ELSEVIER

Journal of Financial Markets 5 (2002) 341–348

Journal of FINANCIAL MARKETS

www.elsevier.com/locate/econbase

Common factor components versus information shares: a reply ☆

Frederick H. deB. Harris[a,*], Thomas H. McInish[b], Robert A. Wood[b]

[a] *Babcock Graduate School of Management, Wake Forest University, Winston-Salem, NC 27109, USA*
[b] *Institute for the Study of Security Markets, University of Memphis, Memphis, USA*

Abstract

Common factor components and information shares provide competing approaches to estimating the parameters of price discovery in cointegrated security markets or trading channels. Using simulated data on Hasbrouck's (J. Financial Markets 5(3) (2002)) parameterized model of the stylized facts in satellite and centralized markets, we show that Gonzalo and Granger's (J. Business Econom. Statist. 13 (1995) 1) procedure for estimating and *testing* common factor components recovers the true information structure in a wide range of financial market microstructure models. In addition, we investigate the role of stochastic process assumptions in estimating price discovery parameters by generalizing the sources of volatility in the implicit efficient price and the level of the signal to noise ratio.

JEL classification: G12

Keywords: Error correction; Common factor components; Information shares; Price discovery

1. Introduction

Common factor components (Gonzalo and Granger, 1995; Harris et al., 1995a; Harris, McInish, and Wood (HMW), 1997, 2002) and information shares (Hasbrouck, 1995, 2002) provide competing approaches to estimating the

☆ Frank deJong and Bruce Lehmann made insightful comments and suggestions; the usual disclaimer applies. Amy Cecil and Joe Huang provided research assistance.

*Corresponding author. Worrell Professional Center, Wake Forest Drive, Wake Forest University, Winston-Salem, NC 27109, USA. Tel.: +1-336-758-5112; fax: +1-336-758-4514.

E-mail address: rick.harris@mba.wfu.edu (F.H. deB. Harris).

PII: S1386-4181(02)00030-7

parameters of price discovery in cointegrated time series. Hasbrouck (2002) uses a simulation model to criticize the Gonzalo-Granger (GG) approach for providing biased estimates of the true price discovery parameters when satellite markets *costlessly* cross uninformative trades at the central market's stale quote midpoint.

Our purpose here is to highlight the pivotal importance of statistical testing of price discovery parameters, something the information shares approach cannot offer and Hasbrouck (2002) does not do. We show that Gonzalo and Granger's Q_{GG} test statistic for the common factor components can be used to recover the exact information structure in Hasbrouck's simulation model. Moreover, our testing elucidates what features of Hasbrouck's trading environment (i.e., the zero half-spread in the satellite market and a low signal to noise ratio) are responsible for the bias in the common factor results. Again, our ability to test subsets of the common factor component vector is the key.

2. Testable price discovery hypotheses

deJong (2002) derives the relationship between the common factor component and information share approaches by showing that the GG permanent–transitory decomposition is a Stock-Watson common stochastic trend (Σw_t) plus an additively separable idiosyncratic transitory disturbance ($\varepsilon_{j,t}$). In HMW (2002) we formulated a microstructure-theoretic model of innovations in a GG trading price sequence $\Delta P_{j,t} = w_t + \Delta\varepsilon_{j,t}$ from which we derived the VECM in the ith of three venues as Eq. (7), repeated here for convenience,

$$\Delta P_{i,t} = \alpha_i + \sum_{j=1}^{3} \gamma_{\perp i,j} \sum_{s=1}^{S} w_{t-s} + \sum_{j=1}^{3} \sum_{s=1}^{S} \beta_{i,j,t-s} \Delta\varepsilon_{j,t-s} + \sum_{j \neq i}^{2} \gamma_{i,j}(\varepsilon_{i,t-1} - \varepsilon_{j,t-1}) + u_{i,t}, \tag{I}$$

where $\gamma_{\perp j}$ are common factor components or weights for each venue, $\sum_{s=1}^{S} w_{t-s}$ is the Stock-Watson common stochastic trend, S is an optimal lag length in the corresponding system of VAR equations, $\varepsilon_{j,t}$ are idiosyncratic transitory disturbances in venue j, $(\varepsilon_{i,t-1} - \varepsilon_{j,t-1})$ is the resultant of the error correction term specified as an intermarket price divergence $(P_{i,t-1} - P_{j,t-1})$, and $u_{i,t}$ is an empirical estimation error that may be contemporaneously correlated across trading venues.

