



ELSEVIER

Journal of Financial Markets 5 (2002) 309–321

Journal of
FINANCIAL
MARKETS

www.elsevier.com/locate/econbase

Price discovery and common factor models[☆]

Richard T. Baillie^a, G. Geoffrey Booth^a, Yiuman Tse^{b,*},
Tatyana Zabotina^c

^a *Eli Broad Graduate School of Management, Michigan State University, USA*

^b *College of Business, The University of Texas at San Antonio, USA*

^c *College of Business and Management, University of Illinois of Springfield, USA*

Abstract

If a financial asset is traded in more than one market, common factor models may be used to measure the contribution of these markets to the price discovery process. We examine the relationship between the Hasbrouck (J. Finance (50) (1995) 1175) and Gonzalo and Granger (J. Bus. Econ. Stat. 13 (1995) 27) common factor models. These two models complement each other and provide different views of the price discovery process between markets. The Gonzalo and Granger model focuses on the components of the common factor and the error correction process, while the Hasbrouck model considers each market's contribution to the variance of the innovations to the common factor. We show that the two models are directly related and provide similar results if the residuals are uncorrelated between markets. However, if substantive correlation exists, they typically provide different results. We illustrate these differences using analytic examples plus a real world example consisting of electronic communications networks (ECNs) and other Nasdaq market makers. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: C32; G10

Keywords: Cointegration; Information shares; Common factor components; ECNs

1. Introduction

A financial asset such as a stock often trades in multiple markets. Its price in any given market is determined (discovered) by news being gathered and interpreted in

[☆]We thank an anonymous referee, Bruce Lehmann (the editor), and Joel Hasbrouck for valuable comments. All errors are our own responsibility. The paper was completed when he and Zabasina were in Binghamton University.

*Corresponding author. Tel.: +1-210-458-6315; fax: +1-210-458-6320.

one or more of these markets. Because only the trading venue differs, intermarket arbitrage keeps the prices in the different markets from drifting apart.¹ Expressed in econometric terms, the prices are cointegrated $I(1)$ variables, which means that the price series share one or more common stochastic factors. If there is only one common factor (and this is often the case), we typically refer to it as the implicit efficient price. It is this price that is driven by news, making it the source of permanent movement in the prices of all markets.

Currently there are two popular common factor models that are used to investigate the mechanics of price discovery: Hasbrouck (1995) and Gonzalo and Granger (1995). Hereafter we refer to these models as information shares (IS) and permanent-transitory (PT), respectively. Both models use the vector error correction model (VECM) as their basis, and Hasbrouck (1996) points out that the VECM is consistent with several market microstructure models in the extant literature. Despite this initial similarity, the IS and PT models use different definitions of price discovery. Hasbrouck (1995) defines price discovery in terms of the variance of the innovations to the common factor. Thus the IS model measures each market's relative contribution to this variance. This contribution is dubbed the market's information share. Gonzalo and Granger (1995), however, are concerned with only the error correction process. This process involves only permanent (as opposed to transitory) shocks that result in a disequilibrium. In the price discovery context, disequilibria occur because markets process news at different rates. The PT model measures each market's contribution to the common factor, where the contribution is defined to be a function of the markets' error correction coefficients.

To more clearly distinguish between the two approaches of measuring price discovery, consider the case a stock being traded in two markets with the prices in these two markets being not only cointegrated but also highly correlated. The high correlation suggests that the two prices move together most of the time. Nevertheless, correlation and cointegration are different statistical concepts. Assume that the first market's price responds to deviations from the second market's price described by the error correction term, but the second market does not respond to deviations from the first market. According to the PT model price discovery only occurs in the second market.² In contrast, the IS metric suggests that both markets contribute to price discovery because of the high correlation between the two markets.

The above example suggests that there are two important and related price discovery questions to be answered. First, from a microstructure perspective, which measure of price discovery is more useful in helping us understand how markets work? Second, how are the IS and PT metrics related and under what conditions, if

¹The arbitrage argument also holds for traded financial assets that share similar characteristics. For example, it is not uncommon for stock index futures and the underlying index (portfolio of stocks) to be cointegrated. Arbitrage arguments also are often evoked in describing the relationship among commodity markets. For literature samples, see Booth et al. (1999) and Goodwin and Schroeder (1991), respectively.

