# DISCUSSION PAPER SERIES

No. 3987

**PRICE DISCOVERY IN FRAGMENTED MARKETS**

Frank de Jong and Peter Schotman

***FINANCIAL ECONOMICS***

**Centre for Economic Policy Research**

**www.cepr.org**

**Available online at:** **www.cepr.org/pubs/dps/DP3987.asp**

*ISSN 0265-8003*

# PRICE DISCOVERY IN FRAGMENTED MARKETS

**Frank de Jong,** Universiteit van Amsterdam and CEPR
**Peter Schotman,** LIFE, Maastricht University and CEPR

Discussion Paper No. 3987
July 2003

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3987

July 2003

# ABSTRACT

## Price Discovery in Fragmented Markets*

This Paper proposes a structural time series model for the intra-day price dynamics of fragmented financial markets. We generalize the structural model of Hasbrouck (1993) to a multivariate setting. We discuss identification issues and propose a new measure for the contribution of each market to price discovery. We illustrate the model by an empirical example using Nasdaq dealer quotes.

JEL Classification: C32 and F31
Keywords: high-frequency data, microstructure and structural time series models

Frank de Jong
Department of Financial Management
Universiteit van Amsterdam
Roetersstraat 11
1018 WB, Amsterdam
THE NETHERLANDS
Tel: (31 20) 525 5815
Fax: (31 20) 525 5285
Email: fdejong@fee.uva.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=115389

Peter Schotman
Maastricht University
LIFE
PO Box 616
6200 MD Maastricht
THE NETHERLANDS
Tel: (31 43) 388 3838
Fax: (31 43) 388 4875
Email: p.schotman@berfin.unimaas.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=117320

*We would like to thank Bruce Lehmann and seminar participants at University of Amsterdam, Stockholm School of Economics, University of Uppsala and Sveriges Riksbanken for helpful comments on an earlier draft. As always, all errors are our own.

Submitted 02 July 2003

# 1 Introduction

The markets in many financial assets are fragmented. To give a few examples, NYSE listed US stocks are often also traded on regional exchanges; many European stocks are cross-listed on the NYSE or Nasdaq; on Nasdaq itself and in the foreign exchange and bond markets there are multiple dealers and the markets for the trading between dealers and their clients is quite separated from the inter-dealer market. Starting with Hasbrouck's (1995) pioneering work, the modeling of microstructure data from such fragmented markets has received considerable attention in the financial literature. This literature was recently surveyed in an issue of the *Journal of Financial Markets* (2002).

The purpose of price discovery models is to describe the dynamic interactions between the quotes or transaction prices from two or more markets, or from two or more dealers of the same asset.[1] The most natural model for quotes $p_{it}$ by dealer $i$ (or prices on market $i$) is that they equal the fundamental value of the asset, $p_t^*$, plus a transitory term:[2]

$$p_{it} = p_t^* + u_{it}. \tag{1}$$

This equation is in the form of an unobserved components model, or a structural time series model in the terminology of Harvey (1989). Prices are observed, but the efficient price $p_t^*$ is not. The fundamental value is a random walk, whereas the dealer dependent transitory term $u_{it}$ is stationary and typically close to white noise. The price changes, $\Delta p_{it}$ therefore have a very typical serial correlation pattern: a strong and negative first order autocorrelation, and small and often negligible higher order autocorrelations.

Despite its intuitive appeal, the unobserved components model is rarely used in empirical work, neither for estimation nor for the definition of measures of price discovery. The standard time series model proposed by Hasbrouck (1995) is the Vector AutoRegression introduced by Sims (1980) in macroeconomics. Since all price series share the same long term (random walk) component, the VAR is subject to cointegration restrictions and estimated as a vector error correction model (VECM). The central quantity of interest is the *information share*, which measures the relative importance of each market in the price discovery process. Hasbrouck defines the infor-

[1] See Hasbrouck (1995) for an example with multiple markets. Huang (2002) is a recent application to multiple dealers.

[2] See e.g. Hasbrouck (1993), Zhou (1996), and Lehmann (2002)

mation share as the fraction of the variance of the random walk component that can be attributed to a particular market (or dealer). The VECM and information share methodology has been applied in many empirical studies, for example by Hasbrouck (1995) and Harris et al. (2002) for US equities traded on the NYSE and regional exchanges; Hupperets and Menkveld (2001) for European equities cross listed in the US; Upper and Werner (2002) for the relation between the cash and futures market in German government bonds; De Jong, Mahieu and Schotman (1998) and Covrig and Melvin (2002) for the foreign exchange market.

In this paper, we revisit the unobserved components model. The information flow is modeled through the simultaneous and lagged covariances between the 'noise' terms in (1) and the innovations in the fundamental values. Working directly within the structural time series model has advantages over the VECM approach in settings with many markets or many dealers. The parsimony of the unobserved components model has advantages both for the statistical inference as well as the definition of information shares. Parsimony is usually not seen as a major concern with high frequency microstructure data. Yet both the VAR approach as well as the definition and measurement of the information share have several drawbacks in applications with many markets (dealers).

First, the particular pattern of autocorrelations in prices (or quotes) is difficult to describe with low order autoregressive models. Autoregressions often require long lags to capture a strong first order autocorrelation but a second autocorrelation that is almost zero. The VECM also suffers from lack of parsimony in the error correction part. In a model with $N$ dealers, the cointegration restrictions lead to $N-1$ different error correction terms in each of the $N$ equations. Related to this is a potential problem with the data. Although microstructure time series have many observations, we do not always have that many observations for *all* markets (dealers). The NYSE is much more active than its regional satellite markets. Foreign exchange dealers are often at a few large banks. Most Nasdaq quotes are issued by a handful of dealers and Electronic Communication Networks (ECN). In these circumstances the time series for a multivariate model of dynamic interactions is sampled at the pace of the slowest market (Harris et al., 2002) or with relatively long fixed calendar intervals. This problem is particularly serious for large dimensional systems, i.e. a setting with multiple markets. When the number of dealers increases, the number of simultaneously available observations generally decreases, but the number of parameters in a VAR increases quadratically with the number of time series.

The other problem with the VAR model is in the information shares. These are not uniquely defined, but depend on the allocation of the covariance terms in the error covariance matrix. Hasbrouck suggests to report upper and lower bounds, obtained by different ordering of the markets. For a two variable system these bounds are sometimes fairly narrow, but there are also applications (for example, Melvin and Covrig, 2002) where the bounds are very wide. In a high dimensional system the number of off-diagonal elements in the covariance matrix increases quadratically in $N$, and will eventually dominate the variance decomposition, so that it is difficult to obtain meaningful estimates of the information shares.