GG's (1995) common factor estimation procedure and likelihood ratio test statistic (Q_{GG}) analyze the individual elements of the vector of common factor weights $\boldsymbol{\gamma}_{\perp j}$ that is orthogonal to the error correction parameter vector $\boldsymbol{\gamma}$ in Eq. (I). Speaking intuitively, the lower the estimated error correction parameter γ_j for a trading venue, the higher the common factor weight $\boldsymbol{\gamma}_{\perp j}$ and the bigger the contribution of that jth venue to revelation of the permanent innovations underlying the common stochastic trend.

Under the *multilateral price adjustment hypothesis*, at various times innovations in the implicit efficient price $\mathscr{P}_t$ (i.e., the w_t) arrive first in each trading venue. Then, within one or two synchronous trades for stocks like those in the DJIA, the w_t are impounded into the trading prices in all three markets. Empirically, the testable *multilateral price adjustment hypothesis* is that all the γ_j and therefore all the common factor components, the $\gamma_{\perp j}$ weights, are statistically significantly different from zero. That is, the central market prices would at times error correct to a divergence from a satellite regional exchange price just as the regionals often error correct to a divergence from New York. The next two sections examine within this VECM framework a testable *one-way price adjustment hypothesis*.

3. Critique of Hasbrouck's one-way price discovery model

To isolate the polar case of one-way price adjustment, Hasbrouck (2002) assumes that satellite market prices $P_{\mathrm{R},t}$ are determined by an idiosynchratic transitory disturbance among regional participants ($\varepsilon_{\mathrm{R},t} = c_{\mathrm{R}} q_{\mathrm{R},t}$) plus a quote midpoint on the central market ($P_{\mathrm{C},t-1} - c_{\mathrm{C}} q_{\mathrm{C},t-1}$) that is assumed to be fully informative but stale (i.e., equal to the lagged implicit efficient price $\mathscr{P}_{t-1}$).[1] The parameters c_{C} and c_{R} are half-spreads in the central and satellite markets, respectively, and $q_{\mathrm{C},t}$ and $q_{\mathrm{R},t}$ are equi-probable random indicator variables signifying the direction of the last trade reported in those markets (+1 for trades at the ask and −1 for trades at the bid).[2] The implicit efficient price evolves according to Eq. (3) in Table 1, where λ reflects a possible sensitivity of $\mathscr{P}_t$ to trading pressure that arises from private values (i.e., noise) in the central market and where provisionally $\lambda = 1$ and $\sigma_w^2 = 1$.[3] The random variables w_t, and $q_{\mathrm{C},t}$ or $q_{\mathrm{R},t}$ are independent standard normal and two bivariate normal random indicator variables with no serial correlation.[4]

Hypothesizing selective execution of uninformed orders, Hasbrouck (2002) investigates satellites that simply execute crosses at the lagged NYSE quote midpoint plus a tiny (perhaps zero) half-spread to cover their much-reduced transaction costs. In his model, the true information share and the true common factor weight for the central market are 1.0, by construction, since the structural

[1] Harris, McInish, and Chakravarty (1995b) and Harris and McInish (2000) provide another motivation for analyzing apparently "stale" quotes and trades by modeling the NYSE as a non-market clearing queue service mechanism.

[2] Some of these assumptions are imperfectly suited to a comparison of alternative permanent–transitory decompositions. In particular, the half-spread times a trade direction indicator variable is an imperfect candidate for the transitory component in a P–T decomposition of security prices because in general c_{C} and c_{R} are not constants. Instead, the adverse selection component of c is a time-varying function of the w_t permanent components – i.e., of the information arrivals.

[3] Traditionally, a quote that is simply old is not "stale" if the w_t information that arrived since it's posting has zero expected value $E(w_t) = 0$. Here, however, Hasbrouck assumes that with $\lambda \neq 0$ in Eq. (3), *any* trades on the central market during the interval t – even pure noise trades – will move the implicit efficient price thereby making the prior quote "stale".

[4] Bertsimas and Lo (1998) criticize the whole approach of using trade direction indicator variables as a transitory component precisely because of the high serial correlation present in the trade direction data.

