Price Discovery in High Resolution

by

Joel Hasbrouck

Discussion

by

Fulvio Corsi
University of Pisa

and

Giuseppe Buccheri, Scuola Normale Superiore di Pisa
Giacomo Bormetti, University of Bologna
Fabrizio Lillo, University of Bologna

SoFiE Lugano, 13 June 2018

Quick summary

General Goal

Study the incorporation of information into prices across markets and timescales

Framework

Hasbrouck's (1995) pioneering work who introduces:

- VECM for prices of the same asset formed in different markets
- Information Share (IS) as measure of price discovery of different markets

Main issue still open in the literature: allocation of correlation terms $\rightarrow$ IS bounds

Approach of the paper

- Shrinking IS bounds by increasing resolution of obs up to microseconds,
- Problem: Huge n$^\circ$ of lags in VECM ($\sim 10^6$)
- Solution: Retain parsimony with "extreme" versions of HAR structure

Main results

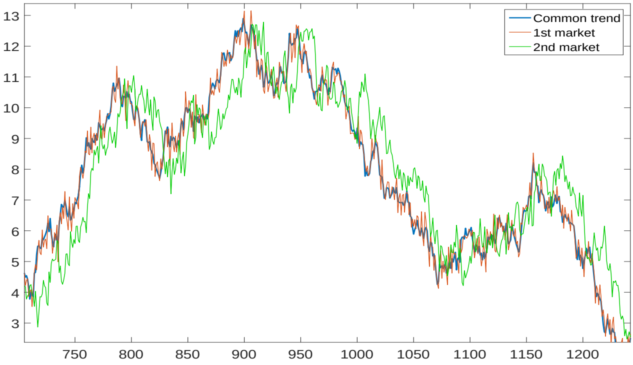
- Successfully collapsing the IS bounds as resolution increases
- Strong information dominance of direct subscribers over consolidated tape
- Slight information dominance of listing exchanges over other exchanges
- Strong information dominance of quotes over trades and lit over dark trades

Ad Hoc Simulation I: VECM with one Lag

- We simulate a VECM for two market prices with 1 lag:

$$\Delta p_t = \gamma B p_{t-1} + \phi_1 \Delta p_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \Omega), \qquad \Omega \text{ is diagonal}$$

- With a chioce of ϕ_1 that induces a sizable lag for the 2^{nd} market



IS 1^{st} market = [0.486, 0.483],

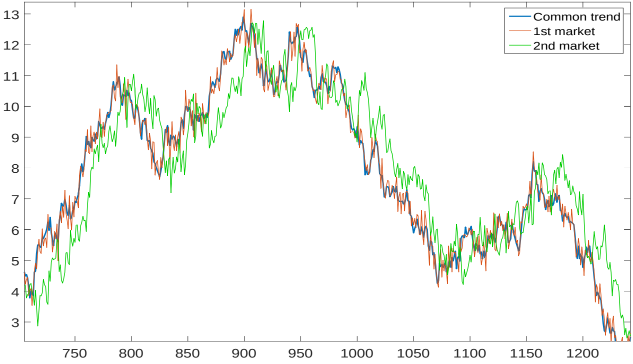
IS 2^{nd} market = [0.513, 0.516],

Ad Hoc Simulation I: VECM with one Lag

- We simulate a VECM for two market prices with 1 lag:

$$\Delta p_t = \gamma B p_{t-1} + \phi_1 \Delta p_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \Omega), \qquad \Omega \text{ is diagonal}$$

- With a chioce of ϕ_1 that induces a sizable lag for the 2^{nd} market



IS 1^{st} market = [0.486, 0.483],

IS 2^{nd} market = [0.513, 0.516],

IS and Autoregressive Lags ϕ_i

- Using the common trends representation

$$p_t = \theta(1)_* \left(\sum_{s=1}^{t} \epsilon_s \right) \iota + \theta^*(L)\epsilon_t \equiv m_t \iota + s_t$$

with $\theta(1)_*$ a generic row of $\theta(1)$ and ι a vector of ones.

- The Information Share (IS) is defined as

$$IS_i = \frac{\theta(1)^2_{*,i} \;\; \Omega_{i,i}}{\theta(1)_* \;\; \Omega \;\; \theta(1)'_*}$$

with $\theta(1)_{*,i}$ the *i-th* element of $\theta(1)_*$

- But, from the Granger's representation theorem we can analytically compute

$$\theta(1) = B_\perp \underbrace{\left[\gamma'_\perp \left(I_N - \sum_{i=1}^{k} \phi_i \right) B_\perp \right]^{-1}}_{\substack{(N-r)\times(N-r) \\ if \;\; r = N-1 \;\Rightarrow\; \text{scalar}}} \gamma'_\perp$$

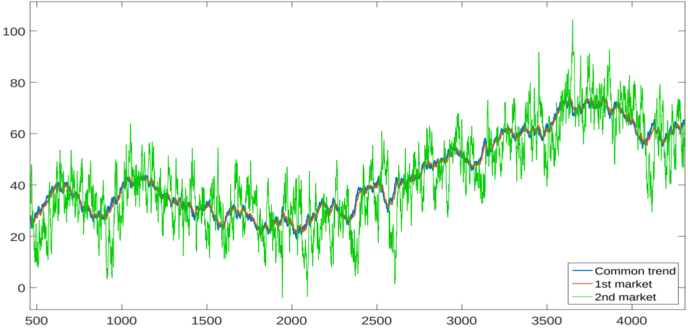
- Hence, the term containing the autoregressive lags ϕ_i cancel out in the IS ratio
$\Rightarrow$ IS depends on γ and Ω, but not DIRECTLY on ϕ's.