The unobserved components model is appealing in these situations. Within this model, we introduce a new measure of the contribution to price discovery. Unlike the traditional information share, which is defined within a reduced form time series model, the new measure is defined directly within the structural time series model and will remain meaningful in high dimensional settings.

The unobserved components model has a drawback of its own. Since it contains the efficient price as a latent variable, there is an inherent identification problem. For the univariate version of the model this identification problem is discussed in depth in Hasbrouck (1993). In the multivariate version of (1), that is of interest for price discovery models, the identification problem turns out to be less severe. Full identification is achieved under plausible assumptions regarding the dealer noise $u_{it}$.

The structure of this paper is as follows. First, we provide a theoretical investigation of the properties of the structural price discovery model. We compare the implications of this model for price dynamics and information shares with the usual VECM approach. Next, we present our alternative measure for the contribution to price discovery. We end with an empirical illustration of the structural approach using Nasdaq multiple dealer quotes.

# 2 A structural time series model

This section explores a structural time series model for market microstructure and price discovery in fragmented markets. The model generalizes the univariate model of Hasbrouck (1993) to a multiple market setting. This section first reviews the results for a univariate pure random walk plus noise model. Then the model is extended to a multivariate random walk plus noise. Finally higher order dynamics are introduced.

## 2.1 Univariate model

Hasbrouck (1993) considers the univariate structural model for $p_t$, the logarithm of the price of a security,

$$\begin{aligned} p_t &= p_t^* + u_t, && \\ p_t^* &= p_{t-1}^* + r_t, && \mathrm{Var}(r_t) = \sigma^2, \\ u_t &= \alpha r_t + e_t, && \mathrm{Var}(e_t) = \omega^2, \end{aligned} \tag{2}$$

where $p_t^*$ is the unobserved efficient price (random walk) and $u_t$ a transitory component. The shocks $e_t$ and $r_t$ are uncorrelated. The coefficient $\alpha$ determines the covariance between transitory and permanent shocks: $\mathrm{Cov}(u_t, r_t) = \alpha\sigma^2$.

We can write the price changes (returns) in this model as

$$\Delta p_t = r_t + \Delta u_t = (1+\alpha)r_t - \alpha r_{t-1} + \Delta e_t \ . \tag{3}$$

The auto-covariances of returns implied by this model are therefore

$$\gamma_0 = \mathsf{E}[\Delta p_t^2] = \sigma^2 \left((1+\alpha)^2 + \alpha^2\right) + 2\omega^2 \tag{4a}$$

$$\gamma_1 = \mathsf{E}[\Delta p_t \Delta p_{t-1}] = -\sigma^2\alpha(1+\alpha) - \omega^2. \tag{4b}$$

All higher order covariances are zero, and therefore the reduced form of the structural model is a first order Moving Average process in the price changes.

From the moment equations, the parameter $\sigma^2$ is uniquely identified as

$$\sigma^2 = \gamma_0 + 2\gamma_1. \tag{5}$$

The parameters $\alpha$ and $\omega^2$ cannot be identified separately. Hence, some identifying restriction is necessary. Hasbrouck (1993) suggests two identifying restrictions: the Beveridge-Nelson (BN) normalization ($\omega^2 = 0$) and the Watson normalization ($\alpha = 0$).

In order to better understand the differences between these normalizations, we first define a range of admissible values for $\alpha$. From the moment conditions we obtain

$$\begin{aligned} \omega^2 &= -\gamma_1 - \alpha(1+\alpha)\sigma^2 \\ &= -\gamma_0\left(\rho_1 + \alpha(1+\alpha)(1+2\rho_1)\right), \end{aligned} \tag{6}$$

where $\rho_1 = \gamma_1/\gamma_0$ is the first order autocorrelation. For microstructure data, the first order autocorrelation is typically negative, but bigger than $-0.5$. Therefore, we assume that $-\frac{1}{2} < \rho_1 \leq 0$. For the interpretation of the model $\omega^2$ must remain positive. This provides a bound on the admissible values of $\alpha$. Equation (6) implies the inequality

$$-\sqrt{1-2\rho_1} \leq (2\alpha+1)\sqrt{1+2\rho_1} \leq \sqrt{1-2\rho_1}. \tag{7}$$

These intervals typically contain both positive and negative values for $\alpha$. Boundary cases are $\rho_1 \to -\frac{1}{2}$, in which case $\alpha$ is not restricted at all, and $\rho_1 = 0$, in which case $-1 \leq \alpha \leq 0$. For a typical first order autocorrelation $\rho_1 = -0.3$, we find the interval $-1\frac{1}{2} \leq \alpha \leq \frac{1}{2}$.

Notice that with a negative first order autocorrelation, the value $\alpha = 0$ is always admissible, but other values for $\alpha$ are also possible. Morley, Zivot and Nelson (2002) study the identification of $\alpha$ in a model with positive first order autocorrelation, which is typical for macro-economic data. In that case, the range of admissible $\alpha$ may not contain zero, and the Watson restriction is not feasible. But since the first order autocorrelation for microstructure return data is almost always negative, the Watson restriction is always feasible for typical microstructure data.

Using (6) we can write the variance of the idiosyncratic term as

$$\begin{aligned} \mathrm{Var}(u_t) &= \alpha^2\sigma^2 + \omega^2 \\ &= -\gamma_0\left(\rho_1 + \alpha(1+2\rho_1)\right) \end{aligned} \tag{8}$$

Since $\gamma_0 > 0$ and $-\frac{1}{2} < \rho_1 \leq 0$, the noise variance attains a lower bound when $\alpha$ is at its maximum value. This occurs when the right inequality in (7) holds as an equality and corresponds to the BN normalization. Hasbrouck shows that the choice of normalization for $\alpha$ may have an important effect on the estimate $\mathrm{Var}(u_t)$ in empirical applications. This completes the summary of Hasbrouck's (1993) model. We now turn to a multivariate generalization of his model.

## 2.2 Multivariate model

Let $p_t$ now be a vector of $N$ prices for the same asset from different markets. The multivariate model reads,

$$\begin{aligned} p_t &= \iota p_t^* + u_t, & & \\ p_t^* &= p_{t-1}^* + r_t, & \mathrm{Var}(r_t) &= \sigma^2, \\ u_t &= \alpha r_t + e_t, & \mathrm{Cov}(r_t,\, e_t) &= 0, \mathrm{Var}(e_t) = \Omega, \end{aligned} \tag{9}$$

where $\alpha$ is a vector, $\iota$ is a vector of ones, and $\Omega$ a $(N \times N)$ matrix. Again, $\mathrm{Cov}(u_t,\, r_t) = \alpha\sigma^2$. As in the univariate model, the innovations in the efficient price and the transitory term may be correlated. By construction, all price series share the same random walk component and are therefore cointegrated.