Table 1
Testing of the one-way price adjustment hypothesis in common factor models

The model to be investigated is

$$P_{R,t} = (P_{C,t-1} - c_C q_{C,t-1}) + c_R q_{R,t}, \tag{1}$$

$$P_{C,t} = \mathscr{P}_t + c_C q_{C,t}, \tag{2}$$

where $P_{R,t}$ and $P_{C,t}$ are the trading prices in the satellite (regional) and central markets, $\mathscr{P}_t$ is the true implicit efficient price, c_C and c_R are half-spreads in the central and satellite markets, $q_{C,t}$ and $q_{R,t}$ are equiprobable random indicator variables (bivariate normal and provisionally assumed independent) signifying the direction of the last trade reported on the central and satellite markets (+1 for trades at the ask and −1 for trades at the bid). The one-way price discovery hypothesis is represented by the stipulated information structure in that the satellite price equals a fully informative but stale central market quote midpoint $(P_{C,t-1} - c_C q_{C,t-1})$ plus a localized half-spread. The implicit efficient price $\mathscr{P}_t$ evolves according to

$$\mathscr{P}_t = \mathscr{P}_{t-1} + \lambda q_{C,t} + w_t, \quad w \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_w^2), \tag{3}$$

where w_t indicates the random arrival of public information about common value, and λ is a possible sensitivity of the implicit efficient price to trading pressure that arises from private information in each market environment.

Letting $\lambda = 1, \sigma_w^2 = 1, c_C = 1$ and $c_R = 1$, these equations establish the baseline model, Model A, that is Hasbrouck's (2002) Price Discovery Example III. Other parametric assumptions in subsequent rows are restrictions on the specifications in Eqs. (1)–(3) defining ten other models (B–K). Each model is simulated with 50,000 observations on the three random variables $w_t, q_{C,t}$ and $q_{R,t}$. Common factor weights $(\gamma_\perp)$ are estimated and tested with reduced-rank regressions and eigenvector calculations based on Gonzalo and Granger (1995), and the one-way price discovery null hypothesis H_0: $[\gamma_{\perp C}\ \gamma_{\perp R}] = [1\ 0]$ is then tested with the Q_{GG} test statistic. The reported parameter estimates are averages that stabilized by the fiftieth simulation. Using the 0.01 or 0.05 p-values typical of microstructure research, in all eleven cases GG common factor estimation and testing fails to reject H_0 and therefore recovers the true information structure specified in Hasbrouck's model.

	Parameter values				Optimal system lag length	Estimates and test results for central market common factor weight (H_0: $\gamma_{\perp C} = 1.0$)			
	λ	σ_w^2	c_R	c_C		Estimated $\gamma_{\perp C}$	Q_{GG}	p-value	Conclusion
Baseline model and restrictions									
Model A	1	1	1	1	3	0.91	0.422	0.52	Accept the
Model B	0	1	1	1	1	0.92	0.365	0.54	null hypothesis
Model C	1	0	1	1	4	0.95	0.112	0.74	
Model D	1	1	0	1	1	0.59	3.176	0.08	
Declining satellite spreads that recover the specified information structure									
Model E	1	1	0.5	1	2	0.76	1.041	0.30	Accept the
Model F	1	1	0.05	1	1	0.63	1.312	0.28	null hypothesis
Model G	1	1	0.005	1	1	0.59	2.857	0.15	
Setting σ_w^2 to reflect heightened information volatility									
Model H	1	8	1	1	1	0.84	0.117	0.73	Accept the
Model I	1	8	0.005	1	1	0.83	0.165	0.69	null hypothesis
Model J	1	8	0	1	1	0.82	0.072	0.79	
Model K	0	8	0.005	1	1	0.96	0.038	0.84	

equations clearly show that $P_{C,t}$ always impounds the contemporaneous implicit efficient price $\mathscr{P}_t$ first, and $P_{R,t}$ never does. The testable empirical question of interest is whether the *estimated* parameter vector γ' on the error correction term in Eq. (I) will be [0 1] and therefore whether the $\gamma'_{\perp}$ vector of common factor weights $[\gamma_{\perp C}\ \ \gamma_{\perp R}]$ will be [1 0]? We refer to this below as the *one-way price adjustment null hypothesis* H_0: $[\gamma_{\perp C}\ \ \gamma_{\perp R}] = [1\ \ 0]$.

Hasbrouck (2002) estimates the central market's information share as 0.9 to 0.98, and he analytically derives the central market's common factor weight as 0.60. He therefore concludes that common factor components are biased and misleading when some trading venues cross orders at much tighter spreads than others. To investigate the pivotal importance of the extreme assumption of zero satellite half-spreads to Hasbrouck's conclusion, in the next section, we vary the satellite market maker's c_R between 1.0 (equal to c_C) and a tiny 0.005 half-spread. In addition, we investigate the role of stochastic process assumptions in price discovery research by generalizing the level of the signal to noise ratio in $\mathscr{P}_t$ to allow $1 \geqslant \lambda \geqslant 0$ and $8 \geqslant \sigma_w^2 \geqslant 0$.

