delta m_t | \Delta x_t^i - E(\Delta x_t^i | \phi_{t-1}^i)]). \quad [17]$$

from [15], equation [17] is equivalent to,

$$I^i = Var(E[\Delta m_t | w_t^i]). \quad [18]$$

under the assumption that w_t , w_t^F and w_t^D are mutually uncorrelated, equation [18] is a proper absolute measure of the portion of the price discovery attributable to the trading activity of market i . The relative information content of the market i 's trading process is,

$$SIS^j = \frac{Var(E[\Delta m_t | w_t^j])}{Var(\Delta m_t)}. \quad [19]$$

Similar measures can be defined for the trade-unrelated shocks.

Every VEC model has an associated common trend model representation implied by the co-integration relationships. The vector moving average (VMA) representation of [16] is,

$$\Delta y_t = \Psi(L)\xi_t, \quad [20]$$

with $\Psi(L)$ is a lag polynomial. Consider the first two equations in [20], $\Delta q_t = \psi(L)\xi_t$, where $\Delta q_t' = (\Delta q_t^D \ \Delta q_t^F)$ and $\psi(L)$ represent the first two rows in $\Psi(L)$. By recursive

substitution and using that $\psi(L) = \psi(1) + (1-L)\psi^*(L)$, with $\psi^*(L) = (\psi(L) - \psi(1))(1-L)^{-1}$,

$$q_t = \psi(1) \sum_{\tau=1}^t \xi_\tau + \psi^*(L) \xi_t. \quad [21]$$

The first term on the RHS of [21] is the long-run (permanent) component, common to both quotes because of the theoretical co-integration relationship between quotes. The second term on the RHS of [21] is a zero-mean weakly stationary (transitory) component. Co-integration entails that $\delta' \Psi(1) = 0$, where δ is any co-integration vector. Under the theoretical assumption that the difference between the market quotes is stationary, $\delta' = (1 \ -1 \ 0 \ 0)$, we have that $\psi_1(1) = \psi_2(1) = \psi_{(1 \times 4)}$, with $\psi_k(1)$ representing the k -th row in $\Psi(1)$. Intuitively, the existence of a common long-run component implies that the long-run impact of a new shock on either the foreign market or the domestic market should have the same permanent impact on both quotes. It follows that $\psi \xi_t$ measures the impact of a shock on the information efficient price. Let $Var(\xi_t) = \Omega_{(4 \times 4)}$. Then, the long-run variance would be given by $Var(\Delta m_t) = \psi \Omega \psi'$. Our aim is to identify the part of this long-run variance that is explained by each market's information. Given the hypothesis of no correlation between the innovations in the trading activity (w_t^D, w_t^F) and the common informative shocks $(\varepsilon_t^D, \varepsilon_t^F)$, a correct measure of [19] would be,

$$SIS^i = \frac{\psi_i^2 \sigma_{w^i}^2}{\psi \Omega \psi'}, \quad [22]$$

where $Var(E[\Delta m_t | w_t^i]) = \psi_i^2 \sigma_{w^i}^2$ and ψ_i is the i -th component of the row vector ψ .

If the innovations in $\xi_t' = (\varepsilon_t^D \ \varepsilon_t^F \ w_t^D \ w_t^F)$ are correlated, the covariance terms in Ω could be attributed to any shock. In this paper, we follow Hasbrouck (1995)'s suggestion of constructing upper and lower bounds for the information shares. Therefore, we orthogonalize the residual variance-covariance matrix using the Cholesky factorization and rotate the ordering of the variables to maximize and minimize the explanatory power of each particular shock. The Hasbrouck modeling framework is problematic where the contemporaneous correlation of shocks across markets is substantive. In that case, Huang (2002) and Booth et al. (2002) show wide gaps between the upper and lower bounds on the information shares. In our particular implementation,

a high frequency sampling may be required to reduce a substantive contemporaneous correlation of trade-related and/or trade-unrelated shocks across markets.

5. Data

In the next sections, we present an empirical application of our methodology using data on Spanish stocks listed in the NYSE as ADRs. We analyze the contribution of each market during the daily two-hour overlapping interval from 9:30 a.m. to 11:30 a.m. (New York Time).⁵ U.S. data are taken from the NYSE's TAQ 2000 database. All quote and trade registers previous to the opening quote of the NYSE are dropped. Trades not codified as "regular trades" have been discarded. Trades performed at the same price, and with the same time stamp are treated as just one trade. Quotes with bid-ask spreads lower than or equal to zero or quoted depth equal to zero have also been eliminated. Trades are classified as buyer or seller initiated trades using the Lee and Ready (1991) algorithm. Price and quote files are coupled using the so-called "five second rule". This rule assigns to each trade the first quote stamped at least five seconds before the trade itself.⁶ The Spanish data are supplied by *Sociedad de Bolsas* (SB), the organization that supervises the computerized Spanish Stock Exchange Interconnection System (SIBE). The SIBE is a pure order-driven market where the most liquid Spanish stocks trade and large trading volumes are handled efficiently and transparently. Since 17 January 2000, SIBE operates in continuous trading between 9:00 a.m. to 5:30 p.m. (Spanish Time). The Spanish database includes a trade and a quote file. The quote file contains all the updates of the five best quotes on the bid and offer sides of the

⁵ Before 17 January 2000 continuous trading was from 9:30 to 17:00 ST; therefore, the overlapping trading period was just one hour and a half. Additionally, the beginning of the daylight saving time in October for Spain and US coincides. However, the end of this daylight saving time is the first Sunday of April in the US and the last Sunday of March in Spain. Hence, during the last week of March, the markets overlap only during one hour.

⁶ Ellis et al. (2000) and Odders-White (2000) have compared alternative classification rules. Both papers find that the Lee and Ready's algorithm outperform other classification routines like the quote rule and the tick rule. However, both papers found that the algorithm misclassifies transactions executed inside the quotes. Ellis et al. propose to apply the tick-rule not only to the midpoint trades, as in Lee and Ready's algorithm, but also to all the trades inside the quotes. The authors show that, for Nasdaq-listed stocks, this procedure improves over extant classification rules. However, there is no test evidencing that this alternative method improves the classification for NYSE-listed stocks. We applied both algorithms and did not find remarkable differences. In any case, Odders-White observes that the biases introduced by the classification rules are more relevant for large and frequently traded stocks. Finally, Blume and Goldstein (1997) showed that the "five-seconds rule" could not be generalized to all sample periods and markets. However, Odders-White (op. cit) also shows that the "five-seconds rule" does not seem to explain much of the bias induced by the Lee and Ready's algorithm.

electronic limit order book time stamped at the nearest hundredth of a second. The trade file includes detailed information about all trades. Due to the real time dissemination of all the information, there are no lags between the reporting time of the updated quotes and the transactions that triggered them. Trades are classified as buyer or seller initiated depending on the side of the book they hit. We transform the SSE quotes into US\$ applying the corresponding intra-daily exchange rate series provided by Reuters. This series has a 1-minute resolution and contains the last exchange rate quoted each minute.

The sample consists of the five Spanish NYSE-listed ADRs in 2000, discarding foreign subsidiaries and participated firms, with NYSE ticker symbols BBV, ELE, REP, SCH and TEF. These stocks are permanently among the 35 most liquid stocks in the SSE. They embody the 66% of the euro value of the 2000 SSE trading, 5% of the dollar value of the NYSE trading of European stocks, and less than 1% of the dollar value of all NYSE trading. These stocks became NYSE-listed before 1990 and an important part of their 2000 revenues come from their business activity in America.