²Harris et al. (1995) use a VECM to examine the price discovery process between the New York Stock Exchange and regional exchanges. Relying only on the error correction coefficients, they report that the price discovery process is focused in New York.

any, do they provide the same or similar answers? The purpose of this paper is to answer the second question and in doing so provide some clues to resolve the issues raised by the first question.

We show that the IS and PT models are directly related and both models' results are primarily derived from the error correction vector in the VECM. They provide similar results if the VECM residuals are uncorrelated. However, if substantial correlation exists, the two models usually provide different results. This is a direct result of Hasbrouck (1995) incorporating contemporaneous correlation in his metric, but Gonzalo and Granger (1995) not doing so in constructing their measure. Hasbrouck (1995) handles this correlation by using Cholesky factorization, which requires that the prices be ordered. Because the IS results are order dependent, Hasbrouck (1995) suggests that different orders be used so that upper and lower information share bounds can be calculated. Unfortunately, the bounds are often very far apart. We provide evidence to support the use of the mean of the bounds to resolve the interpretational ambiguities.

The remainder of this paper is organized in the following fashion. In Section 2 we discuss the theoretical background of cointegration, error correction, and common factors. We show that the IS and PT models are related. Section 3 provides some empirics and is divided into two parts. The first part compares the results of the two models using three analytical examples. These examples are distinguished from each other by different error correction values and each example is examined using different contemporaneous correlation values. The second part revisits Huang (2000) and examines the two price discovery metrics for five groups of Nasdaq participants. We offer concluding remarks in Section 4.

2. Cointegration, error correction, and common factors

Consider two cointegrated $I(1)$ price series, $Y_t = (y_{1t}, y_{2t})'$ with the differential being the error correction term, i.e., $z_t = \beta' Y_t = y_{1t} - y_{2t}$, and, therefore, the cointegrating vector is $\beta = (1, -1)'$. Both the IS and PT models start from the estimation of the following VECM:

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{j=1}^k A_j \Delta Y_{t-j} + e_t, \quad (1)$$

where α is error correction vector and e_t is a zero-mean vector of serially uncorrelated innovations with covariance matrix Ω such that

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}. \quad (1a)$$

σ_1^2 (σ_2^2) is the variance of e_{1t} (e_{2t}) and ρ is the correlation between e_{1t} and e_{2t} . The VECM has two portions: the first portion, $\alpha \beta' Y_{t-1}$, represents the long-run or equilibrium dynamics between the price series, and the second portion, $\sum A_j \Delta Y_{t-j}$, depicts the short-run dynamics induced by market imperfections.

Hasbrouck (1995) transforms Eq. (1) into a vector moving average (VMA)

$$\Delta Y_t = \Psi(L)e_t \quad (2)$$

and its integrated form

$$Y_t = \Psi(1) \sum_{s=1}^t e_s + \Psi^*(L)e_t, \quad (3)$$

where $\Psi(L)$ and $\Psi^*(L)$ are matrix polynomials in the lag operator, L . The impact matrix, $\Psi(1)$, is the sum of the moving average coefficients, with $\Psi(1)e_t$ being the long-run impact of an innovation on each of the prices. If the rows of the impact matrix are identical, the long-run impact is the same for all prices. If we denote $\psi = (\psi_1, \psi_2)$ as the common row vector in $\Psi(1)$, Eq. (3) becomes

$$Y_t = \iota \psi \left(\sum_{s=1}^t e_s \right) + \Psi^*(L)e_t, \quad (4)$$

where $\iota = (1, 1)'$ is a column vector of ones.

Hasbrouck (1995) states that the increment ψe_t in Eq. (4) is the component of the price change that is permanently impounded into the price and is presumably due to new information. Not included in this impact are the transient effects that may be attributed to, for example, bid–ask bounces and inventory adjustments. Hasbrouck (1995) defines this component to be the common efficient price (common factor) between the two prices. His specification is closely related to the common trend representation found in Stock and Watson (1988), i.e.,

$$Y_t = f_t + G_t, \quad (5)$$

where f_t is the common factor and G_t is the transitory component that does not have a permanent impact on Y_t . In Eq. (4), ψe_t is the common factor component and $\Psi^*(L)e_t$ is the transitory portion.