The price changes (returns) in this model are written as

$$\Delta p_t = \iota r_t + \Delta u_t = (\iota + \alpha) r_t - \alpha r_{t-1} + \Delta e_t \tag{10}$$

and the serial covariances are

$$\begin{aligned} \Gamma_0 &= \mathsf{E}[\Delta p_t \Delta p_t'] = \sigma^2 \left( (\iota + \alpha)(\iota + \alpha)' + \alpha \alpha' \right) + 2\Omega & \text{(11a)} \\ \Gamma_1 &= \mathsf{E}[\Delta p_t \Delta p_{t-1}'] = -\sigma^2 \alpha (\iota + \alpha)' - \Omega & \text{(11b)} \end{aligned}$$

The remainder of this section considers identification of $\alpha$ in the multivariate model. First, the sum of lead, current and lag covariances,

$$\Gamma_1' + \Gamma_0 + \Gamma_1 = \sigma^2 \iota \iota' \tag{12}$$

(over-)identifies the variance of the efficient price innovation. Next consider the difference between lead and lag cross-covariances

$$\Gamma_1 - \Gamma_1' = \sigma^2 \left( \iota \alpha' - \alpha \iota' \right). \tag{13}$$

From this, $\alpha$ can be identified up to a translation along $\iota$. Finally, given values for $\sigma^2$ and $\alpha$, the noise covariance matrix $\Omega$ can be identified from equation (11a), or from the sum of the lead and lag covariances

$$\Gamma_1 + \Gamma_1' = -\sigma^2 \left( \alpha \iota' + \iota \alpha' + 2\alpha \alpha' \right) - 2\Omega \tag{14}$$

All parameters in this model are (over)identified, except the vector $\alpha$, which is only identified up to a translation along the unit vector. Let $w$ be a scalar. The entire set of equivalent solutions is characterized by

$$\begin{aligned} \alpha &= \alpha^* - w\iota, & \text{(15a)} \\ \Omega &= \Omega^* + w\sigma^2 \left( (1-w)\iota\iota' + \iota \alpha^{*\prime} + \alpha^* \iota' \right), & \text{(15b)} \end{aligned}$$

where $\alpha^*$ and $\Omega^*$ constitute an admissable solution. Since $\Omega$ is a covariance matrix, it must be positive definite. Therefore not all values for $w$ are admissable, analogous to the univariate case.

For a general $\Omega$, one element (or linear combination of elements) of $\alpha$ needs to be fixed as a normalizing restriction. The identification for $\alpha$ is such that given one $\alpha_i$, the others can be determined from moment conditions derived from equation (13).

The remaining elements of $\alpha$ are even overidentified, as we can use any column of $\Gamma_1 - \Gamma_1'$ to estimate $\alpha$. So, essentially we only need bivariate information (the covariances between an arbitrary $\Delta p_{it}$ and the vector $\Delta p_t$) to estimate the vector $\alpha$. In empirical applications we could use the covariances of individual markets with the central market or a market index to estimate all $\alpha_i$'s (up to a constant).

Various restrictions will lead to full identification of $\alpha$. Analogous the univariate model we first consider the Beveridge-Nelson and Watson restrictions. The BN representation is obtained from the reduced form. The reduced form of the multivariate random walk plus noise model is the first order vector moving average (VMA) process,

$$\Delta p_t = \epsilon_t - C\epsilon_{t-1}, \qquad \mathrm{Var}(\epsilon_t) = \Sigma, \tag{16}$$

where cointegration requires that

$$C = I - \iota\theta' \tag{17}$$

for some vector $\theta$. The common trends representation of the reduced form is

$$p_t = \iota\tilde{p}_t + (I - \iota\theta')\epsilon_t \tag{18a}$$

$$\tilde{p}_t = \tilde{p}_{t-1} + \theta'\epsilon_t. \tag{18b}$$

Under the Beveridge-Nelson restriction, the innovations in the permanent component are equal to an exact linear combination of the VMA innovations: $r_t = \theta'\epsilon_t$. With this structure, we can write

$$\mathrm{Cov}(\Delta p_t, r_t) = \Sigma\theta = \sigma^2(\iota + \alpha), \tag{19}$$

where the last equality follows from (10). This gives a particular choice for $\alpha$, that we shall call the BN value,

$$\alpha_{BN} = \Sigma\theta/\sigma^2 - \iota. \tag{20}$$

For the BN normalization the covariance matrix of $e_t$ is semi-definite

$$\Omega_{BN} = \Sigma - \frac{\Sigma\theta\theta'\Sigma}{\theta'\Sigma\theta}. \tag{21}$$

The BN value of $\alpha$ is the maximal value, since for values $w < 0$, the implied covariance matrix $\Omega$ is not positive semi-definite any more.[3] The Beveridge-Nelson value $\alpha_{BN}$ is always feasible, irrespective of the empirical serial correlation patterns. The range of

[3] In the appendix we provide a proof of this statement.

alternative equivalent combinations of $\alpha$ and $\Omega$ in the multivariate is smaller than in the univariate model. For each price series the univariate restrictions must hold for the diagonal element $\omega_{ii}$ and they must hold jointly. In addition positive definiteness for $\Omega$ is stronger than just positive diagonal elements.

A generalization of the Watson restriction is found by setting $\alpha_i = 0$ for some markets, or more generally $\alpha'\pi = 0$, where $\pi'\iota = 1$. In this case a weighted average of the different price series is unrelated to the change in the efficient price. Imposing the Watson restriction $\alpha_i = 0$ on *every* market, however, leads to $N - 1$ overidentifying restrictions, which may be violated by the data. The interpretation of the Watson restriction is that one market is designated as the central market. In some applications there is a natural choice for the central market. For example, when studying the relation between the NYSE and regional markets in the US, the NYSE would be the central market. As another example, in an application with cross-listed stocks, the home market is the candidate central market. Setting some arbitrary $\alpha_i = 0$ could easily be inadmissable because it will violate the condition that $\Omega$ must be positive definite. Admissability must be checked on a case by case basis and will restrict the potential normalizations of $\alpha$.