Informal conversations with NYSE specialists of non-US stocks suggest that the proper null hypothesis for our empirical exercise should be that the NYSE is a pure satellite of the SSE. A shared opinion among all them is that the contribution of the NYSE to the price-discovery process of ADRs depends, to a large extent, on the degree of development of the domestic market. When the home market is an illiquid, infrequently traded, immature market, the NYSE tends to lead or at least remarkably contribute. This is not the case of the SSE. Because the trading process concentrates on a small set of stocks, the SSE provides particularly high standards of liquidity and activity for the cross-listed stocks, in the framework of a technologically advanced and transparent system. Specialists also agree that, in the case of a strong domestic market, especially when it trades in a different time frame than the US markets, the specialist usually becomes a passive rather than an active player in price formation. Since the specialist's liquidity provision for infrequently traded non-US stocks is expected to be crucial (Bacidore and Sofianos, 2000), we should also wait for NYSE quotes being hardly informative. Using Hasbrouck (1995) methodology, Grammig et al. (2004) and Hupperets et al. (2002) report empirical findings on two Dutch and two German NYSE-listed ADRs, respectively, not necessarily supporting this view. Grammig et al. report 1997-1998 NYSE information shares of 19.6% and 1% for SAP and DT. The NYSE volume shares for these stocks were 24% and 35%, respectively. Hupperets et al. report 1999 average NYSE information shares of 6.5% and 9% for AHO and KPN, with 2.2% and 2.7% NYSE volume shares respectively.

A different issue, however, is whether the SSE traders may glean valuable information from NYSE trades even when the NYSE quotes are uninformative. In this case, a quote-based analysis could erroneously characterize the NYSE as a satellite of the SSE. In Table I–Panel A we report the trading-shares of both markets during the overlapping interval. If the SSE and the NYSE trades were equally informative, we would expect the trade-related information shares to be close to these trading shares. Notice the remarkable differences this panel reports between TEF and the other Spanish stocks.

TABLE I. Activity during the overlapping interval

Panel A: It reports the percentage of the trading activity of the Spanish cross-listed stocks during the overlapping period (15:30-17:30 Spanish Time) that corresponds to the NYSE, both in terms of the volume traded (millions of shares) and trades (thousands). Panel B: It is indicative of the trading activity of the Spanish ADRs listed on the NYSE during the overlapping trading interval between the SSE and the NYSE. The overlapping interval has been divided into equally spaced time intervals from 1, 3 and 5 minutes. The panel B.1 reports the percentage of intervals with at least one new quote register in the TAQ database. The panel B.2 reports the percentage of intervals with at least one new trade register in the TAQ database. Panel C: It reports the empirical size distribution of the NYSE trades during the overlapping interval with respect to the 25%, 50%, 75%, 90%, 95% and 99% percentiles of the empirical size distribution of the SSE trades during the complete Spanish trading session (9:00-17:30 Spanish Time). Under the hypothesis of equality of both distributions, the expected values are 25%, 25%, 25%, 15%, 5%, 4% and 1%, respectively.

Panel A: NYSE trading shares					Panel B: Trading activity					
Stock	Volume		Trades		B.1. Quotes			B.2. Trades		
	Total	% NYSE	Total	% NYSE	1 min	3 min	5 min	1 min	3 min	5 min
BBV	519	1.35	208	2.07	18.71	40.03	54.28	11.73	27.92	39.97
ELE	283	3.08	166	4.17	29.69	60.07	77.45	18.84	45.6	63.68
REP	315	12.03	170	5.85	45.36	77.03	89	24.66	53.98	71.02
STD	710	2.8	247	3.48	31.06	61.33	78.02	21.05	47.23	64.18
TEF	1,795	23.23	651	5.93	74.19	96.28	98.52	61.82	90.23	96.3

Panel C: Size-Distribution of the NYSE trades during the overlapping*									
	Percentil	<=25	25-50	50-75	75-90	90-95	95-99	>99	Obs.
BBV	Buys	18.32	32.17	29.47	12.07	2.57	3.92	1.47	1632
	Sells	17.55	27.35	36.48	12.41	2.75	2.29	1.16	2837
ELE	Buys	27.95	26.91	25.05	11.87	3.41	3.51	1.31	3817
	Sells	0.00	47.03	33.91	10.65	3.82	3.40	1.19	3117
REP	Buys	13.91	13.43	27.66	20.58	8.95	9.16	6.31	4758
	Sells	16.42	7.36	33.91	18.38	10.06	8.00	5.88	5190
STD	Buys	28.06	27.34	25.14	10.78	3.17	3.26	2.24	4415
	Sells	13.56	35.72	30.20	14.09	2.07	2.40	1.95	4202
TEF	Buys	0.00	13.57	14.79	26.28	7.79	17.86	19.70	20732
	Sells	0.00	14.87	21.55	20.84	10.90	17.51	14.32	4505

*Trade-size percentiles of reference: complete SSE trading session (9:00-17:30 ST).

In Table I-Panel B the overlapping period (9:30-11:30 NYT) is divided into equally spaced time intervals of 1, 3 and 5 minutes. Panel B.1 (B.2) shows the percentage of intervals with at least one new quote (trade) register in the TAQ database. This table shows that BBV is an infrequently traded stock at the NYSE; TEF can be

considered as frequently traded; the remaining stocks are intermediate cases. Finally, Table I-Panel C reports the empirical distribution of the NYSE trade size during the overlapping interval with respect to the percentiles of the empirical distribution of the SSE trade size for the entire Spanish trading session. Observe that, for TEF and REP, the allocation of NYSE trades into the highest size-percentiles of the empirical distribution of the SSE trades is larger than expected under the hypothesis that both trade size distributions are equal. Therefore, NYSE trades for these stocks tend to be large in terms of the trade-size standards of the SSE. Additional analyses also reveal that the NYSE trade-size distribution does not change after the SSE closes. The NYSE specialists argue that US trades during the overlapping intervals usually hide large domestic investors trying to exploit profitable opportunities by trading simultaneously at both markets. Since the trades of these institutional traders may be particularly informative (Chakravarty, 2001), Spanish traders may obtain relevant information by monitoring the trading activity at the US market. Figure 1 reports the regular patterns in the intra-daily distribution of the standardized average number of trades per half-hour interval (Spanish Time). The standardized average share volume shows a similar pattern. Consistently with the institutional trading strategy suggested before, the SSE trading activity sharply increases as soon as the NYSE opens. During this period, SSE trades have the largest average size of all the Spanish session. After the SSE closes, the NYSE trading activity suddenly falls and only revives towards the end of the session.

We construct time series for five different clock-time periodicities (from 1 to 5 minutes). A change in quotes is computed as the difference between the logarithm of the quote midpoint at the end and at the beginning of each time interval. The trading process is represented either by the net share volume (NV_t) or by the net number of trades (NT_t) during each interval. We assume that indicators of net (signed) trading activity are more informative about the signed changes in market quotes than indicators of total volume traded. The NV_t is defined as the difference between the buyer-initiated volume and the seller-initiated volume at the interval t . The NT_t is defined as the difference between the number of buyer-initiated and seller-initiated trades. The trade indicator Δx_t^i in [16] is computed as,