Gonzalo and Granger (1995) define the common factor to be a combination of the variables Y_t , such that $f_t = \Gamma Y_t$, where Γ is the common factor coefficient vector. The identification of the common factor is achieved by imposing that the error correction term does not Granger-cause the common factor in the long run. Not only do Gonzalo and Granger show that Γ is the orthogonal to the error correction coefficient vector α , denoted by $\alpha_\perp = (\gamma_1, \gamma_2)'$, but also they develop a statistic to test whether one of the factor components is the sole contributor to the common factor. Their specification of f_t is equivalent to considering it to be the price of a portfolio with Γ serving as the portfolio weights. Thus f_t may be expressed in terms of either price, i.e., y_{1t} and y_{2t} , and the error correction term, z_t . Since both price series are $I(1)$ and the error correction term is $I(0)$ as a result of the price series being cointegrated, f_t is $I(1)$.³ Moreover, because the size of z_t is almost surely small relative to y_{it} , f_t 's evolutionary process is dominated by y_{it} .

³To illustrate that f_t is $I(1)$, recall that $f_t = \gamma_1 y_{1t} + \gamma_2 y_{2t}$ and that $z_t = y_{1t} - y_{2t}$. Substituting the latter equation into the former one results in $f_t = y_{2t} + \gamma_1 z_t = y_{1t} - \gamma_2 z_t$. If Δf_t is not (is) autocorrelated, f_t is (is not) a martingale process.

That $\alpha_{\perp}$ is the common factor coefficient vector is intuitive. To illustrate this point, we switch from financial markets and instead rely on Gonzalo and Granger's (1995) use of Cochrane's (1991) consumption-GNP example. The VECM model between consumption (c) and GNP (y) are

$$\begin{aligned}\Delta c_t &= \alpha_1 z_{t-1} + \cdots + e_{1t}, \\ \Delta y_t &= \alpha_2 z_{t-1} + \cdots + e_{2t},\end{aligned}\tag{6}$$

where the error-correction term, $z_{t-1} = c_{t-1} - y_{t-1}$. Gonzalo and Granger (1995) find that α_1 is not significantly different from zero while α_2 is significant. Thus, income adjusts to z_{t-1} , but consumption does not, indicating that consumption Granger-causes income in the long run. Since $\alpha = (0, 1)'$, $\alpha_{\perp} = (1, 0)'$. The common factor does not include the income series; therefore, a change in income only has transitory effects on consumption and income. Only a change in consumption will have permanent effects. These results parallel the VECM results.

The crux of the IS and PT models is that they decompose the impact of a perturbation and allocate this impact to the markets. The Gonzalo and Granger (1995) model decomposes the common factor into a combination of the two prices. Hasbrouck (1995), however, decomposes the variance of the common factor innovations, i.e., $\text{var}(\psi e_t) = \psi \Omega \psi'$. In other words, the information share of a market is the proportion of variance in the common factor that is attributable to innovations in that market. Although seemingly dissimilar, both models' results are primarily derived from $\alpha_{\perp}$. Continuing the previous notation, we note that Johansen (1991) shows that

$$\begin{aligned}\Psi(1) &= \beta_{\perp} \Pi \alpha'_{\perp}, \\ \Pi &= \left(\alpha'_{\perp} \left(I - \sum_{j=1}^k A_j \right) \beta_{\perp} \right)^{-1},\end{aligned}\tag{7}$$

where $\beta_{\perp}$ is the orthogonal matrix to β , and I is the identity matrix, with Π being a scalar if there is only one common factor in the system.

