Journal of Financial Markets 5 (2002) 259–276

Journal of FINANCIAL MARKETS

www.elsevier.com/locate/econbase

Some desiderata for the measurement of price discovery across markets ☆

Bruce N. Lehmann*

Graduate School of International Relations and Pacific Studies, University of California at San Diego, 9500 Gilman drive, La Jolla, CA 92093-0519, USA

1. Introduction

One of the central functions of secondary markets is price discovery: the efficient and timely incorporation of the information implicit in investor trading into market prices. Trade in identical or related assets often takes place in multiple trading venues and the resulting fragmentation of order flow raises concerns about the efficacy of price discovery and the fairness of trading mechanisms. Accordingly, it is of more than passing economic interest to measure the impact of market fragmentation on the price discovery process.

Hasbrouck (1995) introduced one method for measuring a given market's contribution to price discovery. Prices for the same asset in different markets should tend to converge in the long run but might deviate from one another in the short run due to trading frictions. Under additional, reasonably weak assumptions, such time series are cointegrated and modern time series methods can be used to measure the variance of innovations to the long run price and to decompose it into components, termed information shares, due to each market. However, when price innovations across markets are correlated, the cross-market terms that cannot be allocated cleanly. This technology has spawned a modest-sized cottage industry.

In a sequel to the Harris et al. (1995) error correction model, the lead paper in this volume, Harris et al. (2002a), follows a suggestion from Clive Granger and imports a closely related time series approach from the cointegration literature: the Gonzalo and Granger (1995) permanent/transitory decomposition of cointegrated time series.

☆ I have benefited much from conversations with Graham Elliott, from correspondence with Frank de Jong, who caught a blunder in Section 3, and from the helpful suggestions of Geoff Booth, Rick Harris, Joel Hasbrouck, and Yiuman Tse. I am extremely grateful to Frank de Jong for refereeing and/or reading each paper in this volume. The extraordinary volume of professional service he provided in the process merits sainthood in my view. I am also grateful to Joel Hasbrouck for helping to assemble the papers. The usual disclaimer applies.

*Tel.: +1-858-534-0945; fax: +1-858-534-3939.

E-mail address: blehmann@ucsd.edu (B.N. Lehmann).

PII: S1386-4181(02)00025-3

This procedure identifies a permanent component with two properties: it is a linear combination of contemporaneous data and it is not Granger-caused in the long run by stationary linear combinations of the data. Harris et al. (2002a) associate the Gonzalo-Granger permanent component with long run prices and present evidence on price discovery across three markets based on this approach.

The other three papers in this volume were written to address the reasonableness and interpretation of this identification procedure in price discovery research. Independently of one another, Baillie et al. (2002) and de Jong (2002), the latter as an insight provided in a referee report (!), show precisely what the Gonzalo-Granger permanent component does measure along with its relation to the Hasbrouck information shares and construct different examples illustrating when the two approaches give similar and different answers. Baillie et al. (2002) provides empirical evidence on the differences and similarities between the two as well. Hasbrouck (2002) suggests that the Gonzalo-Granger permanent component does not measure efficient prices because the first desideratum makes little sense in microstructure applications and constructs three examples to illustrate this point. Finally, Harris et al. (2002b) respond to these disparate critiques, focusing their attention on Hasbrouck (2002).

At this point, a typical editorial introduction either ends or provides a few additional details about the papers that follow. At the risk of abusing editorial privilege, I will go on for quite a bit longer for two reasons. First, Hasbrouck (1996, 2002) and elsewhere has consistently emphasized that the economically reasonable a priori restrictions that arise in microstructure applications of cointegration technology simplify matters considerably. Second, my reading of the four papers while preparing this introduction suggested that some synthesis would be desirable.

Accordingly, what follows is my attempt to summarize and synthesize the methods used in this line of research along with the corresponding implications of the papers in this volume for its future. The next section provides a self-contained introduction to the analytics of cointegration and error correction that arises naturally in the econometrics of price discovery. The only tool of modern time series analysis that is required is the ability to evaluate a lag polynomial of the form $C(L)$ at unity—that is, by setting $L = 1$ to obtain $C(1)$—as opposed to the mathematics of singular matrix polynomials typically used to prove the Granger representation theorem, a simplification arising from plausible a priori restrictions. Section 3 reviews the examples and simulations from the papers in this volume and complements them by filling in some analytical gaps, shedding additional light on the circumstances in which the Hasbrouck information shares and Gonzalo-Granger decompositions "work". The penultimate section records some thoughts on the uses of trades and quotes in price discovery research and is followed by a brief conclusion.

2. Methods

The starting point of the analysis of price discovery is the specification of precisely what it is that is being discovered. In broad discussions of market efficiency, this is an almost theological question, what with the epistemological difficulties associated with the identification of fundamental values and abstract notions of heterogeneous private information. These difficulties are side-stepped in the market microstructure literature by defining fundamental value as the price of an asset in the distant future and private information as some knowledge or insight as to where prices are headed. The phrase "distant future" should be thought of as distant in trading time—an interval during which there may be numerous transactions and quote changes—but sufficiently close in calendar time and modest in ex ante risk so that the economically appropriate discount rate is negligible. On this specification, fundamental value is a permanent component of prices and deviations of current prices from fundamental value may be related to private information that is revealed in the price discovery process.

The second element is the characterization of the frictions that prevent prices from continuously reflecting the price that would prevail in the absence of microstructural noise. Since interest is focused here on the general question of price discovery, it makes sense to permit a comparatively unrestrictive stochastic process for deviations of actual prices from friction-free ones. A minimal requirement is that microstructural noise should not affect the price of an asset in the indefinite future and thus should be a transient or temporary component of prices.

Accordingly, suppose a security trades on two venues, cleverly labeled one and two, at potentially different prices p_{1t} and p_{2t}. Since the securities are identical and the microstructural noise is transient, the definition of fundamental value implies that there is a common underlying implicit efficient price m_t satisfying

$$m_t = \lim_{\tau \to \infty} E[p_{1t+\tau}|I_t] = \lim_{\tau \to \infty} E[p_{2t+\tau}|I_t],$$

$$u_t = \Delta m_t = m_t - m_{t-1},$$

$$E[u_t|I_{t-1}] = E[\lim_{\tau \to \infty}\{E[p_{1t+\tau}|I_t] - E[p_{1t+\tau}|I_{t-1}]\}|I_{t-1}] = 0, \tag{1}$$

where I_{t-1} denotes public information available at time $t-1$ that, at minimum, includes past prices. At time t, the prices in the two markets (potentially) differ because of microstructural noise:

$$s_{it} = p_{it} - m_t, \quad E[\lim_{\tau \to \infty} s_{it+\tau}|I_t] = 0, \tag{2}$$

where the second equation reflects the transiene of s_t.