A third way to identify $\alpha$ is by imposing that $\Omega$ is diagonal. Under that assumption the deviations between quotes and the efficient price, $p_{it} - p_t^*$, will only be correlated among dealers because of their joint dependence on the innovation in the efficient price $r_t$. The innovations in the quotes would than have a single factor structure with factor loadings $\alpha$. The factor structure imposes overidentifying restrictions on the reduced form that fully identify $\alpha$. Uncorrelated dealer shocks is a natural assumption in UC models.

Other identification restrictions work through higher order lags. These models will be discussed in the next subsection.

## 2.3 Higher order models

In practice, microstructure data show second order and sometimes even higher order serial covariances. Another empirical motivation for the need to model higher order correlations is given by the correlation pattern in the differences (spreads) between dealer quotes

$$s_{ij,t} = p_{it} - p_{jt} = u_{it} - u_{jt}. \tag{22}$$

In practice, these differences show positive serial correlation, which can be captured by a richer serial correlation structure of the transitory terms. We discuss two ways to model the higher order correlations.

### 2.3.1 Autocorrelation in dealer shocks

The first way to model higher order dynamics is by assuming that the noise term $e_t$ is serially correlated. Looking at the simplest case, the specification for the dealer behavior becomes

$$u_t = \alpha r_t + e_t + \Psi e_{t-1}, \tag{23}$$

with $\Psi$ an $(N \times N)$ matrix. The moment conditions become

$$\begin{aligned}
\Gamma_0 &= \mathsf{E}[\Delta p_t \Delta p_t'] = \sigma^2\left((\iota+\alpha)(\iota+\alpha)' + \alpha\alpha'\right) + \Omega + (\Psi - I)\Omega(\Psi - I)' + \Psi\Omega\Psi', \\
\Gamma_1 &= \mathsf{E}[\Delta p_t \Delta p_{t-1}'] = -\sigma^2\alpha(\iota+\alpha)' + (\Psi - I)\Omega - \Psi\Omega(\Psi - I)', \\
\Gamma_2 &= \mathsf{E}[\Delta p_t \Delta p_{t-2}'] = -\Psi\Omega.
\end{aligned} \tag{24}$$

The additional parameter matrix $\Psi$ is just identified from the second order autocovariance matrix $\Gamma_2$. All implications and identification results remain as in the first order case.

Adding further lags $\Psi_j e_{t-j}$ does not alter anything in the identification of $\alpha$. With more lags the model becomes increasingly more difficult to analyse, but $\alpha$ remains easily connected to the autocovariance structure. The result is given in the form of a theorem.

**Theorem 1** *Let prices be generated by the unobserved components model (9) but with dealer shocks*

$$u_t = \alpha r_t + \sum_{j=0}^{M} \Psi_j e_{t-j}, \tag{25}$$

*where* $\mathsf{E}[e_t r_s] = 0$ *for all* $t$ *and* $s$*. Then*

$$\sum_{i=1}^{M+1} i(\Gamma_i' - \Gamma_i) = \sigma^2(\alpha\iota' - \iota\alpha') \tag{26}$$

The proof is in the appendix. This provides an easy statistic to learn about the dealer responses to changes in the efficient price. In a bivariate setting it amounts to comparing lead and lag cross covariances of returns. The difference in $\alpha$ for two markets is

$$\alpha_i - \alpha_j = \frac{1}{\sigma^2} \sum_{\ell=1}^{M+1} \ell\left(\mathrm{Cov}(\Delta p_{jt}, \Delta p_{i,t-\ell}) - \mathrm{Cov}(\Delta p_{it}, \Delta p_{j,t-\ell})\right) \tag{27}$$

Market $i$ is more informative about the efficient price than market $j$ if lagged prices changes of market $i$ have more predictive power for the current price change in market $j$ than vice versa. The asymmetry in predictive power between markets $i$ and market $j$ is a determinant of the information contents of the prices.

### 2.3.2 Lagged effects of the efficient price

The second way of modeling higher order dynamics is by including lagged effects of the efficient price in the transitory term. With a single lag $r_{t-1}$ we write

$$u_t = \alpha r_t + \phi \Delta r_t + e_t, \tag{28}$$

with $\phi$ a vector, and of course $\mathrm{Cov}(e_t,\, \Delta r_t) = 0$. In this formulation $\alpha$ measures the long run effect of $r_t$ on the dealer quotes. The returns follow as

$$\Delta p_t = (\iota + \alpha + \phi) r_t - (2\phi + \alpha) r_{t-1} + \phi r_{t-2} + \Delta e_t. \tag{29}$$

The serial covariance matrices become

$$\begin{aligned}
\Gamma_0 &= \mathsf{E}[\Delta p_t \Delta p_t'] = \sigma^2 \left( (\iota + \alpha + \phi)(\iota + \alpha + \phi)' + (2\phi + \alpha)(2\phi + \alpha)' + \phi\phi' \right) + 2\Omega, \\
\Gamma_1 &= \mathsf{E}[\Delta p_t \Delta p_{t-1}'] = -\sigma^2 \left( (2\phi + \alpha)(\iota + \alpha + \phi)' - \phi(2\phi + \alpha)' \right) - \Omega, \\
\Gamma_2 &= \mathsf{E}[\Delta p_t \Delta p_{t-2}'] = \sigma^2 \phi (\iota + \alpha + \phi)'.
\end{aligned} \tag{30}$$

Since $\phi$ is now a vector, the second order covariances in $\Gamma_2$ are overidentified. This provides just enough additional structure for the full identification of $\alpha$. Notice however that this needs the full second order covariance matrix. If we use bivariate information only (i.e. only one row or column of $\Gamma_2$) we need to impose one additional restriction on $\alpha$, like in the first order serial correlation case.

Adding further lagged $\Delta r_{t-j}$ is straightforward, but does not provide further insights. As in the previous subsection the same sum of lag and lead covariances identifies relative values of $\alpha$. However, due to the restrictions on the higher order $\Gamma_j$ for $j > 1$, in this case all elements of $\alpha$ are also identified in an absolute sense.

# 3 Information shares

Information measures of price discovery summarise the relation between the change in the efficient price and actual price changes. The most common measure is due to Hasbrouck (1995), who defines information shares within a reduced form model. In this section we suggest a modification of this definition. Instead of the reduced form definition we define the information shares directly within the structural unobserved components model. We also suggest an alternative way to present the information measure in both the reduced form and unobserved components model.