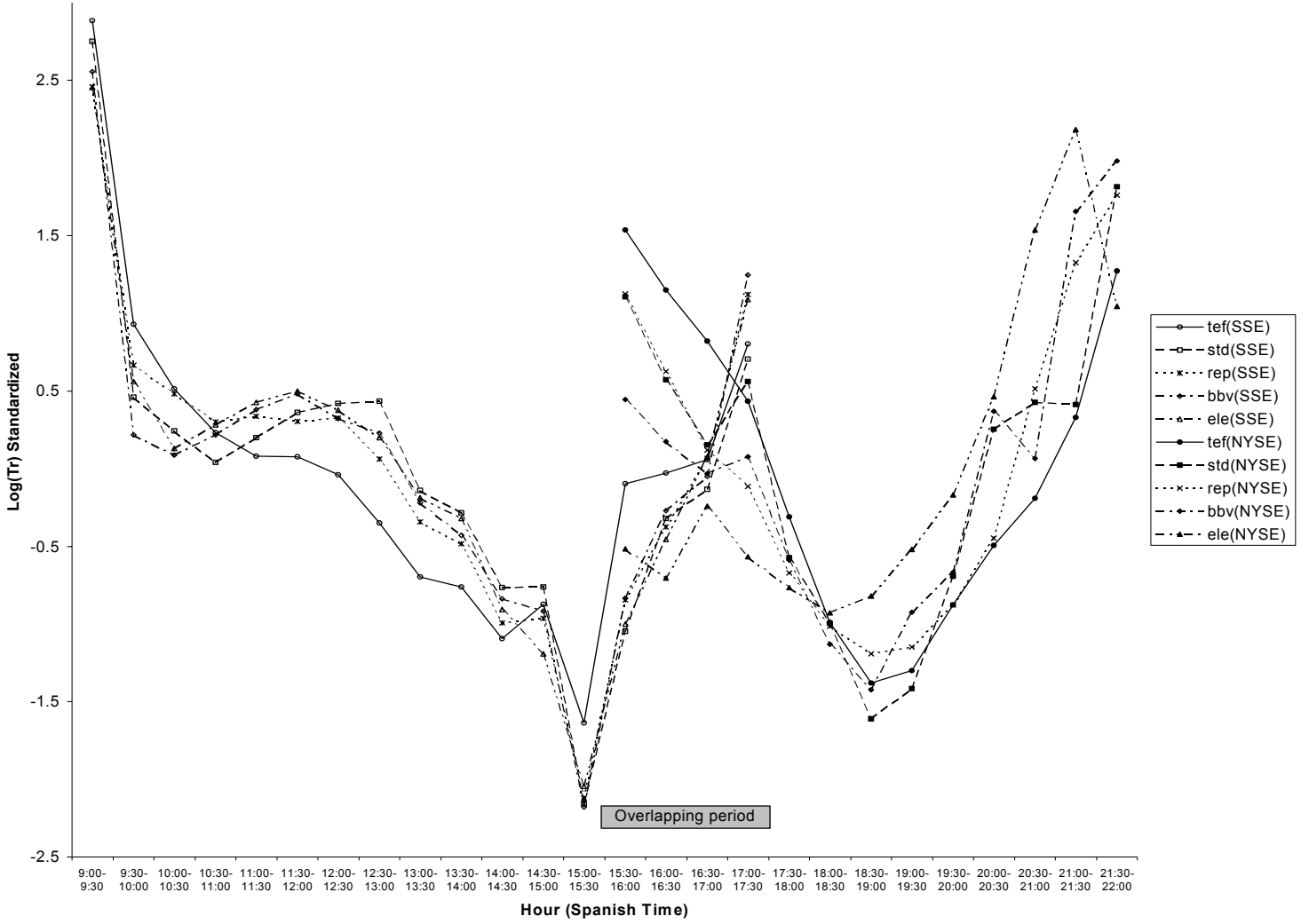
$$\Delta x_t^i = \text{sign}_t \left(\ln \left| \sum_{k=1}^z v_k \right| \right) \quad [23]$$

where $\sum_{k=1}^z v_k$ is either the NV_t or the NT_t , and $sign_t$ equals 1 if $\sum_{k=1}^z v_k > 0$ and -1 if

$\sum_{k=1}^z v_k < 0$. Therefore, x_t^i is computed as $\sum_{j=0}^{t-1} \Delta x_{t-j}^i$, starting with $x_0^i = 0$. All trading

measures have been rescaled: we divide x_t^i by 10000.

FIGURE 1. Intra-daily distribution of trading activity



Using the augmented Dickey-Fuller (1979), Phillips-Perron (1988) and the Kwiatkowski et al. (1992) tests, for all stocks and for all clock-time intervals, we accept a unit root for $y_t' = (q_t^D \ q_t^F \ x_t^D \ x_t^F)$. The optimal lag-length of the VEC model is determined using the Akaike Information Criteria (AIC) on a VAR model for y_t . We

use the Johansen's (1988, 1991, 1992) methodology to test how many co-integration vectors span the co-integration space. Tests for model specification are undertaken jointly with tests for co-integration rank. Regarding the model specification tests, we follow the so-called Pantula (1989) principle (see also Harris, 1995).⁷ In all cases, the model selected has an intercept in the co-integration space, and it includes neither drifts nor linear trends in the short run part of the model. The intercept accounts for differences in the units of measurement of quotes and trading proxies. When we use NV_t as the trading proxy, the trace statistic indicates $r=3$ co-integration vectors, for all stocks and for all time periodicities. Hence, in this model there is only one common trend determining the long-run steady state of all the variables in the system. When we take NT_t , the co-integration rank is $r=3$ for two stocks (TEF and BBV) and $r=2$ for the others. Therefore, some stocks need a second common trend, apparently associated with the trading process, to guarantee convergence to the steady state.

We carry out tests on α and β parameters in [16], using the Johansen methodology implemented in *Cats* for *Rats*. First, we check whether the last two rows of α are zeros, meaning that there is no loss of information from not modeling the determinants of Δx_t^D and Δx_t^F , and it is therefore valid to proceed with a two-equation system. We strongly reject this weak exogeneity hypothesis for all stocks except Δx_t^F for REP and STD in the NT_t model and only for some time-periodicities. Second, we test whether the $(1 \ -1 \ 0 \ 0)$ vector defines one of the structural economic relationships underlying the long-run model. We accept this as a non-binding restriction for all stocks except for STD, but generally with a non-zero intercept, that is $(1 \ -1 \ 0 \ 0 \ *)$. Possible explanations for this intercept are divergences between the market quotes due to market frictions (e.g., the tick size), inventory control by the NYSE specialist, or a less than accurate exchange rate time series. We also considered other theoretical co-integration vectors but we ended with mixed findings.⁸ In the following sections, we report the

⁷ The "Pantula principle" consists in choosing the models with lowest rank and then the model with less deterministic components. We consider two alternative specifications: (a) Intercept in the co-integration space and no deterministic components in the short-run part. (b) Intercept in the short-run part and no deterministic components in the co-integration space. This second case is equivalent to a model with intercepts in both the long run and the short run. It is assumed, however, that the intercept in the co-integration space is cancelled by the intercept in the short run. Other possible specifications have been discarded given the properties of our time series.

⁸ Co-integration between x_t^D and x_t^F alone is a non-binding constraint for 4 stocks in the NV_t model and 2 stocks in the NT_t model. Moreover, q_t^i and x_t^i are themselves co-integrated for 4 stocks in the NV_t model and 3 stocks in the NT_t model. The details of all the tests in this section are omitted because of the limited space, but they are available upon request. The estimation results for the models with an unrestricted and a fully identified β are also available.

results for the VEC model with only the $(1 \ -1 \ 0 \ 0 \ *)$ restriction in the long-run matrix β . Results with an unrestricted and a fully identified β are identical, which is particularly important for STD since we impose the $(1 \ -1 \ 0 \ 0 \ *)$ restriction.

6. Empirical analysis

6.1. Estimation results

This section summarizes the estimation of the VEC model [16] for the five Spanish cross-listed stocks. No overnight returns have been considered and no lags reached back to the previous day. When, for example, the optimal number of lags is four, the dependent vector begins with the fifth observation each overlapping interval. Table II reports the estimated model for TEF with the 1-minute time series when NT_t proxies for trading activity. The estimated model with NV_t is very similar. The lag-length is 3. There are 3 co-integration vectors and the first one is $(1 \ -1 \ 0 \ 0 \ *)$. We also report the residual correlation matrix and the Breusch and Pagan (1980) chi-square test for independence. The model is estimated by SURE using the FGLS algorithm (e.g., Green, 1997, pg. 511-513). The estimation results are consistent across stocks and the main findings derived are independent of the clock-time interval and the trading proxy used.