4. Testing parameter estimates in common factor models

Table 1 summarizes our test results on ten variants of Hasbrouck's most complete Price Discovery Example III. To confirm the time-series properties of each pair of simulated price series, we first tested the order of integration, optimal system lag length, and cointegration using the augmented Dickey-Fuller and Johansen tests. All series pairs were cointegrated $C(1,1)$ at 99%. Using the SAS subroutine TSULMAR, optimal system lag lengths listed in column 5 minimize the AIC for the various pairs of VAR equations and generally mirror the one-period lag structure of the actual specifications. The last four columns of Table 1 display our findings from the same GG (1995) reduced-rank regressions, eigenvector computations, and Q_{GG} testing procedures that we employed on our ISSM sample data in HMW (2002). The parameter estimates and test statistics are averages of fifty simulation runs for each case A through K.

In Model A where λ, σ_w, c_C, and c_R all equal 1, we estimated the central market's common factor component to be 0.91. Although this estimate is 9% lower than the true $\gamma_{\perp C}$ value 1.0, our central point emerges in the last three columns where we report test statistics and p-values for Gonzalo and Granger's (1995) χ^2 statistic Q_{GG}. A 0.422 test statistic and p-value of 0.52 fails to reject the null hypothesis of the central market as the sole source of the common stochastic trend. Indeed, in every case in Table 1, the GG permanent–transitory decomposition and common factor approach do in fact recover the true information structure $[\gamma_{\perp C}\ \ \gamma_{\perp R}] = [1\ \ 0]$ specified by Hasbrouck's (2002) parameterized structural model.

Consider first three subset variants of Model A. Model B reduces the baseline model by setting $\lambda = 0$ thereby simplifying the information structure by removing the direction of the last trade on the central market as a source of volatility in the evolution of the implicit efficient price. Still, the direction of trade indicator variable $q_{C,t}$ influences the stochastic processes in the VECM, which for Model B can be written (using (I)) as,

$$\Delta P_{C,t} \equiv w_t + \Delta q_{C,t} = \sum_{j=1}^{2} \sum_{s=1}^{S} \beta_{j,t-s} \Delta P_{j,t-s} + \gamma_C (w_{t-1} + q_{C,t-1} - q_{R,t-1}) \tag{II}$$

and

$$\Delta P_{R,t} \equiv [w_{t-1}] + \Delta q_{R,t} = \sum_{j=1}^{2} \sum_{s=1}^{S} \beta_{j,t-s} \Delta P_{j,t-s} + \gamma_R ([w_{t-1} + q_{C,t-1}] - q_{R,t-1}). \tag{III}$$

To deduce the *one-way price adjustment null hypothesis* for this restricted model, we examine the correlation between the l.h.s. and the error-correction term on the r.h.s. in parentheses. With Hasbrouck's prior assumptions, $\gamma_C = 0$ in Eq. (II), and $\gamma_R = 1$ in Eq. (III). Consequently, $\gamma' = [0\ \ 1]$, and the *one-way price adjustment null hypothesis* is again H_0: $[\gamma_{\perp C}\ \ \gamma_{\perp R}] = [1\ \ 0]$. Here, we find a $\gamma_{\perp C}$ parameter estimate of 0.92, $Q_{GG} = 0.365$, and the p-value 0.54 again fails to reject H_0.

One can garner insights from other subsets of Model A by selectively zeroing out each of the other parameters in Table 1. For example, Model C assumes that information arrivals about common value are zero (since both $E(w)$ and $\sigma_w^2 = 0$) but that noise about private values expressed through the direction-of-trade random indicator variables q_C and q_R continues to drive the stochastic processes. Using similar reasoning involving the VECM, it can be shown that the testable *one-way price adjustment null hypothesis* is again H_0: $[\gamma_{\perp C}\ \ \gamma_{\perp R}] = [1\ \ 0]$. In this case the estimated $\gamma_{\perp C}$ is 0.95, and the 0.112 test statistic again fails to reject the null. Because Model C represents one-way adjustment to noise on the central market rather than to information, we call this "mimicking behavior" in HMW (2002).

Similarly, assuming $c_R = 0$ in Model D—i.e., a zero half-spread as the satellite market crosses trades at the central market's quote midpoint – results in an estimated $\gamma'_\perp = [0.59\ \ 0.41]$, consistent with Hasbrouck (2002). However, importantly, the $Q_{GG} = 3.176$ test statistic fails to reject the null hypothesis of H_0: $[\gamma_{\perp C}\ \ \gamma_{\perp R}] = [1\ \ 0]$ since the p-value of 0.08 is above the 0.01 and 0.05 levels conventional in microstructure research. In sum, the Gonzalo and Granger (1995) test procedure does in fact recover the true information structure specified in Hasbrouck's structural model.