It is statistically convenient to assume that the efficient price follows a random walk as opposed to a martingale and that the microstructural noise is covariance stationary as opposed to being merely transient. On these assumptions, prices satisfy the structural model:[1]

$$p_t = \iota m_t + s_t, \quad p_t = \begin{pmatrix} p_{1t} \\ p_{2t} \end{pmatrix}, \quad s_t = \begin{pmatrix} s_{1t} \\ s_{2t} \end{pmatrix}, \quad \iota = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

[1] All of the results that follow extend to the N asset case with no substantive modifications. See the other papers in this volume as well as footnote 7.

$$m_t = m_{t-1} + u_t, \quad E[u_t] = E[u_t u_s] = 0 \qquad \forall s \neq t, \quad E[u_t^2] = \sigma_u^2,$$

$$s_t = \Upsilon(L)v_t, \quad \Upsilon(L) = \sum_{j=0}^{\infty} \Upsilon_j L^j, \quad \Upsilon_0 = I,$$

$$E[v_t] = 0, \quad E[v_t v_s'] = 0 \qquad \forall s \neq t, \quad E[v_t v_t'] = \Sigma_v, \tag{3}$$

where the last line follows from the assumption that s_t is covariance stationary. Note that there is no presumption that the efficient price innovation u_t is uncorrelated with future values of the microstructural noise s_t. Put differently, u_t and v_t can be correlated.[2]

Now consider zero net share acquisition trading strategies across the two venues. Such strategies take the form $\varphi(1 \quad -1)$ with $\varphi \neq 0$ and have costs given by

$$\varphi(1 \quad -1)p_t = \varphi s_{1t} - \varphi s_{2t}. \tag{4}$$

Since the difference of two stationary processes is stationary, the cost of a series of such trades constitutes a zero mean covariance stationary stochastic process. That $\varphi(1 \quad -1)p_t$ is covariance stationary is intuitively obvious: zero net share acquisition across trading venues represents exposure only to the microstructural noise that causes prices to differ and not to the efficient price m_t since the long position taken in market 1 (assuming $\varphi > 0$) is offset by the short sale in market 2. In the parlance of modern time series analysis, $\varphi(1 \quad -1)$ is the cointegrating vector that is unique up to the scale factor φ, which is normalized to one—that is, the purchase of one share in the first market and the sale of a share in the second market—in what follows.

From (3) above, the first difference of p_t has the moving average representation

$$\Delta p_t = \Psi(L)\varepsilon_t = \imath u_t + \Delta s_t = \imath u_t + (1-L)\Upsilon(L)v_t, \quad \Psi(L) = \sum_{j=0}^{\infty} \Psi_j L^j,$$

$$\Psi_0 = I \tag{5}$$

so that Δs_t is a covariance stationary but noninvertible moving average. That is, the long run impact of v_t on s_t is $\Upsilon(1)v_t$ while that of Δv_t on Δs_t is $\Upsilon(1)\Delta v_t = (1-L)\Upsilon(L)|_{L=1} v_t = 0$. Since

$$\Psi(L) \equiv \Psi(1) + (1-L)\Psi^*(L), \quad \Psi_i^* = -\sum_{j=i+1}^{\infty} \Psi_j, \tag{6}$$

p_t and Δp_t can always be rewritten as

$$p_t = \Psi(1)\sum_{s=0}^{t} \varepsilon_s + \Psi^*(L)\varepsilon_t,$$

$$\Delta p_t = \Psi(1)\varepsilon_t + \Psi^*(L)\Delta\varepsilon_t = \Psi(1)\varepsilon_t + (1-L)\Psi^*(L)\varepsilon_t, \tag{7}$$

[2] Nonzero means such as those arising from bid/ask spreads are ignored in what follows but the analysis can be easily adapted to accommodate them.

where $\Psi^*(L)\varepsilon_t$ is covariance stationary and its first difference $(1-L)\Psi^*(L)\varepsilon_t$ is a stationary, noninvertible moving average. Since $z_t = (1 \; -1)p_t$ is stationary, it must omit the stochastic trend $\Psi(1)\sum_{s=0}^{t}\varepsilon_s$, implying that $(1 \quad -1)\Psi(1) = 0$ and thus $\Psi(1) = \iota(\psi_1\psi_2)$ (i.e., $\Psi(1)$ has identical rows) which is intuitively obvious since the two prices share the same implicit efficient price.

This representation underlies the Hasbrouck (1995) information shares approach. The long run impact of ε_t, u_t, and v_t on Δp_t may be found by evaluating both $\Psi(L)$and $(1-L)\Upsilon(L)$ in (5) at $L=1$, yielding

$$\Psi(1)\varepsilon_t = \iota u_t \tag{8}$$

and the resulting perfect correlation arising from the relation $\psi'\varepsilon_t = u_t$ implies

$$E[\iota\iota' u_t^2] = E[\Psi(1)\varepsilon_t\varepsilon_t'\Psi(1)'] = E[\iota\psi'\varepsilon_t\varepsilon_t'\psi\iota'] \Rightarrow \sigma_u^2 = \psi'\Sigma_\varepsilon\psi \tag{9}$$

irrespective of the parameters of the structural model (3) and, in particular, of the model for s_t.[3] In other words, market price movements are dominated by those of the efficient price in the long run since differencing a stationary process eliminates its long run impact.

The Hasbrouck information shares involve decomposing $\psi'\Sigma_\varepsilon\psi$ into components attributed to price innovations in the two markets. This attribution is unique when said price innovations are uncorrelated, in which case the decomposition is given by:

$$1 = \frac{\psi_1^2}{\psi'\Sigma_\varepsilon\psi}\sigma_{\varepsilon_1}^2 + \frac{\psi_2^2}{\psi'\Sigma_\varepsilon\psi}\sigma_{\varepsilon_2}^2, \tag{10}$$

where uncorrelated price innovations ε_t implies correlated structural residuals v_t because the correlation induced by the common factor u_t must be 'undone' by the correlation among u_t, v_{1t}, and v_{2t}. When the reduced form residuals are correlated, the decomposition is

$$1 = \frac{\psi_1^2}{\psi'\Sigma_\varepsilon\psi}\sigma_{\varepsilon_1}^2 + 2\frac{\psi_1\psi_2}{\psi'\Sigma_\varepsilon\psi}\sigma_{\varepsilon_1\varepsilon_2} + \frac{\psi_2^2}{\psi'\Sigma_\varepsilon\psi}\sigma_{\varepsilon_2}^2 \tag{11}$$

and there is a range of possible attributions corresponding to different allocations of the covariance term to each market. In the jargon of econometrics, the contribution of each innovation to that of the efficient price is not identified, a point discussed in all four papers.