The identified error correction term (ECT_1) in the two quote equations is statistically significant for all stocks and specifications, and the sign of the coefficient is the expected one. If there is a movement in either the NYSE or the SSE away from the long-run equilibrium in a given period, a proportion of the disequilibrium is corrected the next period. Thus, if $q_{t-1}^S - q_{t-1}^{NY} > 0$ the next period Δq_t^S will decrease and Δq_t^{NY} will increase, rectifying (at least partially) the deviation between both markets. As the SSE quotes also respond to deviations from the NYSE quotes, this result suggests that the price discovery process may be not completely driven by the Spanish market. Consistently, when $q_{t-1}^S - q_{t-1}^{NY} > 0$ then Δx_t^S decreases (larger sell pressure) and Δx_t^{NY} increases (larger buy pressure), which forces quotes to move towards convergence. The magnitude of the coefficient associated to the ECT_1 of the Δq_t^S equation (α_t^S) is always smaller in absolute terms than the coefficient associated to the ECT of the Δq_t^{NY} equation (α_t^{NY}). Statistical tests performed confirm that $|\alpha_t^{NY}| > |\alpha_t^S|$ cannot be rejected

at the 1% level. Therefore, the reaction of the NYSE to the price differentials is faster and larger than that of the SSE.

TABLE II. Estimation of the VEC model

This table reports the estimated coefficients of the VEC model [16] for TEF. The overlapping trading interval between the SSE and the NYSE is divided in 1-minute intervals. For the definition of the variables, see section 5 in the paper. The lag length has been determined using the AIC information criterion. We report the coefficients of the VEC model that uses the transformation [23] of the net number of trades as the proxy for trading activity. The net number of trades is the difference between the number of buyer-initiated trades and the number of seller-initiated trades executed during the corresponding time interval. The first co-integration vector (ECT_1) is restricted to be (1 -1 0 0). We also report the residual correlation matrix with the Breusch and Pagan (1980) chi-square test for independence and the normalized co-integration vectors.

Obs.: 26668	Δq^S	Δq^{NY}	Δx^S	Δx^{NY}
$ECT_1 (t-1)$	-0.0167*	0.0968*	-0.00049*	0.00045*
$ECT_2 (t-1)$	0.00019*	0.0009*	-0.00007*	-6.03E-06*
$ECT_3 (t-1)$	0.00047*	-0.0001	-0.00005*	2.20E-06
$\Delta q^S (t-1)$	-0.1477*	0.1221*	0.00451*	0.00074*
$\Delta q^S (t-2)$	-0.0613*	0.1090*	-0.0012	0.00044*
$\Delta q^S (t-3)$	-0.0102*	0.0718*	0.0008*	0.00024*
$\Delta q^{NY} (t-1)$	-0.0002	-0.3146*	0.0003	-0.0019*
$\Delta q^{NY} (t-2)$	0.0028	-0.1392*	0.0001	-0.00089*
$\Delta q^{NY} (t-3)$	0.0042	-0.0566*	-0.0007	-0.00031*
$\Delta x^S (t)$	3.3233*	0.5643*		
$\Delta x^S (t-1)$	0.13031*	0.9552*	0.2669*	0.0079*
$\Delta x^S (t-2)$	-0.4282*	0.2088*	0.1060*	0.0041*
$\Delta x^S (t-3)$	0.0004	-0.0123*	0.0003	-0.0001
$\Delta x^{NY} (t)$	0.8787*	6.4607*		
$\Delta x^{NY} (t-1)$	0.3885*	2.974*	0.0958*	0.0534*
$\Delta x^{NY} (t-2)$	0.2705	0.8544*	0.0005	0.0234*
$\Delta x^{NY} (t-3)$	-0.0136	0.2306**	0.0087	-0.0010
R^2	0.1732	0.203	0.1446	0.0328

*, **, *** Significant at the 1%, 5% and 10% level respectively

Correlation matrix of residuals				
	\mathcal{E}^S	\mathcal{E}^{NY}	w^S	w^{NY}
\mathcal{E}^S	1			
\mathcal{E}^{NY}	0.0463	1		
w^S	0	0	1	
w^{NY}	0	0	0.0297	1

Breusch-Pagan test: $\chi^2(6) = 80.793$, P-value = 0.0000

Co-integration Vectors					
	q^S	q^{NY}	x^S	x^{NY}	$Int.$
ECT_1	1	-1	0	0	0.0051
ECT_2	1	0.9651	0.3214	-4.265	-6.873
ECT_3	1	0.9684	0.1705	-3.0934	-6.2294

Examining the coefficients of the lagged values of Δq_t^S and Δq_t^{NY} on the quote equations we observe, on the one hand, a significant negative autocorrelation in quotes. This could reflect the correction process that starts after an uninformative shock beats

the quotes. On the other hand, there is a significant positive effect of the lagged Δq_t^S values on the Δq_t^{NY} equation, reflecting co-movements of both quotes led by the SSE. The opposite relationship is also true for TEF for almost all time intervals: the lagged values of Δq_t^{NY} also affect positively to Δq_t^S . However, statistical tests confirm that $|\sum \phi_{S,j}^{NY}| > |\sum \phi_{NY,j}^S|$, that is, Δq_t^{NY} is more sensitive to changes in the lagged values of Δq_t^S than the other way around. The trading activity has a strong positive effect on quotes, independently of the proxy used. Table II evidences that an increase in NT_t either at the SSE or at the NYSE produces an upward adjustment of both quotes. This finding indicates that the trading activity at the NYSE provides some information to both markets, even when the trading activity at the SSE has been taken into account. For the other stocks in the sample, the significance of the effect of the NYSE trading activity on Δq_t^S depends on the time interval considered.

The estimated coefficients for the trade equations report both an expected and an unexpected finding. On the one hand, there is a strong positive autocorrelation in Δx_t^i independently of the proxy used (e.g., Hasbrouck, 1991a). Additionally, we find evidence of clusters of signed trading activity between the SSE and the NYSE, generally led by the Spanish market: positive lagged values of Δx_t^S are followed by positive values in Δx_t^{NY} . This result suggests trading transmission between markets. For the most frequently traded stocks at the NYSE (TEF and REP) we also found evidence of trading transmission from the NYSE to the SSE. On the other hand, positive lagged values of Δq_t^S increase both Δx_t^{NY} and Δx_t^S . That is, after an upward revision in the SSE quotes, both markets experiment a larger pressure to buy. However, and this is the unexpected finding, for the NYSE we obtain the opposite effect. Lagged positive values of Δq_t^{NY} , increase the NYSE pressure to sell (decreases x_t^{NY}). This relationship is not observed for the SSE trading activity. Hence, our explanation is that this result captures the marginal effect of inventory control by the NYSE specialist. Periods of intense demand ($x_t^{NY} > 0$) are linked to increases in the value of the stock ($\Delta q_t^{NY} > 0$), as we have observed before. During these periods, the NYSE specialist will be forced to provide liquidity in order to maintain stable market conditions. This is especially true for infrequently traded stocks (Madhavan and Sofianos, 1998) and non-US stocks (Bacidore and Sofianos, 2000). As a consequence, the specialist could be forced to hold an undesired negative inventory position in the cross-listed stock. During the next period, the specialist will try to motivate traders to introduce market orders to sell in

order to restate their preferred inventory position. This is also consistent with the negative autocorrelation of Δq_t^{NY} discussed before.

Previous results suggest that, as expected, the SSE leads the price discovery process of the Spanish cross-listed stocks. Unfortunately, whether the SSE trading activity accounts for an important part of the SSE contribution, whether the NYSE quotes are a factor in the price discovery process and whether the US trading activity is informative, are questions that cannot be answered through the estimation results alone. This is the aim of the next section.