Hasbrouck's (1995) information shares are defined within an infinite order vector moving average reduced form for an $N$-vector of prices,

$$\Delta p_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j}. \tag{31}$$

The shocks $\epsilon_t$ have zero mean and covariance matrix $\Sigma$. The coefficients $C_i$ are assumed to be absolutely summable, meaning that $\sum_{j=0}^{\infty} j|C_j|$ is bounded. As a normalization we set $C_0 = I$, the $(N \times N)$ identity matrix. Since all prices refer to the same security, the elements in $p_t$ cointegrate and every spread series $p_{it} - p_{jt}$ is stationary. This imposes structure on the persistence matrix $C(1)$, which must satisfy

$$C(1) = \sum_{i=0}^{\infty} C_i = \iota\theta', \tag{32}$$

where $\iota$ is a vector of ones, and $\theta$ an $N$-vector of parameters. Using the structure in (32) we can write the level of prices as the sum of a random walk component and a stationary component. One such decomposition is the common trend representation of Stock and Watson (1988), also known as the Beveridge-Nelson decomposition,

$$p_t = \iota\tilde{p}_t + \sum_{j=0}^{\infty} \tilde{C}_j \epsilon_{t-j}, \tag{33a}$$

$$\tilde{p}_t = \tilde{p}_{t-1} + \theta'\epsilon_t, \tag{33b}$$

where

$$\tilde{C}_j = -\sum_{i=j+1}^{\infty} C_i.$$

The trend component $\tilde{p}_t$ is the efficient price. The innovation of the efficient price, $r_t = \theta'\epsilon_t$, has been used to measure the contribution of each dealer to the price discovery process. Let $\sigma^2 = \mathsf{E}[r_t^2]$ denote the variance of the efficient price. Hasbrouck (1995) proposed the variance decomposition

$$\sigma^2 = \theta'\Sigma\theta = \sum_{i=1}^{N}\sum_{j=1}^{N} \theta_i\theta_j\sigma_{ij} \tag{34}$$

to define information shares for each dealer. If the shocks $\epsilon_{it}$ would be mutually uncorrelated the information shares

$$k_i = \frac{\theta_i^2 \sigma_{ii}}{\sigma^2} \tag{35}$$

would measure the part of the variance of the innovation to the efficient price that is due to the information in dealer $i$'s quotes. When the covariances $\sigma_{ij}$ are not equal to zero, it is not clear how much of the covariance $\theta_i \theta_j \sigma_{ij}$ should be attributed to dealers $i$ and $j$. In empirical work the covariance terms are always large. For large $N$ the covariance terms could even dominate the contributions of the diagonal elements. By varying the order of the variables in $p_t$ in alternative Cholesky decompositions of $\Sigma$ it is possible to obtain an upper and a lower bound.

Although the variance decomposition (35) might not always be very informative, the parameters $\theta$ and $\Sigma$ fully describe the instantaneous relation between the innovation to the efficient price $r_t$ and the price innovations $\epsilon_t$. Instead of reporting upper and lower bounds of the variance decomposition, we can learn more from alternative decompositions of $\epsilon_t$. The essence of the common trends representation is that there is a single random walk component with innovations that are correlated with all elements in $\epsilon_t$. For each individual market (dealer) we can separate the total shock in a permanent part $\beta_i r_t$ and an orthogonal component $e_{it}$ with only transitory effects. Such a permanent/transitory decomposition was introduced by Blanchard and Quah (1989) in the macroeconomics literature and is often used in impulse response analysis of vector autoregressions as an alternative to the Cholesky decomposition. The decomposition for market $i$ follows from the regression of the price innovation $\epsilon_i$ on the change in the efficient price,

$$\epsilon_{it} = \beta_i r_t + e_{it}. \tag{36}$$

The regression parameters are defined as $\beta_i = \mathrm{Cov}(r_t, \epsilon_{it})/\mathrm{Var}(r_t)$. Since $r_t = \theta' \epsilon_t$, this leads to the following expression for $\beta = (\beta_1, \ldots, \beta_N)'$,

$$\beta = \frac{\Sigma \theta}{\sigma^2}. \tag{37}$$

The regression coefficients $\beta_i$ measure how much of each dealer's quote adjustment is due to a change in the efficient price. The quote updates over a longer horizon must eventually be equal to the change in the efficient price, and all $\beta_i$'s will go to one. When $\beta_i$ is less than one, the dealer's quote updates only partially reflect the change in the efficient price.

The residual variance in (36) follows as

$$\mathrm{Var}(e_t) = \Omega_{BN} = \Sigma - \Sigma\theta(\theta'\Sigma\theta)^{-1}\theta'\Sigma \tag{38}$$

The first component in all $\epsilon_{it}$ is the innovation to the efficient price $r_t$, while the remaining $N-1$ components are orthogonal to $r_t$ and are all temporary. After extracting the common permanent component, it is much more plausible that the remaining dealer noise is nearly uncorrelated.[4]

The fraction of the variance of $\epsilon_{it}$ due to a change in the efficient price is

$$R_i^2 = \frac{\beta_i^2\sigma^2}{\sigma_{ii}} = \frac{(\Sigma\theta)_i^2}{\sigma^2\sigma_{ii}} \tag{39}$$

Since it is constructed as a regression $R^2$, we have that $0 \leq R_i^2 \leq 1$. It is also identical to the $R^2$ of the reverse regression of $r_t$ on $\epsilon_{it}$ alone, and therefore equal to the Hasbrouck information share when the series $p_{it}$ is put first in the Cholesky decomposition. Except for exceptional cases this will be the upper bound of the Hasbrouck information share. When all covariances are zero, the measure is identical to Hasbrouck's information share $k_i$. The higher the $R_i^2$ the closer market $i$ is following the efficient price. If several markets closely track the efficient price, the sum of all $R_i^2$ will be (much) larger than one. Rather than summarising the information contents in a single number it is of interest to report both $\beta_i$ and $R_i^2$ for each market.

Both the Hasbrouck information shares and the permanent/transitory decomposition are defined on the reduced form parameters. They would be directly applicable to the reduced form of the unobserved components model. A more direct way to obtain information shares in the unobserved components model is by defining them in terms of the parameters in the unobserved components model. Price innovations in the UC are

$$v_t = \Delta p_t - \mathsf{E}[\Delta p_t | r_{t-\ell}, e_{t-\ell}, \ell > 0] = (\iota + \alpha) r_t + e_t \tag{40}$$

The difference with the reduced form is that the expectation of the price changes is conditional on the unobserved past idiosyncratic shocks and changes in the efficient price instead of past observed prices. The latter would, using the Kalman filter prediction equations, lead back to the reduced form implications.