Now consider nonzero net share acquisition strategies across the two venues. In contrast to the zero net share purchase case, nonzero net cross-market acquisition yields costs with both random walk and stationary components. That is, the simultaneous acquisition of w_1 shares in the first market and w_2 in the second market with $w_1 + w_2 \neq 0$ has cost

$$p_{wt} = w_1 p_{1t} + w_2 p_{2t} = (w_1 + w_2)m_t + w_1 s_{1t} + w_2 s_{2t} = m_t + w s_{1t} + (1-w)s_{2t} \tag{12}$$

[3] More formally, $S_x(\omega) = B(\mathrm{e}^{-i\omega})\Sigma_e B(\mathrm{e}^{-i\omega})'$ is the spectral density matrix at frequency ω of $x_t = B(L)e_t$ where e_t is white noise and $S_x(1) = B(1)\Sigma_e B(1)'$at frequency zero. Hence, $S_{\Delta p}(\omega)$ is given both by $S_{\Delta p}(\omega) = \Psi(\mathrm{e}^{-i\omega})\Sigma_\varepsilon\Psi(\mathrm{e}^{-i\omega})'$ and by $S_{\Delta p}(\omega) = \iota\iota'\sigma_u^2 + (1-\mathrm{e}^{-i\omega})[\Upsilon(\mathrm{e}^{-i\omega})\mathrm{Cov}(v_t, u_t)\iota' + \iota\,\mathrm{Cov}(u_t, v_t')\Upsilon(\mathrm{e}^{-i\omega})'] + (1-\mathrm{e}^{-2i\omega})\Upsilon(\mathrm{e}^{-i\omega})\Sigma_v\Upsilon(\mathrm{e}^{-i\omega})'$ so that $S_{\Delta p}(0) = \Psi(1)\Sigma_\varepsilon\Psi(1)' = \iota\iota'\sigma_u^2$at frequency zero.

after normalizing the total share purchase across markets to one, which is natural since interest here centers on the permanent component m_t.

Zero and nonzero net share acquisitions across trading venues can be used to reparameterize the model as

$$\begin{aligned} p_t &= \eta_w p_{wt} + \delta_w z_t = \eta m_t + \left[\begin{pmatrix} \eta_{1w} \\ \eta_{2w} \end{pmatrix} \begin{pmatrix} w & 1-w \end{pmatrix} + \begin{pmatrix} \delta_{1w} \\ \delta_{2w} \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} \right] s_t \\ &\Rightarrow \eta = \iota \Rightarrow I = \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} w & 1-w \end{pmatrix} + \begin{pmatrix} \delta_{1w} \\ \delta_{2w} \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} \right] \\ &\Rightarrow \delta_{1w} = 1 - w, \quad \delta_{2w} = -w, \\ &\Rightarrow w\delta_{1w} + (1-w)\delta_{2w} = 0 \end{aligned} \tag{13}$$

where the last two lines follow by equating (3) and (13). Hence, we can *always* rewrite (3) as

$$p_t = \iota p_{wt} + \begin{pmatrix} 1-w \\ -w \end{pmatrix} z_t \tag{14}$$

for any choice of w. This reparameterization involves no economics, only arithmetic.

Harris et al. (2002a) suggest that one such reformulation does endow p_{wt} with economic content: the Gonzalo and Granger (1995) common factor representation. In words, w is chosen so that z_t does not Granger cause p_{wt} in the long run, the definition of long run Granger causality being that the cumulative impact of a variable y_t on a cointegrated series x_t is nonzero or

$$\lim_{\tau \to \infty} \frac{\partial E[x_{t+\tau} | I_t]}{\partial y_t} \neq 0. \tag{15}$$

Since m_t is a random walk, any estimate of it should, at minimum, satisfy such a condition.

The implications of long run Granger non-causality of p_{wt} by z_t involve the vector error correction representation of p_t. Consider the autoregressive representation in levels:[4]

$$\Phi(L)p_t = \varepsilon_t, \quad \Phi(L) = \sum_{j=0}^{K} \Phi_j L^j, \quad \Phi_0 = I \tag{16}$$

and the arithmetically equivalent representation in first differences

$$\begin{aligned} &\Delta p_t = \Phi(1)p_{t-1} + A(L)\Delta p_{t-1} + \varepsilon_t, \quad A(L) = \sum_{j=1}^{K-1} A_j L^{j-1}, \\ &A_i = \sum_{j=i+1}^{K} \Phi_i, \qquad i = 1, \ldots, K-1. \end{aligned} \tag{17}$$

[4] The assumption that the order of the autoregression is finite is for convenience; most of what follows does not require finite lag lengths.

Since Δp_t is stationary in first differences, $A(L)\Delta p_{t-1}$ represents a stationary component of price changes. Tedious but straightforward algebra reveals that $\Phi(1) = \gamma\begin{pmatrix}1 & -1\end{pmatrix}$ via

$$\begin{aligned}
\Delta p_t = \Psi(L)\varepsilon_t &\Rightarrow (1-L)p_t = \Psi(L)\varepsilon_t \Rightarrow (1-L)\Phi(L)p_t = \Phi(L)\Psi(L)\varepsilon_t \\
&\Rightarrow (1-L)\varepsilon_t = \Phi(L)\Psi(L)\varepsilon_t \Rightarrow (1-L)I = \Phi(L)\Psi(L) \\
&\Rightarrow 0 = \Phi(1)\Psi(1) = \Phi(1)\iota\psi' \\
&\Rightarrow \exists\gamma\ s.t.\ \Phi(1) = -\gamma(\,1 \quad -1\,)
\end{aligned} \tag{18}$$

since $(\,1 \quad -1\,)\Psi(1) = (\,1 \quad -1\,)\iota\psi' = 0$ so that the vector error correction representation

$$\Delta p_t = \gamma z_{t-1} + A(L)\Delta p_{t-1} + \varepsilon_t, \quad E[\varepsilon_t] = E[\varepsilon_t\varepsilon_s] = 0 \qquad \forall t,\ s, \quad E[\varepsilon_t\varepsilon_t'] = \Sigma_\varepsilon \tag{19}$$

follows from the first difference representations (5) and (17).

These relations can be used to identify the portfolio weight w_f that defines the Gonzalo-Granger common factor $f_t = w_f p_{1t} + (1-w_f)p_{2t}$. Transform the vector error correction model for Δp_t into one for Δf_t and Δz_t via

$$\begin{pmatrix}\Delta f_t \\ \Delta z_t\end{pmatrix} = \Pi\Delta p_t = \gamma^\pi z_{t-1} + A^\pi(L)\begin{pmatrix}\Delta f_{t-1} \\ \Delta z_{t-1}\end{pmatrix} + \varepsilon_t^\pi,$$

$$\Pi = \begin{pmatrix} w_f & 1-w_f \\ 1 & -1\end{pmatrix}, \quad \gamma^\pi = \begin{pmatrix}\gamma_f \\ \gamma_1 - \gamma_2\end{pmatrix}, \quad \gamma_f = w_f\gamma_1 + (1-w_f)\gamma_2,$$

$$A^\pi(L) = \Pi A(L)\Pi^{-1} \tag{20}$$

which, to eliminate the redundant Δz_t term, can be rewritten as one for Δf_t and z_t through