6.2. *Information shares. General findings*

Table II reports the residual correlation matrix for the VEC model [16] estimated using TEF data and with a 1-minute resolution. Because of the causality assumption, trade-related and trade-unrelated noises are uncorrelated. Additionally, ε_t^S and ε_t^{NY} are slightly but significantly correlated. Of course, this correlation increases with time aggregation. Similarly, w_t^S and w_t^{NY} are also significantly correlated, suggesting common shocks in the trading process. This correlation also increases, but less, with time aggregation. Therefore, we expect the information share bounds described in section 4 to be tighter as we decrease time aggregation. This matrix also indicates that the scenario with noisy trade-unrelated shocks described in section 2 might be a proper characterization of the underlying price discovery process: the simultaneous trade-unrelated shocks are not perfectly correlated, suggesting that both markets may differ in their ability to judge noisy public disclosures.

We derive the VMA representation of the VEC model [16] directly from the estimated VEC model following Watson (1994). The changes in the US\$/€ exchange rate might cause some distortions in the computation of the information shares. However, Lieberman et al. (1999) conclude that the correlation between the changes in quotes and the changes in the exchange rates is irrelevant. Moreover, Grammig et al. (2004) find that the exchange rate is not a relevant determinant in price discovery. Therefore, we assume that any bias induced by shocks in the exchange rate is negligible. Table III-Panel A contains the lower and upper bounds of the information shares for the 1-minute resolution.

TABLE III. Information shares during the overlapping interval

This table contains the information shares (in %) for trade-related and trade-unrelated shocks originated at the NYSE and at the SSE. The information share is the proportion of variance in the efficient price that is attributable to a given innovation, either trade-related (w^S, w^{NY}) or trade-unrelated ($\varepsilon^S, \varepsilon^{NY}$). The table provides the lower bound and the upper bounds (in parenthesis) based on the estimation of the VEC model [16] (panel A) and the refinements discussed in section 6 (panels B to E). Results are reported for a 1-minute periodicity.

	BBV				ELE				REP				STD				TEF			
	ε^S	ε^{NY}	w^S	w^{NY}	ε^S	ε^{NY}	w^S	w^{NY}	ε^S	ε^{NY}	w^S	w^{NY}	ε^S	ε^{NY}	w^S	w^{NY}	ε^S	ε^{NY}	w^S	w^{NY}
Panel A: Model [16]																				
NV	82.1	1.0	16.5	0.1	85.0	2.3	12.2	0.2	78.4	4.1	16.7	0.4	83.1	0.9	15.9	0.1	68.8	6.3	20.3	1.9
	(82.3)	(1.2)	(16.6)	(0.2)	(85.3)	(2.6)	(12.2)	(0.2)	(78.7)	(4.4)	(16.8)	(0.4)	(83.1)	(0.9)	(15.9)	(0.1)	(71.1)	(8.6)	(20.7)	(2.3)
NT	77.2	0.0	22.7	0.0	78.9	2.2	18.6	0.0	69.6	4.9	24.9	0.2	76.2	0.7	23.0	0.0	61.9	6.6	28.4	0.9
	(77.2)	(0.0)	(22.7)	(0.1)	(79.2)	(2.4)	(18.6)	(0.0)	(69.9)	(5.2)	(25.0)	(0.2)	(76.2)	(0.7)	(23.0)	(0.0)	(63.9)	(8.6)	(28.7)	(1.2)
Panel B: Model [16] with no-trading dummies																				
NV	82.2	1.1	16.3	0.2	84.9	2.3	12.2	0.2	78.4	4.2	16.7	0.4	83.1	0.9	15.9	0.1	68.8	6.2	20.3	1.9
	(82.4)	(1.2)	(16.4)	(0.2)	(85.2)	(2.6)	(12.3)	(0.2)	(78.6)	(4.5)	(16.8)	(0.5)	(83.1)	(0.9)	(15.9)	(0.1)	(71.2)	(8.6)	(20.7)	(2.4)
NT	76.5	0.8	22.5	0.1	78.9	2.2	18.6	0.0	68.8	5.0	25.6	0.2	76.2	0.7	23.0	0.0	61.8	6.6	28.4	0.9
	(76.6)	(0.8)	(22.5)	(0.1)	(79.2)	(2.4)	(18.6)	(0.0)	(69.2)	(5.3)	(25.7)	(0.2)	(76.2)	(0.7)	(23.0)	(0.0)	(63.9)	(8.6)	(28.7)	(1.2)
Panel C: Model [16] with overnight information																				
NV	81.9	1.1	16.5	0.2	84.5	2.7	12.2	0.3	78.4	4.1	16.7	0.4	83.1	0.8	16.0	0.1	68.6	6.2	20.5	1.9
	(82.1)	(1.3)	(16.6)	(0.2)	(84.8)	(3.0)	(12.2)	(0.3)	(78.7)	(4.4)	(16.8)	(0.5)	(83.1)	(0.9)	(16.0)	(0.1)	(71.0)	(8.6)	(20.9)	(2.4)
NT	77.3	0.0	22.7	0.0	78.5	2.6	18.5	0.0	68.7	5.5	25.1	0.3	74.9	0.7	24.4	0.0	63.4	5.5	28.1	0.8
	(77.3)	(0.0)	(22.7)	(0.1)	(78.8)	(2.9)	(18.5)	(0.0)	(69.0)	(5.8)	(25.2)	(0.4)	(74.9)	(0.8)	(24.4)	(0.0)	(65.3)	(7.3)	(28.4)	(1.1)
Panel D: Model [16] with medium-sized and large-sized trades																				
NV	79.5	1.5	18.6	0.2	77.3	3.8	18.3	0.4	71.9	5.7	21.4	0.7	82.4	0.0	17.6	0.0	54.4	13.2	24.5	3.6
	(79.7)	(1.7)	(18.6)	(0.3)	(77.5)	(4.0)	(18.4)	(0.4)	(72.1)	(5.9)	(21.5)	(0.8)	(82.4)	(0.0)	(17.6)	(0.0)	(58.0)	(16.8)	(25.2)	(4.4)
NT	73.0	1.3	25.5	0.0	78.3	3.1	18.3	0.0	66.4	5.7	27.2	0.5	74.3	0.8	24.9	0.0	46.6	15.3	33.0	1.4
	(73.2)	(1.5)	(25.5)	(0.0)	(78.6)	(3.4)	(18.3)	(0.0)	(66.7)	(6.0)	(27.2)	(0.4)	(74.3)	(0.8)	(24.9)	(0.0)	(49.8)	(18.5)	(33.5)	(1.9)
Panel E: Model [16] with medium-sized trades (stealth trading)																				
NV	86.6	1.7	11.4	0.1	82.4	4.5	12.7	0.1	80.1	5.2	14.1	0.3	89.4	1.4	9.1	0.0	55.3	17.7	19.0	2.7
	(86.8)	(1.9)	(11.4)	(0.1)	(82.7)	(4.7)	(12.7)	(0.1)	(80.4)	(5.5)	(14.1)	(0.3)	(89.5)	(1.5)	(9.1)	(0.0)	(60.3)	(22.6)	(19.4)	(3.0)
NT	82.2	2.0	15.5	0.1	74.7	3.9	21.1	0.1	72.6	5.8	21.3	0.0	84.1	0.8	15.0	0.0	56.8	16.4	21.2	1.0
	(82.5)	(2.2)	(15.5)	(0.1)	(74.9)	(4.2)	(21.1)	(0.1)	(72.9)	(6.2)	(21.3)	(0.0)	(84.2)	(0.8)	(15.1)	(0.0)	(61.1)	(20.7)	(21.5)	(1.2)

The SSE is, as we expected, the leading market in the price discovery process of the five Spanish cross-listed stocks during the overlapping interval. The information shares for the SSE trade-unrelated informational shocks (ε_t^S) are between 62% and 85% depending on the stock and the trading measure. Table III also indicates that the contribution of the NYSE trade-unrelated shocks (ε_t^{NY}) is negligible for all stocks. Only TEF and REP, with a 6.3%-6.6% and 4.1%-4.9% of the variance of the efficient price explained by the NYSE trade-unrelated shocks, might generate some sort of doubt about its statistical significance. Even in this case, the contribution is marginal. For the remaining stocks, the ε_t^{NY} contribution approaches zero. In the context of the scenario proposed in section 2, these results suggest that the SSE generally does judgments about public disclosures that are more precise and disseminates this information quicker than the NYSE. However, for the most frequently traded Spanish stocks, the NYSE might

have some advantage over the SSE in processing some particular event, probably news that concern the American investments of these firms, their industries in general, or the US economy in particular.