Ad Hoc Simulation II: VECM with Different Variances

- We simulate a VECM for two markets (with no lags):

$$\Delta p_t = \gamma B p_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \Omega), \qquad \Omega \text{ is diagonal}$$

  - 1^{st} market is fast adapting (γ_1 small) and with small innovation variance
  - 2^{nd} market is slower (γ_2 larger) and with substantial innovation variance



IS 1^{st} market = [0.380, 0.389],

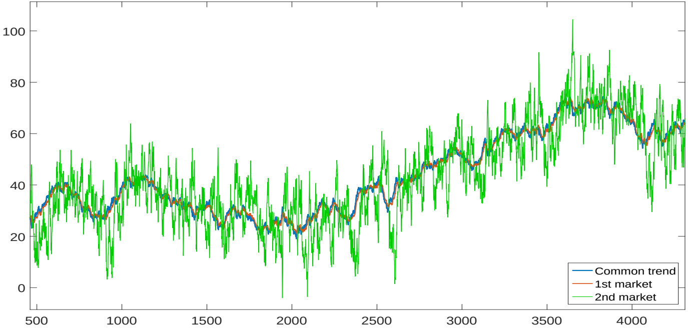
IS 2^{nd} market = [0.619, 0.610].

Ad Hoc Simulation II: VECM with Different Variances

- We simulate a VECM for two markets (with no lags):

$$\Delta p_t = \gamma B p_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \Omega), \qquad \Omega \text{ is diagonal}$$

  - 1^{st} market is fast adapting (γ_1 small) and with small innovation variance
  - 2^{nd} market is slower (γ_2 larger) and with substantial innovation variance



IS 1^{st} market = [0.380, 0.389],

IS 2^{nd} market = [0.619, 0.610],

IS and Market Variances $\Omega_{i,i}$

- In the simulations with a reduced form VECM, the IS of a market depends:
  - Positively on the speed of adjustment of the market:

    $$\text{small } \gamma_i \Rightarrow \text{high } IS_i$$

  - Positively on the noise variance of the market:

    $$\text{large } \Omega_{i,i} \Rightarrow \text{high } IS_i$$

- $\Rightarrow$ IS of a mkt with sufficiently large $\Omega_{i,i}$ can be larger than a mkt with a small γ_j (Update: see also the presentation here at SoFiE pre-conf of Christian Nguenang)
- Not clear if the reduced form model VECM contains MM noise or not (see Dias, Fernandes, Scherrer 2018)
- Difficulties of interpretation might stem from the reduced form nature of VECM
- Look for a structural model (as in De Jong & Schotman 2010) but with a lagged adjustment mechanism
- We consider a multivariate generalization of the partial price adjustment model of Hasbrouck & Ho (1987)

Structural Model with Lagged Adjustments

- In the proposed model the efficient price m_t is the scalar process

$$m_t = m_{t-1} + u_t, \qquad u_t \sim \text{NID}(0, \Sigma)$$

- The latent prices with lagged adjustment X_t is assumed to be

$$X_t = X_{t-1} + \Psi(m_t \,\iota - X_{t-1})$$

  the matrix Ψ is the speed of adjustment matrix.

  If $\Psi = \mathbb{I}$, then $X_{i,t} = m_t$, i.e. instantaneous price adjustment in every market

- Finally, the vector of observed log-prices is:

$$P_t = X_t + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, H)$$

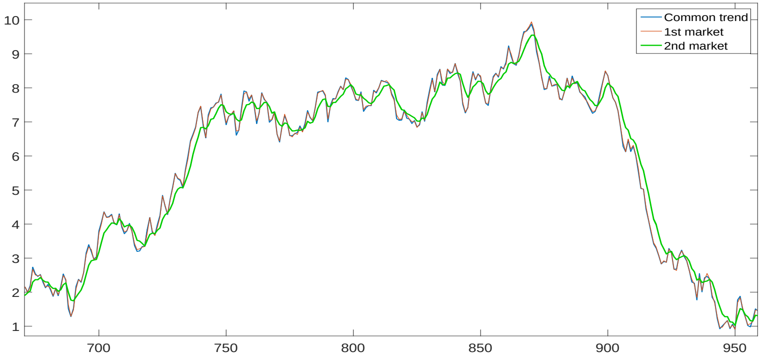
- This model can be conveniently estimated with the standard Kalman filter (treating zeros as missing values) assuming (identification assumption) $u_t \perp \varepsilon_t$

Structural Model with Partial Price Adjustment

- We simulate the structural model for two markets:

$$P_t = X_{t-1} + \Psi(m_t \iota - X_{t-1}) + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Omega), \qquad \Omega \text{ is diagonal}$$

  - 1^{st} market is fast adapting ($\Psi_{1,1} = 0.9$)
  - 2^{nd} market is slower ($\Psi_{2,2} = 0.3$)



IS 1^{st} market = [0.997, 0.087]

IS 2^{nd} market = [0.912, 0.002]

Concluding Remarks

Some additional questions:

- What are the effects of large number of zero returns on IS?
- What is the effect of the strong one-second periodicity observed in UHF data?

What I liked about the paper:

- Clean and effective solution to the problem of non-uniqueness of IS
- Heterogeneous/Multifrequency view of the market
- Unified approach to modeling price dynamics across a full range of frequencies
- Consistency of results across resolutions
- Emergence of clearer pictures as resolution increases.