$$\begin{aligned}
\begin{pmatrix}\Delta f_t \\ z_t\end{pmatrix} &= \begin{pmatrix}0 \\ 1\end{pmatrix} z_{t-1} + \gamma^\pi z_{t-1} + A^\pi(L)\begin{pmatrix}1 & 0 \\ 0 & 1-L\end{pmatrix}\begin{pmatrix}\Delta f_{t-1} \\ z_{t-1}\end{pmatrix} + \varepsilon_t^\pi \\
&= \begin{pmatrix}0 \\ 1\end{pmatrix} z_{t-1} + \gamma^\pi z_{t-1} + \begin{pmatrix}A_{11}^\pi(L) & A_{12}^\pi(L)(1-L) \\ A_{21}^\pi(L) & A_{22}^\pi(L)(1-L)\end{pmatrix}\begin{pmatrix}\Delta f_{t-1} \\ z_{t-1}\end{pmatrix} + \varepsilon_t^\pi
\end{aligned} \tag{21}$$

yielding the autoregressive representation

$$\begin{pmatrix} B_{11}^{\pi}(L) & B_{12}^{\pi}(L) \\ B_{21}^{\pi}(L) & B_{22}^{\pi}(L) \end{pmatrix} \begin{pmatrix} \Delta f_t \\ z_t \end{pmatrix} = \begin{pmatrix} 1 - A_{11}^{\pi}(L)L & -[\gamma_f + A_{12}^{\pi}(L)(1-L)]L \\ -A_{21}^{\pi}(L)L & [(\gamma_2 - \gamma_1) - A_{22}^{\pi}(L)(1-L)]L \end{pmatrix}$$
$$\begin{pmatrix} \Delta f_t \\ z_t \end{pmatrix} = \varepsilon_t^{\pi} \tag{22}$$

and the corresponding moving average representation

$$\begin{pmatrix} \Delta f_t \\ z_t \end{pmatrix} = \begin{pmatrix} B_{11}^{\pi}(L) & B_{12}^{\pi}(L) \\ B_{21}^{\pi}(L) & B_{22}^{\pi}(L) \end{pmatrix}^{-1}$$
$$\varepsilon_t^{\pi} = \begin{pmatrix} 1 - A_{11}^{\pi}(L)L & -[\gamma_f + A_{12}^{\pi}(L)(1-L)]L \\ -A_{21}^{\pi}(L)L & [(\gamma_2 - \gamma_1) - A_{22}^{\pi}(L)(1-L)]L \end{pmatrix}^{-1} \varepsilon_t^{\pi}. \tag{23}$$

The long run impact of z_t on f_t is zero if that of $\varepsilon_{zt}^{\pi} = \begin{pmatrix} 1 & -1 \end{pmatrix} \varepsilon_t$ is zero which, in turn, requires the 1, 2 element of $B^{\pi}(1)^{-1}$ to be zero since $B^{\pi}(1)^{-1}$ is triangular if $B^{\pi}(1)$ is. Accordingly, the error correction term $\gamma_f = 0$ in the Gonzalo-Granger representation and so $w_f = -(\gamma_1 - \gamma_2)^{-1}\gamma_2 = -\delta_2$ and the corresponding common factor estimate is given by $f_t = (\gamma_1 - \gamma_2)^{-1}[\gamma_1 p_{2t} - \gamma_2 p_{1t}]$.

Now it is clearly desirable to select a common factor estimate that expunges the error correction term. However, it is not desirable for an estimate to depend only on contemporaneous prices because Δf_t is the sum of u_t, the innovation in the efficient price, and the stationary term $(\gamma_1 - \gamma_2)^{-1}(-\gamma_2 \quad \gamma_1)\, A(L)\Delta p_{t-1}$. Put differently, $u_t = \Delta m_t$ Granger causes price changes and is not Granger caused by them implying, in this setting, that u_t cannot be predicted from lagged price changes or z_{t-1}, a condition much stronger than the absence of long-run Granger causality alone. Purging a common factor estimate of all microstructural noise is desirable but impossible if the covariance matrix Σ_v is nonsingular. However, differenced common factor estimates can be rendered unpredictable if they are linear combinations of current and lagged prices.[5] This point is made forcefully in Sections 2 and 3 of Hasbrouck (2002).

That said, the portfolio weights themselves do measure an interesting quantity. Recall that the impact matrix $\Psi(1) = \iota(\psi_1 \psi_2)$ reflects the long run response of future prices $p_{t+\tau}$ to innovations in the current price ε_t where τ is sufficiently far in the future so that the effects of s_t can reasonably be assumed to be negligible; that is, $\lim_{\tau \to \infty} \partial p_{t+\tau} / \partial \varepsilon_t' = \iota \psi'$.[6] Manipulations similar to those performed in (18) above

[5] Weak-form efficiency cannot justify inclusion of only p_t on the hypothesis that transient microstructure effects do not generate exploitable profit opportunities; if anything, weak form efficiency suggests treating $(\gamma_1 - \gamma_2)^{-1}(-\gamma_2\gamma_1)\varepsilon_t$ as the new information that is impounded in market prices.

[6] In microstructure applications, measured stationary components typically die out quite rapidly so T will typically be relatively small in practice. This intuition must be false since returns are more predictable at low frequencies, low in this case being measured in weeks, months, and years. However, such time-varying expected returns should have a negligible impact in the high frequency data that arise in microstructure applications.

serve to identify ψ because

$$\begin{aligned}\Phi(L)p_t = \varepsilon_t &\Rightarrow \Psi(L)\Phi(L)p_t = \Psi(L)\varepsilon_t = \Delta p_t = (1-L)p_t\\ &\Rightarrow (1-L)I = \Psi(L)\Phi(L)\\ &\Rightarrow \Psi(1)\Phi(1) = 0 \Rightarrow \left(\iota\psi'\right)\left[\gamma\begin{pmatrix}1 & -1\end{pmatrix}\right] = 0 \Rightarrow \psi'\gamma = 0\\ &\Rightarrow \psi = \gamma_\perp\end{aligned} \tag{24}$$

where $\gamma_\perp$ satisfies $\gamma'\gamma_\perp = 0$.[7]

It is the implications of this last relation for the application to price discovery that may be found in de Jong (2002) and Baillie et al. (2002), who derived them independently. While the Gonzalo-Granger common factor estimates may not be of economic interest, the portfolio weights underlying them are because

$$\begin{pmatrix} w_f \\ 1-w_f \end{pmatrix} = \frac{1}{\gamma_1-\gamma_2}\begin{pmatrix} -\gamma_2 \\ \gamma_1 \end{pmatrix}, \tag{25}$$

which is clearly orthogonal to γ. Put differently, the random walk component of the Gonzalo-Granger common factor is the usual Stock-Watson common trend as noted by both de Jong and Baillie et al. as well as in Proposition 5 of Gonzalo and Granger (1995). Note that this link suggests that a potentially better estimate of the unobserved efficient price can be obtained from the Stock-Watson common trend representation (see also (7) above)

$$p_t = \Psi(1)\sum_{s=0}^{t}\varepsilon_s + \Psi^*(L)\varepsilon_t = \iota\tau_t + \Psi^*(L)\varepsilon_t,$$

$$\tau_t = \tau_{t-1} + \gamma'_\perp\varepsilon_t \Rightarrow f_t = \tau_t + \gamma'_\perp\Psi^*(L)\varepsilon_t, \tag{26}$$