In the structural unobserved components model the parameters $\beta_i$ in (36) become

$$\beta_i = \mathrm{Cov}(\Delta p_{it},\, r_t)/\mathrm{Var}(r_t) = \mathrm{Cov}(v_{it},\, r_t)/\mathrm{Var}(r_t) = 1 + \alpha_i. \tag{41}$$

[4] In the absence of a series that is identically equal to the efficient price, the residual covariance matrix $\Omega$ can not be fully diagonal, since it must be singular by construction: $\theta'\Omega_{BN}\theta = 0$.

This measure will depend on the particular normalization for $\alpha$, but notice that for the BN normalization, $\alpha = \Sigma\theta - \iota$, the value of $\beta$ is exactly the one from the reduced form MA model, $\beta = \Sigma\theta/\sigma^2$. This is also the maximum possible value for $\beta$, because the BN normalization gives the highest possible value for $\alpha$, i.e. the maximal correlation between $r_t$ and the idiosyncratic component $u_t$.

The fraction of the variance of $v_{it}$ due to a change in the efficient price in the structural model is

$$R_i^2 = \frac{\beta_i^2 \sigma^2}{\beta_i^2 \sigma^2 + \omega_{ii}}, \tag{42}$$

which also depends on the normalization or identification for $\alpha$. Using the measures from the unobserved components model therefore requires identification. Without full identification of $\alpha$ it depends on the range of admissable values for $w$ whether the identification problem is serious, which is an empirical issue.

# 4 Empirical Application

To illustrate the various models we consider a set of Nasdaq dealer quotes. For the five most active dealers for Intel we considered midquotes for the six month period February-July 1999 containing 123 trading days. The five top dealers are the two ECN's Island (ISLD) and Instinet (INCA) and the three wholesale dealers Spear, Leeds & Kellogg Capital (SLKC), Mayer and Schweitzer (MASH), and Knight/Trimark Securities (NITE). Quotes are sampled at two minutes intervals. Since Intel is a liquid stock, there are hardly any missing values at this sampling frequency.[5] The total number of observations for all series is 24,108.

The purpose of the illustrations is to highlight the differences between alternative specifications. From the example we can get an impression whether a UC model violates typical moments in high frequency quote data and give rise to misleading implications about the information contents of quotes or the interactions among dealers.

Results depend on the sample autocovariance matrices of the quotes changes. All sample covariances are estimated omitting the overnight returns. The contemporaneous covariance matrix and the first two lags are reported in table 1. Contemporaneous

[5] At higher frequencies we do not observe quote updates for the less active dealers in many time periods. Various ways to deal with these missings have been suggested, see for example Harris, McInish, Shoesmith and Wood (1995) and DeJong, Mahieu and Schotman (1998). For clarity in this empirical illustration of the parameterization issues, we decided to keep the econometrics as simple as possible and work with data at the two-minutes frequency.

correlations among the quotes changes is only around 0.4. Since cointegration implies that the long-run correlation must be equal to one, enough dynamic structure remains despite the relatively low two minutes sampling frequency. All first order autocorrelations are negative. Most first order autocorrelations are around -0.20, except for INCA, where it is only -0.10. Since the INCA quotes are much closer to a random walk than the others, we should expect that most of the price discovery will go through INCA. Second order covariances are negligible, except for SLKC and NITE.

The variance of the random walk component can be estimated from the long run covariance matrix

$$\bar{\Gamma} = \Gamma_0 + \sum_{i=1}^{L}(\Gamma_i + \Gamma_i') = \sigma^2 \iota\iota' \tag{43}$$

It is clear from table 1 that with $L = 2$ not all elements in $\bar{\Gamma}$ are the same, nor that all correlations are equal to one. For the three wholesale dealers the diagonal elements are still larger than for the two ECN's. Given the large number of observations, the differences are significant. Further lags must add some negative autocorrelations for the three dealers. We did not obtain full equality of all elements of $\bar{\Gamma}$ by adding a small number of lags. On the other hand, a few more lags hardly affects the estimate of the random walk variance $\sigma^2$. We therefore estimate all models with a maximum of second order lags, with cointegration as a maintained hypothesis. Applying GMM to estimate $\sigma^2$ from the ten moments in $\bar{\Gamma}$ gives $\hat{\sigma}^2 = 2.54$ with a standard error of 0.06.

Implications for $\alpha$ can be obtained from the moment matrix

$$D(\Gamma) = \Gamma_1' - \Gamma_1 + 2(\Gamma_2' - \Gamma_2) = \sigma^2(\alpha\iota' - \iota\alpha') \tag{44}$$

Elements of $D(\Gamma)$ scaled by $\sigma^2$ are reported in the last panel of table 1. In the table all columns of $D(\Gamma)$ are in deviation of the first element, assuming that $\alpha_{ISLD} = 0$. With this normalisation all columns should be equal and show estimates of the other $\alpha_i$'s. The sample moments in the table indeed exhibit a structure with almost identical columns. The magnitudes are the same in all columns. The $\alpha$ of ISLD is the biggest in all columns, while those of SLKC and NITE are the two smallest. The $\alpha$'s of INCA and MASH are about the same and close to ISLD.

## 4.1 Vector Error Correction Model

A VECM is the most common model for estimating information shares. We estimated the model with second order dynamics,

$$\Delta p_t = c + As_{t-1} + D\Delta p_{t-1} + \epsilon_t, \tag{45}$$

where $s_t$ is the vector of differences between the midquote of ISLD and each of the other four dealers, and $A$ a $(5 \times 4)$ matrix of error correction parameters, and $D$ a $(5 \times 5)$ matrix. The most salient features of the VECM are reported in table 2.

The estimates of the information shares confirm the results of Huang (2002) that the ECN's dominate the price discovery on Nasdaq. Individual information shares of either ECN's or regular dealers are, however, in extremely wide intervals. For example, the lower and upper bound for ISLD are 3% and 70% respectively. The wide intervals are caused by the strong contemporaneous correlations of the errors. The errors of ISLD have a correlation of 0.70 with INCA, the other ECN.[6]

From the alternative decomposition of the dealer shocks (36) we find that all $\beta_i$ are close to one with the exception of $\beta_{NITE} = 0.76$. NITE appears to adjust slower to changes in the efficient price. Since the maximum information shares in table 2 are equal to the regression $R^2$ in (36), we also conclude that the quotes of the two ECN's have less volatility around the efficient price than the quotes of the wholesale dealers.