The trade-related informational shares show a completely different picture. On the one hand, for the 1-minute resolution the SSE trade-related shocks (w_t^S) explain between 12.2% and 28.4% of the long-term variance of the Spanish cross-listed stocks, again depending on the stock and trading measure. The information share for the NYSE trade-related shocks (w_t^{NY}), however, is less than 1% in almost all the cases. We define the NYSE relative trade-related informational share (RTRIS) as the trade-related information share of the NYSE over the sum of all the information shares attributed to the trading activity, that is, $RTRIS^{NY} = IS(w_t^{NY}) / (IS(w_t^{NY}) + IS(w_t^S))$. If we compare this measure with the “trading shares” in Table I, the conclusion must be that the NYSE trading activity has no information content: 23% of TEF and 12.03% of REP traded volume takes place at the NYSE. However, the corresponding RTRIS are almost zero. Similar results are obtained with the number of trades.

6.3. *Refinements*

In this subsection, we discuss some refinements of the empirical model [16] that provide new insights into the price discovery process of the Spanish cross-listed stocks. Table III reports the information shares for some of these “improved” model specifications.⁹

Intra-daily patterns: Figure 1 evidenced deterministic patterns in trading activity during the overlapping trading interval between the NYSE and the SSE. Since trading patterns may signal regular information flow changes, the information shares may conditionally change over time. We divide the overlapping period in three intervals and repeat the preceding analysis for each of them. The first period (15:30-16:00 ST) encloses the opening of the NYSE and the third period (17:00-17:30 ST) comprises the close of the SSE. We find that the contribution of the NYSE is not uniform during the overlapping period. As the close of the SSE approaches, the NYSE clearly becomes a pure satellite. Indeed, for TEF the NYSE information shares are zero in almost all cases.

No-trading periods: The measures of net (signed) trading activity we use to represent the trading process do not differentiate between zeros due to no trading and

⁹ Tables with detailed estimation results are available upon request from the authors.

zeros due to an equilibrium between buyer and seller initiated trading. The information content of a without-trading period, however, may differ from that of a with-trading period, either balanced or not (e.g., Easley and O'Hara, 1992). In our sample, zeros in NV_t are mainly due to no-trading periods (95% SSE, 99% NYSE). Zeros in NT_t are due to balanced trading for the SSE (52%-89%) and no trading for the NYSE (80-98%). Of course, the proportion of zeros in either trade proxy is much larger for the NYSE than for the SSE. These differences might introduce some sort of bias in the computation of the information shares. To deal with this problem, we add no-trading dummy variables to model [16]. The NYSE (SSE) dummy equals 1 whenever there is a contemporaneous no-trading zero in the NYSE (SSE) trading proxy and 0 otherwise. Table III-Panel B provides the information shares. To sum up, we do not find any remarkable difference with the results reported before, suggesting that, in our sample, no-trading periods provide the same information than balanced periods.

Overnight information: In model [16], we are assuming that nothing significant to the model has happened in the 22 hours between the close of the SSE and the opening of the NYSE the next day. Namely, we are implicitly imposing that NYSE opening residuals do not serially depend on the information revealed during the last 22 hours. There are several reasons to think this could be the case. Prior the NYSE opens, the SSE has been trading during 6½ hours. This period may suffice to incorporate all the public information revealed during the market closure in New York or during the overnight period. Indeed, we would expect the initial NYSE shocks to reflect new information not revealed during the first 6½ hours of trading in Spain. The sudden drop in the SSE activity just before the NYSE opens (see Figure 1), suggests that Spanish traders are expecting new information to be revealed with the opening of the US market. In this case, the initial residuals of the NYSE continuous-trading session would be uncorrelated with public disclosures during the non-overlapping period. We empirically evaluate the impact of the overnight information and the 6½ hours of price discovery in Spain on the information shares by adding additional exogenous variables to model [16]. For each trading day d , we compute the return, net volume, and net number of trades in New York between 17:30($d-1$) (SSE closes) and 22:00($d-1$) (NYSE closes), and in Spain between 9:00(d) a.m. (SSE opens) and 15:30(d) (NYSE opens). We also obtain the overnight returns between 22:00($d-1$) and 9:00(d) a.m. Then, we construct seven time series indicators that equal the daily value of the corresponding variable for the first half hour of the overlapping interval and 0 otherwise. There is only one systematic empirical pattern across stocks: the SSE returns during the 6½ hours preceding the NYSE opening are positively correlated with the initial NYSE trade-unrelated shocks. Despite that, Table III-Panel C evidences that the information shares for this model specification do not remarkably differ from those reported in previous panels.

Alternative trading measures: Despite the fact that the main conclusions do not vary, Table III - Panels A to C evidence that the information shares may depend on the trading proxy used in model [16]. Indeed, microstructure research suggests alternative and, *a priori*, equally acceptable trading proxies. In several competitive and strategic asymmetric information models, the size of trades is positively related with the quality or precision of the information possessed by informed traders (e.g., Easley and O'Hara, 1987; Kim and Verrecchia, 1991; Holden and Subrahmanyam, 1992). In contrast, Barclay and Warner (1993) and Chakravarty (2001), provide evidence supporting the so-called stealth-trading hypothesis, that privately informed traders concentrate their trades in medium sizes to camouflage their positions so that most of a stock's cumulative price change will take place on medium-size trades. If every informed trader focuses on medium and large sizes, including small-sized trades in the NV_t and NT_t trading measures adds an unnecessary uninformative noise. What is more, when the stealth-trading hypothesis is true even large-sized trades could be discarded. Table III reports the information shares of model [16] when NV_t and NT_t are computed using all trades but the small-sized ones (Panel D) and just medium-sized trades (Panel E). All trades with size lower than or equal to (larger than) the 25% (75%) percentile of the empirical distribution of the trade size for each particular stock are classified as small-sized (larger-sized) trades. Panel D shows that, when small-sized orders are discarded, the NYSE trade-unrelated information share for TEF remarkably increases. When the largest trades are eliminated, the NYSE contribution augments and additional 2-4% in this stock. Therefore, under the assumption that the stealth-trading hypothesis applies, the NYSE cannot be considered as a pure satellite market of the SSE during the overlapping interval for the most frequently traded cross-listed stock. Nonetheless, this US contribution is absolutely trade-unrelated.