[7] These results extend trivially to the N asset case as can be seen by replacing $(1\ \ -1)$ with the (nonunique) cointegrating matrix α throughout the preceding equations. For example, two such matrices and their error correction terms are given by

$$\alpha_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \quad z_t^{\alpha_1} = \alpha_1 p_t$$

and

$$\alpha_2 = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \equiv T_{\alpha_2}\alpha_1; z_t^{\alpha_2} = \alpha_2 p_t = T_{\alpha_2}\alpha_1 p_t = T_{\alpha_2} z_t^{\alpha_1}$$

in the three asset case. The corresponding error correction representations are

$$\Delta p_t = \Gamma^{\alpha_1} z_{t-1}^{\alpha_1} + A(L)\Delta p_{t-1} + \varepsilon_t = (\Gamma^{\alpha_1} T_{\alpha_2}^{-1})T_{\alpha_2} z_{t-1}^{\alpha_1} + A(L)\Delta p_{t-1} + \varepsilon_t = \Gamma^{\alpha_2} z_{t-1}^{\alpha_2} + A(L)\Delta p_{t-1} + \varepsilon_t$$

and thus

$$\gamma'_\perp\Gamma^{\alpha_1} = 0 \Rightarrow \gamma'_\perp\Gamma^{\alpha_2} = \gamma'_\perp\Gamma^{\alpha_1}T_{\alpha_2}^{-1} = 0.$$

where ε_t estimates can be obtained from the estimated error correction model.[8] As is obvious from (8), this representation underlies Hasbrouck's information shares as well.

Accordingly, the difference between the Hasbrouck information shares approach and the Gonzalo and Granger identification scheme employed in Harris et al. lies not in the latter's use of the common factor f_t but rather in their differential usage of normalized $\gamma_\perp$. Harris et al. (2002a) equate it with price adjustment dynamics: $\gamma_\perp$ measures the long-run impact of current price innovations on future prices—that is, $\lim_{\tau\to\infty} \partial p_{t+\tau}/\partial\varepsilon'_t$—and not its contribution to price discovery in any direct sense.[9] The information shares obtained from $\gamma'_\perp \Sigma_\varepsilon^{1/2}$ for any $\Sigma_\varepsilon = \Sigma_\varepsilon^{1/2}(\Sigma_\varepsilon^{1/2})'$ represent one way to link these multipliers to price discovery and there may well be others. Of course, $\gamma_\perp$ can always be used to answer questions about long run impact multipliers.

3. Models

That there remains much confusion over just what normalized estimates of $\gamma_\perp$ imply for price discovery is amply demonstrated by the dueling simulations and examples that populate the papers in this volume. This muddled state is unsurprising because the error correction model is a reduced form and the role of each market in price discovery depends on the parameters of the structural model. Accordingly, it is a simple matter to construct examples in which $\gamma_\perp$ is an arbitrarily good or bad measure of price discovery.

The papers in this volume contain two sorts of examples, corresponding to simple first order autoregressive and moving average models for price changes, respectively. Baillie et al. and de Jong assume that the autoregressive polynomial $A(L)$ is zero in the error correction model (19). Hence, price changes satisfy

$$\Delta p_t = \gamma z_{t-1} + \varepsilon_t \tag{27}$$

[8] The question of whether the common trend τ_t is superior to the Gonzalo-Granger common factor f_t in finite samples is an empirical one that depends on both the size of the population stationary component $f_t - \tau_t$ and the precision of the model estimates. With respect to the latter point, a natural a priori supposition is that the speed with which the stationary microstructural noise vanishes coupled with the huge number of observations available in microstructure applications suggests that the parameters underlying short-run security price dynamics will typically be estimated with considerable precision, mitigating concern about the measurement error in the ε_t estimate.

[9] Harris et al. observe that large sample inference for $\gamma_\perp$ estimates is straightforward using the eigenvector computations discussed in Gonzalo and Granger (1995). In models like this one in which the cointegrating vector is specified a priori, it seems to me to be more sensible to obtain $\hat{\gamma}$ from OLS estimation of the VECM and base inference on the δ-method applied to $\hat{\gamma}_\perp = f(\hat{\gamma})$, making large sample inference straightforward for information shares for any allocation of the covariance term to each market. This observation is easily extended to the N asset case because a basis for the space of cointegrating vectors can be specified a priori; see the discussion in footnote 7.

implying, after some manipulation, that the stationary component is a first order autoregression:

$$s_t = \left[I + \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} \right] s_{t-1} + v_t$$
$$= \begin{pmatrix} 1+\gamma_1 & -\gamma_1 \\ \gamma_2 & 1-\gamma_2 \end{pmatrix} s_{t-1} + v_t, \quad E[u_t v_t] = 0. \tag{28}$$

The substitution of (28) into (3) reveals that this model is a typical, reduced form error correction or partial adjustment mechanism via

$$p_t - \iota m_t = \begin{pmatrix} 1+\gamma_1 & -\gamma_1 \\ \gamma_2 & 1-\gamma_2 \end{pmatrix} (p_{t-1} - \iota m_{t-1}) + v_t. \tag{29}$$

Both de Jong and Baillie et al. contemplate the conceptual experiment of varying the correlation structure of the microstructural noise while holding the common factor weights $\gamma_\perp$ fixed; see, in particular, Eqs. (13) and (14) of de Jong and Table 1 in Baillie et al. They show that the common factor portfolio weights implicit in $\gamma_\perp$ are compatible with any information share from zero to one for different specifications of Σ_v in this model with the two approaches yielding the same answer when $\Sigma_v = \kappa I$ for some $\kappa > 0$. Note that the correlation matrix is far from diagonal in the empirical example in Baillie et al. while only one of the variances is very different from the others. Note also that the Gonzalo-Granger common factor is identical to the Stock-Watson stochastic trend up to scale in this case since

$$\Delta f_t = \frac{1}{\gamma'_\perp \iota} \gamma'_\perp \Delta p_t \propto \gamma'_\perp \varepsilon_t = \tau_t - \tau_{t-1}, \tag{30}$$

a further indication that the role of $\gamma_\perp$ in common factor estimation is distinct from its use in the measurement of price discovery. In addition, $\mathrm{Var}[\Delta f_t] > (<) \mathrm{Var}[\Delta \tau_t]$ if $\gamma'_\perp \iota < (>) 1$.

Hasbrouck (2002) contains three versions of a simple model he used in prior research on this general type of signal extraction problem. In each of these models, price changes follow the first order moving average process

$$\Delta p_t = \varepsilon_t + \Psi_1 \varepsilon_{t-1}, \quad \Sigma_\varepsilon = E[\varepsilon_t \varepsilon'_t] \tag{31}$$

which is related to structural model (3) via

$$\begin{aligned} C &= E[\Delta p_t \Delta p'_t] = \Sigma_\varepsilon + \Psi_1 \Sigma_\varepsilon \Psi'_1, \\ M &= E[\Delta p_t \Delta p'_{t-1}] = \Psi_1 \Sigma_\varepsilon \end{aligned} \tag{32}$$

Accordingly, Ψ_1 may be found by solving the quadratic matrix equation

$$M - \Psi_1 C + \Psi_1 M \Psi'_1 = 0 \tag{33}$$

for the value of Ψ_1 for which $I + \Psi_1 L$ is invertible. Table 1 provides the population analogues of the simulation output produced by Hasbrouck (2002) and Harris et al.