Since $\beta = (1, -1)'$, $\beta_{\perp} = (1, 1)'$. We can prove that

$$\Psi(1) = \begin{pmatrix} \psi \\ \psi \end{pmatrix} = \Pi \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_2 \end{pmatrix}.\tag{8}$$

Therefore, $\psi = (\psi_1, \psi_2)$ is directly related to Γ , i.e.,

$$\frac{\psi_1}{\psi_2} = \frac{\gamma_1}{\gamma_2}.\tag{9}$$

Eq. (9) demonstrates that what is important is the relative values of ψ_j or γ_j and not their individual values. In other words, γ_j 's are unique up to a scale factor that drops out in the IS and PT measures.

As Hasbrouck (1995) argues, if Ω is diagonal, then $\psi \Omega \psi'$ will consist of two terms, the first (second) represents the contribution to the common factor innovation from the first (second) market. The information share of market j is

defined as

$$S_j = \frac{\psi_j^2 \sigma_j^2}{\psi \Omega \psi'}. \quad (10)$$

Substituting (9) into (10) yields

$$S_j = \frac{\gamma_j^2 \sigma_j^2}{\gamma_1^2 \sigma_1^2 + \gamma_2^2 \sigma_2^2}, \quad (11)$$

$$\frac{S_1}{S_2} = \frac{\gamma_1^2 \sigma_1^2}{\gamma_2^2 \sigma_2^2}. \quad (12)$$

Eqs. (11) and (12) are valid only if there is no correlation between the error terms. The equations show that the relative information share of market j is the square of its common factor component, γ_j , weighted by its variance, σ_j^2 . If σ_1^2 and σ_2^2 are similar in value (as likely for similar informationally linked markets), results of the IS and PT are qualitatively similar.⁴

This bivariate system can easily be generalized to an n -variate system. Substituting (9) into (11) and (12) yields

$$S_j = \frac{\gamma_j^2 \sigma_j^2}{\sum_{i=1}^n \gamma_i^2 \sigma_i^2}, \quad (11a)$$

$$\frac{S_j}{S_k} = \frac{\gamma_j^2 \sigma_j^2}{\gamma_k^2 \sigma_k^2}. \quad (12a)$$

However, if the price innovations are significantly correlated across markets, Eq. (11) does not hold. In this case, Hasbrouck (1995) uses the Cholesky factorization of $\Omega = MM'$ to eliminate the contemporaneous correlation, where M is a lower triangular matrix. As proved in the following, the largest (smallest) information share value occurs when the variable is first (last) in the sequence, assuming the cross correlation ρ is positive.

The information shares are given as follows:

$$S_j = \frac{([\psi M]_j)^2}{\psi \Omega \psi'}, \quad (13)$$

where $[\psi M]_j$ is the j th element of the row of matrix ψM . Denote

$$M = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho \sigma_2 & \sigma_2(1 - \rho^2)^{1/2} \end{pmatrix}. \quad (14)$$

Using (9) and (13),

$$\frac{S_1}{S_2} = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_2 m_{22})^2}. \quad (15)$$

⁴ Martens (1998) provides a result identical to Eq. (11), but he does not show its relevance to the PT model.

By noting that $S_1 + S_2 = 1$,

$$S_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, \quad (16)$$

$$S_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}. \quad (17)$$

Eqs. (16) and (17) show that the information shares only depend on α (or its orthogonal) and Ω . They also show that the factorization imposes a greater information share on the first price, unless $m_{12} = 0$ (i.e., no correlation between market innovations.) Hasbrouck (1995) considers the upper (lower) bound of market j 's information share when market j is the first (second) variable in the factorization. Moreover, Eqs. (16) and (17) also indicate that the higher correlation, the greater (smaller) the upper (lower) bound. Intuitively, the upper bound incorporates the series' own contribution from σ_1 (as represented by m_{11} in (14)) and its correlation with the other series from $\rho\sigma_2$ (as indicated by m_{12}). The lower bound only considers the series' "pure" contribution that is uncorrelated with the other series as indicated by $m_{22} = \sigma_2(1 - \rho^2)^{1/2}$.