Liquidity shocks: As discussed earlier, trade-related and trade-unrelated shocks in model [16] may be affected by omitted endogenous decision variables. A solid candidate is liquidity. As far as liquidity varies over time, it affects trading costs, volatility and encourages strategic behavior by traders.¹⁰ In this paper, we use net depth to measure liquidity. Recently, Chakravarty et al. (2001) have shown that depths rather

¹⁰ Handa and Schwartz (1996), Parlour (1998) and Foucault (1999) propose models where the state of the limit order book influences the order flow composition. In these models an unbalanced limit order book reflects the market sentiment. Similarly, Huang and Stoll (1994) claim that the asymmetric depth reflects information asymmetries. Seppi (1997) and Kaniel and Liu (2001) theoretically show that, contrary to the usual claim, informed traders would prefer to submit limit orders rather than market orders under certain conditions. Empirically, Harris and Panchapagesan (2003), Madhavan and Panchapagesan (2000), Corwin and Lipson (2000) and Coppejans and Domowitz (2002), among others, provide evidence on the information content of the limit order book.

than spreads are first to impound new information that leads to new quote trends. Moreover, Coppejans et al. (2001) evidence that positive depth shocks on the offer (bid) side of the market tend to lower (raise) returns. Since both liquidity traders and informed traders may submit limit orders, the dynamics of the size quotes have an informative and an uninformative component. The market i 's depth process (d_t^i) can be modeled as a function of both public and private information. Similar to [11], $\Delta d_t^i = \pi^i \Delta m_t^* + \kappa_t^i$, with κ_t^i being the unpredictable component of liquidity given the common information set. Table IV shows the information shares for a six-equation VEC model with trade-related shocks and liquidity-related shocks as measured by net depth.¹¹ In summary, liquidity-shocks play a negligible role in the price-discovery process of the Spanish ADRs, independently of trading proxy used. In line with Coppejans and Domowitz (1999), our results suggest that liquidity-shocks are not informative once we control for trade-related information.

In summary, the NYSE is a pure satellite market of the SSE at least for four of the five Spanish cross-listed stocks. The information content of the trading activity in the US market is negligible for all stocks ($\lambda_F \approx 0$). Only for the most frequently traded stock and under the assumption that all the informed trades concentrate in medium sizes (stealth trading) or medium and large sizes, the NYSE trade-unrelated contribution is noticeable.

¹¹ Model [16] is expanded to incorporate the depth processes so that we consider a six-equation model. We maintain the causal relationship from trades to quotes and we add causality from depth to trades and quotes. Intuitively, depth is a function of the bids and offers, the result of submitting passive limit orders to the system. These orders naturally precede transactions. In addition, since limit orders may transmit information by their own, they may also cause quote adjustments even without trading activity. The depth proxy for each market is the net depth at the best quotes weighted by time. We apply the transformation [23] to get Δd_t^i . The series in levels d_t^i is computed by accumulation of the series in first differences. The unit root tests indicate that d_t^i (Δd_t^i) is a unit root (stationary) process for both markets and for all stocks. Details about model specification tests and estimation results are available upon request.

TABLE IV. Information shares with liquidity shocks

This table contains the estimated information shares (in %) for trade-related, liquidity-related and trade-and-liquidity-unrelated shocks originated at the NYSE and the SSE. The information share is the proportion of variance in the efficient price of the stock that is attributable to a given innovation, either trade-related (w^S , w^{NY}), liquidity-related (κ^S , κ^{NY}) or trade-and-liquidity-unrelated (ε^S , ε^{NY}). The table provides the lower bound and the upper bounds (in parenthesis) based on the estimation of the VEC model [16], expanded to incorporate net depth. Results are reported for the 1-minute clock time resolution.

	NV						NT					
	ε^S	ε^{NY}	w^S	w^{NY}	κ^S	κ^{NY}	ε^S	ε^{NY}	w^S	w^{NY}	κ^S	κ^{NY}
BBV	78.47 (78.70)	1.81 (2.04)	18.79 (18.87)	0.29 (0.37)	0.30 (0.30)	0.03 (0.03)	71.83 (72.08)	1.84 (2.09)	25.43 (25.45)	0.16 (0.18)	0.46 (0.46)	0.01 (0.01)
ELE	82.23 (82.50)	2.22 (2.49)	14.21 (14.22)	0.20 (0.21)	0.86 (0.86)	0.00 (0.00)	77.27 (77.51)	2.23 (2.47)	19.45 (19.44)	0.02 (0.02)	0.76 (0.76)	0.04 (0.04)
REP	78.47 (78.73)	3.87 (4.13)	15.90 (15.92)	0.18 (0.20)	1.29 (1.30)	0.00 (0.00)	70.10 (70.41)	3.90 (4.21)	24.29 (24.30)	0.03 (0.04)	1.37 (1.37)	0.00 (0.00)
STD	83.46 (83.48)	1.24 (1.27)	15.16 (15.16)	0.06 (0.07)	0.05 (0.05)	0.00 (0.00)	76.38 (76.45)	0.99 (1.06)	22.54 (22.54)	0.00 (0.00)	0.02 (0.02)	0.00 (0.00)
TEF	65.04 (67.29)	6.48 (8.73)	21.56 (22.06)	1.86 (2.37)	0.60 (0.59)	1.72 (1.72)	57.95 (59.98)	6.54 (8.57)	29.73 (30.13)	1.04 (1.44)	0.57 (0.57)	1.74 (1.74)

7. Conclusions

This paper studies the price discovery process of cross-listed securities during the period when the different markets simultaneously trade (their trading sessions overlap). Building on Hasbrouck (1995) information share approach, we have introduced a methodology that distinguishes two sources of information asymmetries between markets: trade-related and trade-unrelated information. By separating the quote innovations into trade-related and trade-unrelated shocks, we are able to decompose the global information share of each market into its trade-related and trade-unrelated components. By modeling the trading processes of the different markets, we also allow for incremental information contained in trade shocks, beyond that reflected in each market's quotes. This approach improves on previous empirical methodologies in several ways: it considers other pieces of information than the market quotes; it provides a more accurate characterization of a dominant-satellite relationship when it exists; it accounts for alternative sources of information provision and, therefore, information asymmetries; it allows to evaluate the information content of the trading process of each market; and it is flexible enough as to accommodate further extensions, like introducing liquidity shocks.

As an empirical application of this methodology, we studied the contribution of the NYSE in the price discovery process of the Spanish cross-listed stocks. This study

centered on the daily overlapping trading interval between both the NYSE and the SSE. Not surprisingly, our results indicate that the SSE leads the price discovery. The SSE trade-related shocks account for 12% to 30% of the long-run variance of the Spanish cross-listed stocks. NYSE trade-related shares, however, are undistinguishable from zero. Indeed, the NYSE global contribution is negligible. Only for the most frequently traded stock and under the assumption that only medium and large sized trades are informative, the NYSE trade-unrelated shocks account for 15-17% of the price-discovery process.

References

- Angel, J., 1992, Limit orders versus market orders, *Working Paper*, Georgetown University.
- Bacidore, Jeffrey M., and George Sofianos, 2000, NYSE specialist trading in non-US stocks, *NYSE Working Paper*, #00-05.
- Barclay, Michael J., and Jerold B. Warner, 1993, Stealth trading and volatility, *Journal of Financial Economics*, 34, 281-305.
- Blume, Marshall E., and Michael A. Goldstein, 1997, Quotes, order flow, and price discovery, *Journal of Finance*, 52, 221-244.
- Booth, G. Geoffrey, Ji-Chan Lin, Teppo Martikainen, and Yiurman Tse, 2002, Trading and pricing in upstairs and downstairs stock markets, *Review of Financial Studies*, 15, 4, 1111-1135.
- Breusch, T., and A. Pagan, 1980, The LM test and its implications to model specification in econometrics, *Review of Economic Studies*, 47, 239-254.
- Chakravarty, Sugato, 2001, Stealth-trading: Which traders' trades move stock prices?, *Journal of Financial Economics*, 61, 289-307.
- Chakravarty, Sugato, Frederick H. deB. Harris, and Robert A. Wood, 2001, Do bid-ask spreads or bid and ask depths convey information first?, *Working Paper*, Wake Forest University.
- Chan, Kalok C., Wai-Ming Fong, Bong-Chan Kho, and René Stulz, 1996, Information, trading and stock returns: lessons from dually listed securities, *Journal of Banking and Finance*, 1161-1187.
- Chiang, Kevin C.H., and Ji-Chai Lin, 1999, Inferring price discovery from frictional markets, *Working Paper*, E.J. Ourso College of Business Administration, Louisiana State University.
- Chowdhry, Bhagwan, and Vikram Nanda, 1991, Multimarket trading and market liquidity, *Review of Financial Studies*, 4, 3, 483-511.
- Coppejans, Mark, and Ian Domowitz, 1999, Screen information, trader activity, and bid-ask spreads in a limit order market, *Working Paper*, Duke University.
- Coppejans, Mark, and Ian Domowitz, 2002, An empirical analysis of trades, orders, and cancellations in a limit order market, *Working Paper*, Duke University.
- Coppejans, Mark, Ian Domowitz, and Ananth Madhavan, 2001, Liquidity in an automated auction, *Working Paper*, Duke University.
- Dickey, D.A., and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 427-431.