Table 1
Selected population error correction model parameters for the structural model

$$\Delta p_t = \begin{pmatrix} 1 & \lambda + c_1 & 0 \\ 0 & 0 & c_2 \end{pmatrix} \begin{pmatrix} u_t \\ q_{1t} \\ q_{2t} \end{pmatrix} + \begin{pmatrix} 0 & -c_1 & 0 \\ 1 & \lambda & -c_2 \end{pmatrix} \begin{pmatrix} u_t \\ q_{1t} \\ q_{2t} \end{pmatrix}, \quad \mathrm{Var}\left[\begin{pmatrix} u_t \\ q_{1t} \\ q_{2t} \end{pmatrix} \right] = \begin{pmatrix} \sigma_u^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad c_1 = 1 \tag{37'}$$

								Hasbrouck Information Share Bounds		Gonzalo-Granger Weights	
σ_u^2	λ	c_2	$\sigma_{\Delta m}^2$	$\sigma_{\Delta f}^2$	$\rho_{\varepsilon_1,\varepsilon_2}$	ψ_1	ψ_2	Low	High	w_f	$1 - w_f$
1	1	1	2	4.667	0.090	0.614	0.072	0.987	0.997	0.895	0.105
1	0	1	1	2.000	0.286	0.577	0.155	0.859	0.965	0.789	0.211
0	1	1	1	5.000	0.000	0.500	0.000	1.000	0.000	1.000	0.000
1	1	0	2	2.000	0.196	0.600	0.400	0.936	0.984	0.600	0.400
1	1	0.500	2	3.333	0.140	0.609	0.182	0.968	0.992	0.770	0.230
1	1	0.050	2	2.020	0.195	0.600	0.395	0.937	0.984	0.603	0.397
1	1	0.005	2	2.000	0.196	0.600	0.400	0.936	0.984	0.600	0.400
8	1	1	9	11.12	0.147	0.839	0.065	0.991	0.999	0.928	0.072
8	1	0	9	9.000	0.229	0.833	0.167	0.977	0.998	0.833	0.167
8	1	0.005	9	9.000	0.229	0.833	0.167	0.977	0.998	0.833	0.167
8	0	0.005	8	8.000	0.300	0.889	0.111	0.977	0.999	0.889	0.111

(2002b) based on (32) and (33) along with selected additional parameters of the corresponding error correction models.[10]

The simplest version, example 1 of Section 4, involves no private information or trading on stale prices and takes the form of (3) with v_t both serially uncorrelated and uncorrelated with u_t at all leads and lags. In Hasbrouck's notation, $v_{it} = c_i q_{it}$ where c_i is a cost parameter and the normalized trade size q_{it} has variance one. A tedious calculation reveals that ψ is given by

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \kappa \begin{pmatrix} c_1^{-2} \\ c_2^{-2} \end{pmatrix} = \frac{\kappa}{c_1^2 c_2^2} \begin{pmatrix} c_2^2 \\ c_1^2 \end{pmatrix},$$

$$\kappa = \frac{-\sigma_u^2}{2} + \frac{\sigma_u \sqrt{c_1^2 + c_2^2}}{2c_1^2 + c_2^2} \sqrt{(c_1^2 + c_2^2)\sigma_u^2 + 4c_1^2 c_2^2} \tag{34}$$

so that the population Gonzalo-Granger portfolio weights in this model are simply

$$\begin{pmatrix} w_f \\ 1 - w_f \end{pmatrix} = \begin{pmatrix} c_1^{-2}/(c_1^{-2} + c_2^{-2}) \\ c_2^{-2}/(c_1^{-2} + c_2^{-2}) \end{pmatrix} = \begin{pmatrix} c_2^2/(c_1^2 + c_2^2) \\ c_1^2/(c_1^2 + c_2^2) \end{pmatrix} \tag{35}$$

independent of σ_u^2. In Table 1 of Hasbrouck (2002), $c_1 = c_2 = \sigma_u^2 = 1$ (i.e., the *trade* innovation covariance matrix $\Sigma_v = I$), resulting in Gonzalo-Granger portfolio weights equal to $\frac{1}{2}$ but in population information shares ranging from 0.21 to 0.79 due to u_t, the equicorrelated component of price changes. In accord with Hasbrouck's simulations, $\mathrm{Var}[\Delta f_t] > \mathrm{Var}[\Delta m_t]$ in this case because $c_1^2 + c_2^2 > c_1^2 c_2^2$ and Δf_t follows a first order moving average process in contrast to the iid nature of Δm_t.[11] As in the Baillie et al. and de Jong examples, the Gonzalo-Granger portfolio weights diverge from $\frac{1}{2}$ with unequal costs and nonzero correlation between v_{1t} and v_{2t} distorts both the Gonzalo-Granger portfolio weights and the Hasbrouck information shares.

[10] Both Hasbrouck (2002) and Harris et al. (2002b) use simulations with sample sizes of 50,000 or 100,000 to assess the impact of alternative parameterizations on both information shares and the Gonzalo-Granger portfolio weights and common factor estimates. Their simulation output sometimes differs a bit from the population parameters calculated from (32) and (33). Footnote 6 of Hasbrouck (2002) explains why; 10–20 lag but finite autoregressions imperfectly approximate the theoretically infinite order autoregressive representation arising from the inversion of the MA(1) polynomial.

[11] That is

$$\Delta f_t = \Delta m_t + \frac{c_2^2}{c_1^2 + c_2^2} c_1 \Delta q_{1t} + \frac{c_1^2}{c_1^2 + c_2^2} c_2 \Delta q_{2t} = \Delta m_t + \frac{c_1 c_2}{c_1^2 + c_2^2}[c_2 \Delta q_{1t} + c_1 \Delta q_{2t}],$$

$$\sigma_{\Delta f}^2 = \sigma_{\Delta m}^2 + \left(\frac{c_1^2}{c_1^2 + c_2^2}\right)(2c_2^2 + 2c_1^2) > \sigma_{\Delta m}^2 = \sigma_{\Delta \tau}^2.$$

Examples 2 and 3 incorporate stale prices as well as public and private information, yielding the structural model

$$\left.\begin{aligned} p_{1t} &= m_t + c_1 q_{1t} \\ p_{2t} &= m_{t-1} + c_2 q_{2t} = m_t + c_2 q_{2t} - \Delta m_t \end{aligned}\right\} \Rightarrow p_t = \imath m_t + s_t,$$

$$s_t = \begin{pmatrix} c_1 q_{1t} \\ c_2 q_{2t} - \Delta m_t \end{pmatrix},$$

$$m_t = m_{t-1} + \lambda q_{1t} + u_t, \tag{36}$$

where c_1 and c_2 are cost parameters while q_{1t} and q_{2t} are iid, (signed) trade indicator variables with unit variances that are mutually uncorrelated at all leads and lags and that are each uncorrelated with the public information component of the efficient price change—that is, u_t—at all leads and lags as well. In this model, all price discovery takes place in the first market and the efficient price is the *only* source of serial correlation in price levels (i.e., $s_t = v_t$ and $\Upsilon(L) = I$).