## 4.2 Reduced Form Vector Moving Average

The reduced form VMA with second order dynamics is

$$\Delta p_t = c + \epsilon_t + (\iota\theta' - I - B_2)\epsilon_{t-1} + B_2\epsilon_{t-2} \tag{46}$$

The 45 parameters in $\theta$, $\Sigma$ and $B_2$ are estimated by GMM using the 65 moment conditions for $\Gamma_0$, $\Gamma_1$ and $\Gamma_2$. Table 3 shows estimation results. Hansen's J-statistic rejects the 20 overidentifiying moment conditions that result from the cointegration restriction $C(1) = \iota\theta'$. The empirical violation of this restriction in the model with second order lags was already evident in table 1. Although the VECM and VMA are not nested, it seems that the VMA fits the data better: all diagional elements of $\Sigma$ and also the determinant are smaller for the VMA.

[6] The wide intervals for the information shares are not an artefact of the sampling frequency: Huang (2002) finds similar wide intervals for Intel at the one minute frequency. Huang (2002) uses slightly different data though, since he aggregates individual dealers into categories.

Implications for the information shares are similar to the VECM results. Both minimum, maximum ($R_i^2$) and $\beta_i$ are close to the VECM estimates. Direct estimates of $\theta$ are close to the implied values from the VECM. The difference between $\beta_{ISLD}$ and $\beta_{INCA}$ is slightly bigger, indicating that Island (ISLD) is the dominating market maker, despite the higher maximum information share of Instinet (INCA). The high information share of Instinet is mainly caused by its low residual variance.

## 4.3 Unobserved Components

By reparameterising the VMA we obtain alternative observationally equivalent unobserved components representations with second order dynamics as in (23). Of these equivalent models table 3 already reports results for the Beveridge-Nelson representation. In table 4 we report results for a model in "Watson" format ($\sum_i \alpha_i = 0$) and two models with a diagonal covariance matrix $\Omega$. The first is the model with $e_{t-1}$ as lagged variable, the other with $r_{t-1}$ as lagged variable in the equation for the dealer noise $u_t$.

The implications are not different from either VECM or VMA. The two ECN's still dominate. Also, NITE has the lowest $\alpha$ and its quote updates have the lowest correlation $R^2$ with the efficient price. NITE is the dealer that contributes least to the price discovery process. It is the only dealer with a significantly different $\alpha_i$. For the "Watson" representation we can not reject the hypothesis that the differrences $\alpha_i - \alpha_{NITE}$ are all the same using the GMM test based on the difference of the J-statistic of restricted and unrestricted models.

Diagonality appears a good modeling assumption. Considering the large sample size, the restriction is only marginally rejected against the VMA (and its equivalent UC representations). It seems surprising that diagonality provides such a good fit to the data, since the correlation between the shocks of ISLD and INCA was -0.47 in the "Watson" model. Note, however, that $\alpha$ and $\Omega$ are not unrelated. A translation of $\alpha$ in the direction of $\iota$ induces a compensating change in the structure of $\Omega$. The results mainly show that we can shift $\alpha$ such that $\Omega$ becomes almost diagonal. Assuming a diagonal $\Omega$ is a structural modeling assumption about the behavior of dealers.

The fit of the model with lagged effects of the efficient price is significantly worse than all other models. Yet the estimates of the interesting parameters remain very similar.

# 5 Conclusion

In this paper we proposed an Unobserved Components model for price discovery in fragmented markets. The model decomposes the observed prices in an underlying common efficient price and market-specific transitory components. We show how this model is related to the usual VAR or VECM models for price discovery, and argue that the unobserved components model is a natural and parsimonious way of modeling price discovery. The parameters in the unobserved components model have natural interpretations as the variance of the efficient price, variances and covariances of the transitory terms, and correlations between transitory terms and the efficient price. Because of this structure, it is easy to impose economically interesting or plausible restrictions on the model, for example diagonality of the transitory term covariance matrix. Moreover, the dynamic structure (lag length) of the model can be easily adapted to the serial correlation pattern observed in the data.

We also propose a new measure for the contribution to price discovery based on a permanent/ transitory decomposition of the error terms instead of the usual Cholesky decomposition. This measure is based on the covariance between the transitory components and the efficient price and can also be applied in the context of the usual VECM models.

Our empirical example using Nasdaq quotes illustrates the approach. We conclude that the key parameters of interest can be estimated from a parsimonious unobserved components model. These parsimonious models could prove useful for applications on smaller data sets, for example around specific events as corporate announcements.

# Appendix A Maximum $\alpha$

In this appendix we show that the BN normalization of $\alpha$ is the maximum possible value. Recall that for the BN normalization the covariance matrix of $e_t$ is semi-definite

$$\Omega_{BN} = \Sigma - \frac{\Sigma\theta\theta'\Sigma}{\theta'\Sigma\theta}$$

For other normalizations of $\alpha$, we can write

$$\Omega = \Omega_{BN} + w(\iota\theta'\Sigma + \Sigma\theta\iota') - w(w+1)\sigma^2\iota\iota' \tag{A1}$$

We now show that this implies that only positive values for $w$ are allowed. First, pre- and post-multiply the expression for $\Omega$ by $\theta$ and note that $\theta'\Omega_{BN}\theta = 0$,

$$\theta'\Omega\theta = 2w\sigma^2(\theta'\iota) - w(w+1)\sigma^2(\theta'\iota)^2 \tag{A2}$$

The right hand side of this equation is a quadratic function of $w$ with $w = 0$ and

$$w_1 = \frac{2}{\theta'\iota} - 1 \tag{A3}$$

As long as $0 < \theta'\iota < 2$, $w_1$ is positive and $\theta'\Omega\theta$ is positive for values $0 < w < w_1$. Negative values for $w$ are not allowed, like too high positive values (too low values of $\alpha$). The condition $0 < \theta'\iota < 2$ seems reasonable. Consider for example the MA model for $\theta'\Delta p_t$,

$$\theta'\Delta p_t = \theta'\epsilon_t - \theta'(I - \iota\theta')\epsilon_{t-1}, \tag{A4}$$

which can be written as

$$\theta'\Delta p_t = e_t - (1 - \theta'\iota)e_{t-1}, \tag{A5}$$

with $e_t = \theta'\epsilon_t$. An MA coefficient $1 - \theta'\iota$ between 0 and 1 seems reasonable for stationary microstructure data with negative first order serial correlation. If $\theta'\iota = 1$, then $\theta' p_t$ is a weighted average of individual prices which follows a random walk, equal to the efficient price $p_t^*$. In the empirical applications we always find that $0 < \theta' iota < 2$.