- Easley, David, and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics*, 19, 69-90.
- Easley, David, and Maureen O'Hara, 1992, Time and the process of security price adjustment, *Journal of Finance*, 47, 2, 577-605.
- Ellis, Katrina, Roni Michaely, and Maureen O'Hara, 2000, The accuracy of trade classification rules: evidence from Nasdaq, *Journal of Financial and Quantitative Analysis*, 35, 4, 529-551.
- Engle, Robert, and C. Granger, 1987, Co-integration and error correction: Representation, estimation and testing, *Econometrica*, 35, 251-276.
- Foucault, Thierry, 1999, Order flow composition and trading costs in a dynamic limit order market, *Journal of Financial Markets*, 2, 99-134.
- Glosten, L., 1994, Is the electronic open limit order book inevitable? *Journal of Finance*, 47, 577-606.
- Gonzalo, Jesus, and Clive Granger, 1995, Estimation of common long-memory components in co-integrated systems, *Journal of Business and Economic Statistics*, 13, 1-9.
- Grammig, Joachim, Michael Melvin, and Christian Schlag, 2004, Price discovery in international equity trading, *Journal of Empirical Finance*, forthcoming.
- Green, William H., 1997, Econometric Analysis, Prentice-Hall.
- Handa, Puneet, and Robert A. Schwartz, 1996, Limit order trading, *Journal of Finance*, 51, 1835-1861.
- Harris, Milton, and Arthur Raviv, 1993, Differences of opinion make a horse race, *Review of Financial Studies*, 6, 473-506.
- Harris, Frederick H. deB, Thomas H. McInish, Gary L. Shoesmith, and Robert A. Wood, 1995, Co-integration, error correction, and price discovery on internationally linked security markets, *Journal of Financial and Quantitative Analysis*, 30, 4, 563-579.
- Harris, Frederick H. deB, Thomas H. McInish, and Robert A. Wood, 2002, Security price adjustment across exchanges: an investigation of common factor components for Dow stocks, *Journal of Financial Markets*, 5, 277-308.
- Harris, Lawrence E., and Venkatesh Panchapagesan, 2003, The information-content of the limit order book: evidence from NYSE specialist actions, *Working Paper*, Marshall School of Business, University of Southern California.
- Harris, Richard, 1995, Co-integration analysis in econometric modeling, Prentice Hall.
- Hasbrouck, Joel, 1988, Trades, quotes and information, *Journal of Financial Economics*, 22, 229-252.
- Hasbrouck, Joel, 1991a, Measuring the information content of stock trades, *Journal of Finance*, 46, 179-207.

- Hasbrouck, Joel, 1991b, The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies*, 4, 571-595.
- Hasbrouck, Joel, 1995, One security, many markets: determining the contributions to price discovery, *Journal of Finance*, 50, 4, 1175-1199.
- Hasbrouck, Joel, 1996, Modeling market microstructure time series, G.S. Maddala and C.R. Rao, eds.: *Handbook of Statistics*, vol. 14, Statistical Methods in Finance (Elsevier, North-Holland, Amsterdam).
- Hasbrouck, Joel, 2002, Stalking the “efficient price” in market microstructure specifications: An overview, *Journal of Financial Markets*, 5, 3, 329-339.
- Hasbrouck, Joel, George Sofianos, and Deborah Sosebee, 1993, New York Stock Exchange systems and trading procedures, NYSE *Working Paper* #93-01.
- Holden, Craig W., and Avanidhar Subrahmanyam, 1992, Long-lived private information and imperfect competition, *Journal of Finance*, 47, 247-270.
- Huang, Roger D., 2002, The quality of ECN and Nasdaq market maker quotes, *Journal of Finance*, 57, 3, 1285-1319.
- Huang, Roger D., and Hans R. Stoll, 1994, Market microstructure and stock return predictions, *Review of Financial Studies*, 7, 1, 179-213.
- 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.
- Huuperets, Erik, and Bert Menkveld, 2002, Intra-daily analysis of market integration: Dutch blue chips traded in Amsterdam and New York, *Journal of Financial Markets*, 5, 57-82.
- Johansen, Soren, 1988, Statistical analysis of co-integration vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.
- Johansen, Soren, 1991, Estimation and hypothesis testing of co-integration vectors in Gaussian vector autoregressive models, *Econometrica*, 59, 1551-1580.
- Johansen, Soren, 1992, Determination of the co-integration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics*, 54, 383-397.
- Kaniel, Ron, and Hong Liu, 2001, So what order do informed traders use?, *Working Paper*, University of Texas at Austin.
- Kempf, Alexander, and Olaf Korn, 1999, Market depth and order size, *Journal of Financial Markets*, 2, 29-48.
- Kim, Oliver, and Robert E. Verrecchia, 1991, Market reaction to anticipated announcements, *Journal of Financial Economics*, 273-309.
- Kim, Oliver, and Robert E. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics*, 17, 41-67.

- Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, 54, 159-178.
- Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica*, 53, 1315-1335.
- Lee, Charles M., and Mark J. Ready, 1991, Inferring trade direction from intra-daily data, *Journal of Finance*, 46, 733-746.
- Lieberman, Offer, Uri Ben-Zion, and Schmuël Hauser, 1999, A characterization of the price behavior of international dual stocks: an error correction approach, *Journal of International Money and Finance*, 18, 2, 289-304.
- Madhavan, Ananth, and Venkatesh Pachapagesan, 2000, Price discovery in auction markets: a look inside the black box, *Review of Financial Studies*, 13, 3, 627-658.
- 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.
- Madhavan, Ananth, and George Sofianos, 1998, An empirical analysis of NYSE specialist trading, *Journal of Financial Economics*, 48, 189-210.
- Odders-White, Elizabeth R., 2000, On the occurrence and consequences of inaccurate trade classification, *Journal of Financial Markets*, 3, 3, 259-286.
- Pantula, S., G., 1989, Testing for unit roots in time series data, *Econometric Theory*, 5, 256-271.
- Parlour, Christine A., 1998, Price dynamics and limit order markets, *Review of Financial Studies*, 11, 4, 789-816.
- Pascual, Roberto, Alvaro Escribano, and Mikel Tapia, 2003, "Adverse Selection Costs, Trading Activity and Liquidity in the NYSE: An Empirical Analysis", *Journal of Banking and Finance*, forthcoming.
- Phillips, P.C.B., and P. Perron, 1988, Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.
- Seppi, Duane, 1997, Liquidity provision with limit orders and a strategic specialist, *Review of Financial Studies*, 10, 1, 103-150.
- Subrahmanyam, Advanidhar, 1997, Multi-market trading and the informativeness of stock trades: an empirical intra-daily analysis, *Journal of Economics and Business*, 49, 515-531.
- Tse, Yiuman, 2000, Further examination of price discovery on the NYSE and regional exchanges, *Journal of Financial Research*, 23, 3, 331-351.
- Watson, Mark W., 1994, Vector autoregressions and co-integration, in Engle R.F. and D.L. McFadden (Eds.) *Handbook of Econometrics*, Vol. IV, Elsevier Science, North-Holland, Amsterdam, 2844-1915.

Zellner, Arnold, 1962, An efficient method of estimating Seemingly Unrelated Regressions and tests for aggregation bias, *Journal of the American Statistical Association*, 57, 348-368.