The structure of example 2 and Hasbrouck's simulation of example 3—the focus of the Harris et al. (2002b) reply—is particularly simple because there are only two sources of uncertainty in the model in this case. In example 2, the efficient price is driven entirely by the information revealed in the trades in the first market (i.e., $u_t = 0$) and so price changes follow

$$\begin{aligned} \Delta p_t &= \begin{pmatrix} \lambda + c_1 & 0 \\ 0 & c_2 \end{pmatrix}\begin{pmatrix} q_{1t} \\ q_{2t} \end{pmatrix} + \begin{pmatrix} -c_1 & 0 \\ \lambda & -c_2 \end{pmatrix}\begin{pmatrix} q_{1t-1} \\ q_{2t} \end{pmatrix} \\ &= \varepsilon_t + \Psi_1 \varepsilon_{t-1}, \quad \varepsilon_t = \begin{pmatrix} (\lambda + c_1) q_{1t} \\ c_2 q_{2t} \end{pmatrix}, \quad \Psi_1 = \begin{pmatrix} \dfrac{-c_1}{\lambda + c_1} & \\ \dfrac{\lambda}{\lambda + c_1} & -1 \end{pmatrix} \end{aligned} \tag{37}$$

yielding the long-run impact matrix $I + \Psi_1 = (\lambda/(c_1+\lambda))\begin{pmatrix} \imath & 0 \end{pmatrix} = (\lambda/(c_1 + \lambda))\imath\begin{pmatrix} 1 & 0 \end{pmatrix}$. Hence, both the Hasbrouck information shares and the Gonzalo-Granger portfolio weights are $\begin{pmatrix} 1 & 0 \end{pmatrix}$ in this case; both methods succeed when all price discovery takes place in one market and trades in that market are the sole driver of the efficient price. Once again, $\mathrm{Var}[\Delta f_t] > \mathrm{Var}[\Delta \tau_t]$ since $\lambda/(c_1 + \lambda) < 1$ for $\lambda > 0$ and Δf_t follows a first order moving average process.

When the cost parameter c_2 is set to zero in example 3, price changes are given by

$$\Delta p_t = \begin{pmatrix} \lambda + c_1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} q_{1t} \\ u_t \end{pmatrix} + \begin{pmatrix} -c_1 & 0 \\ \lambda & 1 \end{pmatrix}\begin{pmatrix} q_{1t-1} \\ u_{t-1} \end{pmatrix} \tag{38}$$

which is a first order moving average with serially uncorrelated innovations. Tedious arithmetic reveals that both the Gonzalo-Granger portfolio weights and ψ take the simple form

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} w_f \\ 1 - w_f \end{pmatrix} = \begin{pmatrix} \beta_{21} \\ 1 - \beta_{21} \end{pmatrix}, \quad \beta_{21} = \frac{\lambda(c_1 + \lambda) + \sigma_u^2}{(c_1 + \lambda)^2 + \sigma_u^2} \tag{39}$$

due to the stochastic singularity created by setting $c_2 = 0$.[12] Hence, the Gonzalo-Granger portfolio weights are (0.6 0.4) for $c_1 = \lambda = \sigma_u^2 = 1$, the values used by Hasbrouck and in all but two of the models in Harris et al. (2002b). In contrast, the lower bound on the population Hasbrouck information share attributes nearly all price discovery to the first market for these parameter values. Note that Δf_t has the same variance as Δm_t and is not serially correlated in this case.

The situation is a bit more complicated when $c_2 > 0$ as in example 3 of Hasbrouck and in models A–C, E–I, and K of Harris et al. (2002b) because there are now three sources of uncertainty underlying these two prices. Another tedious calculation reveals that the Gonzalo-Granger portfolio weights are given by

$$\begin{pmatrix} w_f \\ 1 - w_f \end{pmatrix} = \begin{pmatrix} \dfrac{-Cov[\Delta p_{2t}, \Delta z_t]}{Var[\Delta z_t]} \\ \dfrac{Cov[\Delta p_{1t}, \Delta z_t]}{Var[\Delta z_t]} \end{pmatrix} + \kappa_{GG} \frac{\lambda(\lambda + c_1) + \sigma_u^2}{Var[\Delta z_t]} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$\kappa_{GG} = \frac{\sqrt{\lambda^2(\lambda + 2c_1)^2 c_2^4 + \sigma_u^2(c_1^2 + c_2^2)\{\sigma_u^2(c_1^2 + c_2^2) + 2c_2^2[\lambda^2 + 2c_1(\lambda + 2c_1)]\}}}{c_1^2\sigma_u^2 + c_2^2(\lambda^2 + \sigma_u^2)} \tag{40}$$

in this case, from which ψ can be obtained via:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \frac{1 + \kappa_{GG}}{2} \begin{pmatrix} w_f \\ 1 - w_f \end{pmatrix} \tag{41}$$

since the Gonzalo-Granger weights are scaled versions of ψ.

While these expressions may not be all that intuitive, a familiar pattern emerges. When $\sigma_u^2 = 1$, the population Hasbrouck information shares are above 0.9 in all but one case while all of the population Gonzalo-Granger portfolio weights lie below 0.9. Matters are similar when $\sigma_u^2 = 8$; the population Hasbrouck information shares are all above 0.977 while the population Gonzalo-Granger portfolio weights lie between 0.833 and 0.928. As before, correlation between signed trades creates problems for both approaches.

In summary, these examples and simulations complement the insights obtained from the analytics of cointegration. The Hasbrouck information shares correctly measure price discovery only when price change innovations are uncorrelated while the Gonzalo-Granger portfolio weights generically do so only when price change innovations have the same variance as well. When price discovery takes place in only one market and the Hasbrouck information shares correctly measure price discovery, the Gonzalo-Granger portfolio weights do so as well. Matters are a bit more complicated in the examples with uncorrelated trade innovations across markets in which price discovery takes place in only one market with the other trading on stale prices. In this case, the Hasbrouck information shares attribute the bulk of price discovery to the first market. In contrast, the Gonzalo-Granger allocation to the first market lies below its information share and deviates from one

[12] That is, we cannot simply write ε_t as a linear combination of q_{1t} and u_t along the lines of (37) because q_{1t} and u_t have no contemporaneous impact on p_{2t}.

increasingly as trading costs in the second market shrink to zero since unequal costs correspond to unequal trade innovation variances in this model. Finally, changes in the Gonzalo-Granger common factors are generically more volatile than efficient price changes.