Hasbrouck (1995) reports that the upper and lower bounds in his study of price discovery between the NYSE and off-NYSE quotes are almost the same. The reason is that he uses one-second interval of prevailing quotes. For such a high frequency, the contemporaneous correlation is insignificant (see also Tse, 2000). However, all other studies using the model with a lower frequency report significant differences between the upper and lower bounds. Huang (2002), for instance, uses one-minute intervals to examine the nature of the price discovery between the electronic communications networks (ECNs) and various Nasdaq dealers. The lower and upper bounds of the Island (an ECN) for Yahoo, the last company in Table 5a of Huang (2002), are 79.5% and 30.6%, respectively, for the month of January 1998. For the month of November 1999 (Table 5b), the upper and lower bounds are 47.7% and 8.4%. In their study of Finnish upstairs and downstairs stock markets, Booth et al. (2002) use trading intervals averaging approximately 30 min. They report the information share upper and lower bounds for the downstairs market to be, on average, 99.2% and 13.0%, respectively.⁵

In Hasbrouck (1995) and all the studies using this model, the information shares are estimated using the VMA representation of the VECM. However, Eqs. (15)–(17) show that the information shares can be calculated by the VECM directly without using the VMA procedure. This makes the estimation process much easier. Moreover, these three equations can be generalized to an n -variate system, i.e.,

$$\frac{S_j}{S_k} = \frac{\left[\sum_{i=j}^n \gamma_i m_{ji}\right]^2}{\left[\sum_{i=k}^n \gamma_i m_{ki}\right]^2}. \quad (15a)$$

⁵The differences between the lower and upper bounds reported in Martens (1998) and Tse (1999) in international futures markets are also very large.

The upper (lower) bound of a market's information share with the series being the first (last) series is given by S_1 (S_n):

$$S_1 = \frac{[\sum_{i=1}^n \gamma_i m_{i1}]^2}{[\sum_{i=1}^n \gamma_i m_{i1}]^2 + [\sum_{i=2}^n \gamma_i m_{i2}]^2 + \cdots + (\gamma_n m_{nn})^2}, \quad (16a)$$

$$S_n = \frac{(\gamma_n m_{nn})^2}{[\sum_{i=1}^n \gamma_i m_{i1}]^2 + [\sum_{i=2}^n \gamma_i m_{i2}]^2 + \cdots + (\gamma_n m_{nn})^2}. \quad (17a)$$

3. Some empirics

As indicated in the previous section, the IS model results typically depend on the ordering of the variables in the Cholesky factorization of the innovation covariance matrix. The first (last) variable in the ordering tends to have a higher (lower) information share, and this discrepancy may be very large if the series' innovations are highly contemporaneously correlated. In this section we empirically explore this phenomenon and compare IS model information shares to PT model common factor weights.

3.1. Three analytical examples

We first investigate three analytic examples using the following error correction model for two cointegrated series x_1 and x_2 :

$$\Delta x_{1t} = -\alpha_1(x_{1,t-1} - x_{2,t-1}) + e_{1t},$$

$$\Delta x_{2t} = \alpha_2(x_{1,t-1} - x_{2,t-1}) + e_{2t},$$

$$\text{cov}(e_1, e_{2t}) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad (18)$$

where α_i and α_2 are positive constants and $e_{it} \sim N(0, 1)$.

In Table 1 we report the IS model information shares and the PT model common factor weights for three different examples of error correlation behavior. Example A is a symmetric system with $\alpha_1 = \alpha_2 = 0.05$. Example B is a strongly asymmetric system that is constructed so that only x_2 reacts to a disequilibrium. In this example, $\alpha_1 = 0.0$ and $\alpha_2 = 0.05$. Example C depicts a moderately asymmetric system with $\alpha_1 = 0.025$ and $\alpha_2 = 0.05$. In each example, four different values of ρ are used. These range from 0.0 to 0.9. Because the information shares and common factor weights for x_1 and x_2 are complementary, we report only the results for x_1 . We also report the upper and lower bound values for this variable. The upper (lower) bound for x_1 is associated with x_1 (x_2) being the first variable in the Cholesky factorization. The reported common factor weights are derived directly from the error correction coefficients.