In Models E-G, we analyze the magnitude of the downward bias on $\gamma_{\perp C}$ (as the satellite market half spread declines) that Hasbrouck (2002) first detected. With $c_R = \frac{1}{2} c_C$ in Model E, the estimated $\gamma_{\perp C}$ is 0.76. With $c_R = 1/20$th of c_C in Model F, the estimated $\gamma_{\perp C}$ is 0.63. Finally, with a c_R of just 0.005 – i.e., 200 times smaller than the central market half-spread, the estimated $\gamma_{\perp C}$ is 0.59. In all three cases, the p-values for rejecting the null hypothesis of $\gamma_{\perp C} = 1$ are well above the 0.01 or 0.05 levels. Thus, contrary to Hasbrouck (2002), substantial differences in spreads between the central and satellite markets do not pose a problem for drawing correct price discovery inferences from estimates of the Gonzalo-Granger common factor weights.

5. A clarification on the information volatility parameter

The VECM in (I) also provides an analytical framework for clarifying how the standard normal w_t information arrivals interact with the $q_{C,t}$ and $q_{R,t}$ direction-of-trade disturbances that we interpret as noise about private values. Scaling up the variance of w_t by a factor of 8 to reflect the enhanced information volatility in fast markets (see Model H) results in a $\gamma_{\perp C}$ parameter estimate of 0.84, statistically insignificantly different from 1.0 (with p-value 0.73). Similarly scaling up the variance of w_t while the spread differential remains at $c_C = 1$ and $c_R = 0.005$ (see Model I) results in a much increased $\gamma_{\perp C} = 0.83$ relative to the 0.59 estimate for comparable slow markets (compare Model G). Even a *zero* half-spread in the satellite market (see Model J) results in an increased estimated $\gamma_{\perp C} = 0.82$ in these fast markets. Again, the 0.165 and 0.072 Q_{GG} test statistics for Models I and J (p-value 0.69 and 0.79, respectively) fail to reject the true null H_0 that $\gamma_{\perp C} = 1.0$. Finally, since the VECM that includes only common value information arrivals as a source of volatility in P_t is sufficient, as we saw earlier in Equations II and III for Model B, to derive the null hypothesis $[\gamma_{\perp C} \;\; \gamma_{\perp R}] = [1 \;\; 0]$, we investigated Model K where again information volatility is great but where $\lambda = 0$. Here too the common factor results ($\gamma_{\perp C} = 0.96$, $Q_{GG} = 0.038$ and p-value $= 0.84$) exactly recover the information structure of Hasbrouck's specifications.

6. Conclusion

Hasbrouck's (2002) insightful parameterization of the stylized facts relating satellite markets and the NYSE has stimulated several new directions for microstructure research on asymmetric half-spreads, determinants of volatility in the implicit efficient price, and on the role of stale quotes. By generalizing and testing a one-way price adjustment hypothesis in a VECM framework, we show that Gonzalo and Granger's (1995) common factor weights and Q_{GG} test statistics provide an attractive method for recovering the information structure of these asset pricing models.

References

Bertsimas, D., Lo, A.W., 1998. Optimal control of execution costs. Journal of Financial Markets 1, 1–50.

deJong, 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., 2000. A regime-level empirical model of the specialist quote revision process. Review of Quantitative Finance and Accounting 14, 399–417.

Harris, F.H.deB., McInish, T.H., Shoesmith, G., Wood, R.A., 1995a. Cointegration, error correction and price discovery on informationally-linked security markets. Journal of Financial and Quantitative Analysis 30, 563–579.

Harris, F.H.deB., McInish, T.H., Chakravarty, R., 1995b. Bids and asks in disequilibrium market microstructure. Journal of Banking and Finance 19, 323–345.

Harris, F.H.deB., McInish, T.H., Wood, R.A., 1997. Common long-memory components for intraday stock prices: A measure of price discovery. Working paper, Babcock School, Wake Forest University, Winston-Salam, NC.

Harris, F.H.deB., McInish, T.H., Wood, R.A., 2002. Security price adjustment across exchanges: an investigation of common factor components for Dow stocks. 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., 2002. Stalking the "efficient price" in market microstructure specifications: an overview. Journal of Financial Markets, this issue.