4. Data

The previous discussion was deliberately vague about the precise data to be used in price discovery research. In one sense, one need not be concerned with the potential inability to buy at p_{1t} and sell at p_{2t} as would be the case if, for example, both prices were midquotes, bids, or offers or if φ were so large that the trades would exhaust available depth at these prices in either or both markets. The random walk efficient price signal plus stationary microstructure noise model presumes that feasible trade prices $p_t^* \neq p_t$ take the form

$$p_t^* = \iota m_t + s_t^*, \quad p_t^* = \begin{pmatrix} p_{1t}^* \\ p_{2t}^* \end{pmatrix}, \quad s_t^* = \begin{pmatrix} s_{1t}^* \\ s_{2t}^* \end{pmatrix} \tag{42}$$

so that $\varphi(1 \quad -1)(p_t^* - p_t) = \varphi(1 \quad -1)(s_t^* - s_t)$, which is also a stationary, zero mean stochastic process ignoring any unconditional means induced in s_t^* by bid–ask spreads and the like. Of course, the difference between p_t^* and p_t is of interest both to potential cross-market arbitrageurs and to financial econometricians interested in the determinants of microstructural noise.

This reasoning suggests that data should be collected in a way that maximizes the precision of γ (and thus $\gamma_\perp$) and Σ_ε estimates while respecting the stationarity requirement and that, if possible, yields a structural model with diagonal Σ_ε. This consideration suggests the use of midquotes as in most other papers in the price discovery literature, including Baillie et al. (2002) in this volume. For models estimated in calendar time, quotes are generally available continuously in most markets and the potential for stale quotes is of little concern since such quotes would simply have no impact on price discovery. Frequent sampling is generally desirable as it tends to result in low correlations among the innovations in quote midpoints. For models estimated in event time—that is, sampled only at quote changes—this approach will generate more data points than transactions prices.

One problem with using quote changes alone or transactions prices arises from nonsynchronous events across markets. Harris et al. (2002a) use transactions prices and circumvent this problem by discarding all trades save for those in which trade in one market occurs within a prespecified time of trade in the other market. At first blush, this resolution would not seem problematic since one would ordinarily suppose that data omitted in this fashion would typically be missing at random, potentially generating only a loss of efficiency relative to the inclusion of all data in a correctly specified price generating process.

However, this procedure can systematically discard important information-relevant trades or quote changes. For example, suppose trades or quote changes in a satellite market often dry up around large block trades. This circumstance generates a bias against observing data pairs initiated by trade or quote changes at stale prices in the satellite market, thereby omitting some observations that would reveal that the block trade price and not the prior stale price impinged on future prices. Similarly, any bias against trading in the satellite market immediately following large trades, perhaps due to uncertainty about their information content, would excise observations indicating that the satellite price simply followed that of the primary market. More generally, any tendency for trade or quote changes to diminish in one market when price discovery was taking place in another market naturally yields samples that understate the amount of price discovery taking place in the latter market. Put differently, one is measuring the contributions of different markets to price discovery only for episodes of nearly contemporaneous trading activity in these circumstances.

This potential problem suggests that a superior approach might blend the use of time-sampled quotes, quote changes, and trades along with descriptors of their attributes. In the stylized models analyzed in Hasbrouck (2002) and Harris et al. (2002b), observations on trades *perfectly* reveal the underlying efficient price. More generally, economic reasoning and casual empirical observation suggest that error correction processes—that is, the stationary component of prices and the corresponding error correction parameters γ and $\gamma_{\perp}$—differ across event types. One would naturally expect small trades to have little long run impact and large trades both to contribute more to price discovery and to have sizeable temporary components. That is, the market for trades of different size in a given market might be considered distinct (albeit related) venues even in the absence of a comprehensive model for both prices and quantities. Similarly, long times between quote changes may signal quote revisions that have different information content for efficient price changes. Factors such as these can be accommodated by permitting γ to change across prices differentiated in various dimensions.[13]

5. Conclusion

This special issue of the *Journal of Financial Markets* was born of the hope that these pages could resolve the apparent differences between the Hasbrouck information shares and Gonzalo-Granger approaches to price discovery research. While claiming complete victory in this regard is surely unwarranted, it is fair to say that concrete progress has been made. We knew that the Hasbrouck information shares decompose the variance of efficient price changes into components

[13] Joel Hasbrouck has pointed out that one can mix regularly sampled quotes by defining the variables of interest to be the observed quote and the last trade irrespective of the time that has elapsed since the last trade. Such an approach makes sense if the resulting first differences are covariance stationary, a circumstance that places implicit restrictions on the trade generating process.

attributable to different markets with unavoidable ambiguity when price change innovations are correlated across markets. We now know that:

- the Gonzalo-Granger portfolio weights do not resolve the inherent identification problem associated with the measurement of the contribution of different markets to price discovery but rather are proportional to the long run impact multiplier $\lim_{\tau \to \infty} \partial p_{t+\tau}/\partial \varepsilon_t'$;
- hence, the Stock-Watson common trend should be used to estimate efficient prices;
- these portfolio weights generically identify markets in which all price discovery takes place in large samples only when the Hasbrouck information shares do so as well; and
- in a number of examples involving uncorrelated trade innovations across markets in which all price discovery takes place in one market but neither method attributes all price discovery to it, the Hasbrouck information shares allocate more price discovery to it with the Gonzalo-Granger weights declining sharply as the trading costs in the second market and, thus, its trade innovation variance shrink.

These considerations suggest that it is sensible to report both $\gamma_\perp$ and Σ_ε estimates in addition to the Hasbrouck information shares. Coupled with the data considerations noted in Section 4 and the many other insights scattered throughout this volume, we have identified a number of desiderata for the measurement of price discovery across markets.

References

Baillie, R.T., Booth, G.G., Tse, Y., Zabotina, T., 2002. Price discovery and common factor models, Journal of Financial Markets, this issue.

de Jong, F., 2002. Measures of the contribution to price discovery: a comparison, Journal of Financial Markets, this issue.

Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics 13, 1–9.

Harris, F.H.deB., McInish, T.H., Shoesmith, G., Wood, R.A., 1995. Cointegration, error correction, and price discovery on informationally linked securities markets. Journal of Financial and Quantitative Analysis 30, 563–657.

Harris, F.H. deB., McInish, T.H., Wood, R.A., 2002a. Security price adjustment across exchanges: An investigation of common factor components for Dow stocks. Journal of Financial Markets, this issue.

Harris, F.H. deB., McInish, T.H., Wood, R.A., 2002b. Common factor components versus information shares: a reply, Journal of Financial Markets, this issue.

Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. Journal of Finance 50, 1175–1201.

Hasbrouck, J., 1996. Modeling microstructure time series. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics 14: Statistical Methods in Finance. Elsevier North Holland, Amsterdam, pp. 647–692.

Hasbrouck, J., 2002. Stalking the "efficient price" in market microstructure specifications: An overview, Journal of Financial Markets, this issue.