Table 1

IS and PT metrics for three analytic examples

This table investigates three analytic examples using the following error correction model for two cointegrated series x_1 and x_2 :

$$\Delta x_{1t} = -\alpha_1(x_{1,t-1} - x_{2,t-1}) + e_{1t},$$

$$\Delta x_{2t} = \alpha_2(x_{1,t-1} - x_{2,t-1}) + e_{2t},$$

$$\text{cov}(e_{1t}, e_{2t}) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where α_i and α_2 are positive constants and $e_{it} \sim N(0, 1)$.

Example	ρ	IS model information share for x_1			PT model common factor weight for x_1
		Upper bound	Lower bound	Midpoint	
A. Symmetric System: $\alpha_1 = \alpha_2 = 0.05$					
	0.0	0.50	0.50	0.50	0.50
	0.1	0.55	0.45	0.50	0.50
	0.5	0.75	0.25	0.50	0.50
	0.9	0.95	0.05	0.50	0.50
B. Strong asymmetric system: $\alpha_1 = 0.00$ and $\alpha_2 = 0.05$					
	0.0	1.00	1.00	1.00	1.00
	0.1	1.00	0.99	1.00	1.00
	0.5	1.00	0.75	0.88	1.00
	0.9	1.00	0.19	0.60	1.00
C. Moderate asymmetric system: $\alpha_1 = 0.025$ and $\alpha_2 = 0.05$					
	0.0	0.80	0.80	0.80	0.67
	0.1	0.82	0.73	0.78	0.67
	0.5	0.89	0.43	0.66	0.67
	0.9	0.98	0.09	0.54	0.67

Let us first review the results of Example A, the symmetric system. The information shares in Eqs. (16) and (17) can be simplified as

$$S_1 = \frac{1 + \rho}{2}, \quad (16b)$$

$$S_2 = \frac{1 - \rho}{2}. \quad (17b)$$

If $\rho = 0.0$, the upper and lower bound values for the information shares are 0.50. This value is identical to the common factor weights. However, this equivalency holds only if there is no contemporaneous correlation. As this correlation increases, the upper and lower bounds of the information shares change, with the upper bound becoming larger and the lower bound getting smaller. If $\rho = 0.90$, the spread between the two bounds is so large that any economic inference is meaningless. The common factor weights does not change, however, because this metric does not

depend on ρ . The mean (in this case the midpoint as well) of the upper and lower boundaries is always 0.50.

A somewhat different picture emerges in Example B, a strong asymmetric system. Once again if $\rho = 0.0$, the information share upper and lower bound values are the same and, in turn, equal to the common factor weight. Instead of 0.50, however, the values of all the metrics are 1.00. As the contemporaneous correlation increases, the information share of the upper bound remains at 1.00 but the lower bound value becomes progressively smaller and the spread, from an economic perspective, becomes uncomfortably large. Moreover, the midpoint of the bounds also decreases as ρ increases.

Example C, a moderate symmetric system, represents a situation somewhere between the previous two examples. If $\rho = 0.0$, the information share upper and lower bound values are the same, but they are different from the common factor weight. A re-examination of Eq. (11) indicates this is because the relative information share of x_1 is the square of its relative common factor component.⁶ Similar to the other examples, as ρ increases, the information share upper (lower) bound value of x_1 increases (decreases). In addition, the midpoint decreases as ρ increases, but this metric modestly tracks the common factor weight.

The above three examples support the notion that the IS and PT models measure different aspects of the cointegrated system. They show that if the residual correlation is zero or small and the residual variances are equal, both models provide consistent results for a symmetric system and a system that is strongly asymmetric. If the contemporaneous correlation of the innovations is very large, however, the results can be very different. Permitting unequal variances aggravates these difference. In microstructure applications, this inequality of variances (induced by temporary effects) can be significant, as shown by the following empirical results.

3.2. Price discovery by ECNs

We now turn to a real market situation. An ECN is a computer-mediated market that widely disseminates buy and sell limit orders from its subscribers.⁷ It is based on an open limit order book and automatically matches orders submitted electronically and anonymously by its customers. When an order cannot be matched, the ECN posts its best bid or ask quote on the Nasdaq trading screen under the ECN's own name. In this case, the trade will be executed if the ECN's quote is at the inside, and someone who sees it on the screen enters an offsetting order at that price. Therefore, the ECN acts both as a quasi-stock exchange and as a broker. ECNs are used by large institutions and market makers. Market makers often balance their inventory positions by trading on the ECNs anonymously instead of offering attractive quotes on the Nasdaq system. Although the Securities and Exchange Commission (SEC)

⁶ According to Eq. (12), $0.80/0.20 = (0.67/0.33)^2$. Note that the variance of the innovation of x_i are equal, thereby canceling each other.

⁷ Today, there are 10 operating ECNs on the market, with Instinet (INCA) and Island (ISLD) being the two most popular.

rule implemented in 1997 requires quotes sent to ECNs by market makers to be included in the Nasdaq national best bid and offer montage, the possibility of trading anonymously and absence of market makers attracts informed investors to the ECNs.⁸

Huang (2002) uses July 1998 data for the 30 most active Nasdaq stocks to analyze price discovery of quotes by ECNs and Nasdaq marketmakers. In this paper, we use the mid-quotes (or the average of the bid and ask quotes) of Yahoo (YHOO) in March 1999 with a sample size of 9003. The results are similar if we use bid or ask quotes only. Similar to Huang (2002), we adopt a one-minute time interval and identify five groups of Nasdaq participants: ECNs (x_1), wholesalers (x_2), wire houses (x_3), institutional brokers (x_4), others (x_5).

Based on the results of Johansen (1991), the five series are cointegrated with one common factor. For brevity, we do not report these expected results. According to the PT model, the coefficient vector of the common factor, $\Gamma = \alpha_{\perp} = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)' = (0.693, 0.150, 0.036, 0.001, 0.119)'$. The model also indicates that only the error correction coefficients of x_1 and x_2 are significant at the 1% level. These results indicate that ECNs are the common factor's main component. More specifically, all the markets significantly adjust to the deviations from the ECN quotes, but the ECNs only respond moderately to the deviations from wholesalers, and do not respond to the deviations from wire houses, institutional brokers, and others.

As shown in the previous section, we can estimate the information shares by using Γ and the covariance matrix of the residuals, Ω , directly. In particular,

$$\Omega = \begin{bmatrix} 0.358 & & & & \\ 0.254 & 0.353 & & & \\ 0.206 & 0.207 & 0.451 & & \\ 0.242 & 0.235 & 0.229 & 0.805 & \\ 0.266 & 0.236 & 0.208 & 0.239 & 0.338 \end{bmatrix}. \quad (19)$$

The correlation matrix, Ξ , derived from this covariance matrix is

$$\Xi = \begin{bmatrix} 1.000 & & & & \\ 0.713 & 1.000 & & & \\ 0.552 & 0.557 & 1.000 & & \\ 0.451 & 0.440 & 0.409 & 1.000 & \\ 0.715 & 0.637 & 0.534 & 0.427 & 1.000 \end{bmatrix}. \quad (20)$$

Ξ shows that the residuals are highly correlated, and therefore, the upper and lower bounds of the information shares will be considerably different. Using Eqs. (16a) and (17a), the upper and lower bounds of ECNs are calculated to be 0.968 and 0.204 with a midpoint of 0.586. These results are similar to Example C in Section 3.

⁸ Huang (2002) provides a detailed discussion of the trading mechanisms of the ECNs.

If the midpoint is used as the representative information share value, the IS model shows that ECNs contribute 58.6% of the innovations to the implicit efficient price.⁹ The other four market makers contribute the remainder (41.4%). The PT model indicates that although all the market makers adjust to the deviations from the ECNs' quotes, ECNs do not adjust to the deviations from other markets' quotes. The ECNs comprise 69.3% of the implicit efficient price, with the other four markets contributing 30.7%. Both measures show that the ECNs dominate other four markets in price discovery. The ECN dominance, however, is less material according to the IS model than the PT model. The reason is that the IS model incorporates the correlation between the series, while the latter does not.

4. Concluding remarks

The common stochastic factor, or implicit efficient price, in cointegrated financial series has received substantial attention from the researchers continuously trying to refine their estimations of market quality and efficiency. We explore the relationship between the popular common-factor models developed by Hasbrouck (1995) and Gonzalo and Granger (1995).

Both models use the VECM as their starting point. The feature that distinguishes them from each other is that Hasbrouck (1995) decomposes the implicit efficient price variance. Relying on the premise that the price volatility reflects the flow of information, he attributes a greater share of the efficient price discovery to the market that contributes the greatest share to this volatility. In contrast, Gonzalo and Granger's (1995) approach is to decompose the common factor itself. In doing so, the Gonzalo and Granger (1995) model ignores the correlation among the markets and attributes the leading role solely to the market that adjusts least to the price movement in the other markets.

We show that these two models are related, a connection that substantially simplifies the estimation of Hasbrouck's (1995) model. In particular, relative values of Hasbrouck's (1995) information shares are proportional to the square of the relative markets' common factor loadings from the Gonzalo and Granger (1995) model weighted by their variances, if the residuals from this model are uncorrelated. In markets affected by the same information flow (i.e., with similar volatility), the two models produce consistent results, i.e., the market with the greatest contribution to the price discovery has the largest loading on the common factor. If the residuals are correlated, however, the Hasbrouck (1995) model's upper and lower bounds are not the same and the spread between the two bounds is positively related to the degree of correlation. This correlation is determined partially by the information flows between the markets and partially by the frequency of the price data, with very high frequency data typically being less correlated. In this case, we argue that the average of Hasbrouck's (1995) upper and lower bounds provides a sensible estimate of the markets' roles in the production of the efficient price.

⁹The midpoint is also close to the average of all the orderings of the Cholesky factorization.

In conclusion, it would be satisfying to declare that one of the two models provides a better measure of price discovery. We cannot do this, however, because it depends on whether price discovery is considered to be solely an error correction phenomenon or if it should also consider the correlations among the markets' innovations. Nevertheless, we believe that Hasbrouck's (1995) reliance on the relative innovation variances has more general economic appeal and interpretation.

References

- Booth, G.G., So, R., Tse, Y., 1999. Price discovery in the German equity derivatives markets. *Journal of Futures Markets* 19, 619–643.
- Booth, G.G., Lin, J.-C., Martikainen, T. and Tse, Y., 2002. Trading and pricing in upstairs and downstairs markets. Forthcoming in *Review of Financial Studies*.
- Cochrane, J.H., 1991. Univariate vs. multivariate forecasts of GNP growth and stock returns: Evidence and implications for the persistence of shocks, detrending methods, and tests of the permanent income hypothesis. Working paper 3427, NBER, Cambridge, MA.
- Goodwin, B.K., Schroeder, T.C., 1991. Cointegration tests and spatial price linkages in regional cattle markets. *American Journal of Agricultural Economics* 73, 452–464.
- Gonzalo, J., Granger, C.W.J., 1995. Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics* 13, 27–36.
- Harris, R., McNish, T.H., Shoesmith, G.C., Wood, R.A., 1995. Cointegration, error correction, and price discovery informationally linked security markets. *Journal of Financial and Quantitative Analysis* 30, 563–579.
- Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. *Journal of Finance* 50, 1175–1199.
- Hasbrouck, J., 1996. Modeling microstructure time series, in: *Handbook of Statistics 14: Statistical Methods in Finance*, G.S. Maddala and C.R. Rao, eds. (Elsevier North Holland, Amsterdam), 647–692.
- Huang, R.D., 2002. Price discovery by ECNs and Nasdaq market makers. *Journal of Finance*, forthcoming.
- Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–1580.
- Martens, M., 1998. Price discovery in high and low volatility periods: Open outcry versus electronic trading. *Journal of International Financial Markets, Institutions & Money* 8, 243–260.
- Stock, J.H., Watson, M.W., 1988. Variable trends in economic time series. *Journal of Economic Perspectives* 2, 147–174.
- Tse, Y., 1999. Round-the-clock market efficiency and home bias: Evidence from international Japanese government bonds futures markets. *Journal of Banking and Finance* 23, 1831–1860.
- Tse, Y., 2000. Further examination of price discovery on the NYSE and regional exchanges. *Journal of Financial Research* 23, 331–351.