# Appendix B Proof of theorem 1

The representation for the price change is

$$\Delta p_t = (\iota + \alpha)r_t - \alpha r_{t-1} + \Psi_0 e_t + \sum_{i=1}^{M}(\Psi_i - \Psi_{i-1})e_{t-i} - \Psi_M e_{t-M-1} \tag{B1}$$

We start by analysing the covariance structure of the series $\Psi_0 e_t + \sum_{j=1}^{M}(\Psi_j - \Psi_{j-1})e_{t-j} - \Psi_M e_{t-M-1}$. The auto-covariances are

$$\begin{aligned}
\tilde{\Gamma}_{M+1} &= -\Psi_M\Psi_0' \\
\tilde{\Gamma}_M &= -\sum_{i=M-1}^{M}\Psi_i\Psi_{i-M+1}' \\
\tilde{\Gamma}_j &= (\Psi_j - \Psi_{j-1})\Psi_0' + \sum_{i=j+1}^{M}(\Psi_i - \Psi_{i-1})(\Psi_{i-j} - \Psi_{i-j-1})' - \Psi_M(\Psi_{M-j+1} - \Psi_{M-j})' \\
&= -\sum_{i=j-1}^{M}\Psi_i\Psi_{i-j+1}' + 2\sum_{i=j}^{M}\Psi_i\Psi_{i-j}' - \sum_{i=j+1}^{M}\Psi_i\Psi_{i-j-1}' \qquad 1 < j < M
\end{aligned} \tag{B2}$$

Summing the elements in (B2) gives

$$\sum_{j=1}^{M+1} j\tilde{\Gamma}_j = -\sum_{i=0}^{M} \Psi_i \Psi_i', \tag{B3}$$

since all terms of the form $\sum_{i=j}^{M} \Psi_i \Psi_{i-j}'$ cancel because the coeffients $(-(j+1)+2j-(j-1))$ are always zero. Putting the efficient price changes $(\iota+\alpha)r_t - \alpha r_{t-1}$ back in, the same sum of the moments of $\Delta p_t$ follows as

$$\sum_{j=1}^{M+1} j\Gamma_j = -\sigma^2 \alpha(\iota+\alpha)' - \sum_{i=0}^{M} \Psi_i \Psi_i' \tag{B4}$$

Subtracting the transpose of this matrix all symmetric terms cancel and we are left with the result

$$\sum_{j=1}^{M+1} j(\Gamma_j' - \Gamma_j) = \sigma^2(\alpha\iota' - \iota\alpha') \tag{B5}$$

Analogous algebra establishes the same result for the specification (28).

# References


Blanchard, O.J. and D. Quah (1989), The Dynamic Effects of Aggregate Demand and Supply Disturbances, *American Economic Review*, **79**, 655–673.

Covrig, V. and M. Melvin (2002), Asymmetric Information and Price Discovery in the FX Market: Does Tokyo Know more about the Yen?, *Journal of Empirical Finance*, **9**, 271–285.

De Jong, F., R.J. Mahieu, and P.C. Schotman (1998), Price Discovery in the Foreign Exchange Market, *Journal of International Money and Finance*, **17**, 5–27.

Harris, F.H. deB., T.H. McInish, G.L. Shoesmith, and R.A. Wood (1995), Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets, *Journal of Financial and Quantitative Analysis*, **30**, 563–579.

Harris, F.H. deB., T.H. McInish, and R.A. Wood (2002), Security Price Adjustments across Exchanges: An Investigation of Common Factor Components for Dow Stocks, *Journal of Financial Markets*, **5**, 277–308.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

Hasbrouck, J. (1993), Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement, *Review of Financial Studies*, **6**, 191–212.

Hasbrouck, J. (1995), One Security, Many Markets: Determining the Contributions to Price Discovery, *Journal of Finance*, **50**, 1175–1199.

HUANG, R.D. (2002), The Quality of ECN and NASDAQ Market Maker Quotes, *Journal of Finance*, **57**, 1285–1319.

HUPPERETS, E.C.J. AND A.J. MENKVELD (2002), Intraday Analysis of Market Integration: Dutch Blue Chips traded in Amsterdam and New York, *Journal of Financial Markets*, **5**, 57–82.

LEHMANN, B.N. (2002), Some Desiderata for the Measurement of Price Discovery across Markets, *Journal of Financial Markets*, **5**, 259–276.

SIMS, C.A. (1980), Macroeconomics and Reality, *Econometrica*, **48**, 1–48.

STOCK, J.H. AND M.W. WATSON (1988), Testing for Common Trends, *Journal of the American Statistical Association*, **83**, 1097–1107.

UPPER, C. AND T. WERNER (2002), Tail wags dog? Time-varying information shares in the Bund market, Discussion paper 24/02, Economic Research Centre of the Deutsche Bundesbank.

ZHOU, B. (1996), High-Frequency Data and Volatility in Foreign-Exchange Rates, *Journal of Economics and Business Statistics*, **14**, 45–52.

Table 1: Data (Auto-)Covariances

The table reports the sample covariances (*correlations*) for the time series of quote changes of the five most active dealers in Intel in the period February-July 1999. The entry on row $i$ and column $j$ for $\Gamma_\ell$ refers to the covariance $\mathsf{E}[\Delta p_{it} \Delta p_{j,t-\ell}]$. The long-run covariance matrix is defined as $\bar{\Gamma} = \Gamma_0 + \sum_{i=1}^{2}(\Gamma_i + \Gamma_i')$. The dealer information matrix is defined as $D(\Gamma) = \sum_{i=1}^{2} i(\Gamma_i' - \Gamma_i)/\sigma^2$. Columns in this matrix are shown in deviation of the first element, so that the row corresponding to dealer ISLD consists of zeros by construction. The scaling factor $\sigma^2$ is a GMM estimate from $\bar{\Gamma}$. Dealer acronyms are ISLD (Island), INCA (Instinet), SLKC (Spear, Leeds & Kellogg Capital), MASH (Mayer and Schweitzer) and NITE (Knight/Trimark Securities).

| | Dealer | ISLD | INCA | SLKC | MASH | NITE |
|---|---|---|---|---|---|---|
| Lag 0 | ISLD | 5.37 | *0.54* | *0.38* | *0.40* | *0.28* |
| ($\Gamma_0$) | INCA | 2.36 | 3.50 | *0.50* | *0.48* | *0.37* |
| | SLKC | 2.37 | 2.49 | 7.08 | *0.36* | *0.29* |
| | MASH | 2.51 | 2.41 | 2.59 | 7.24 | *0.27* |
| | NITE | 1.84 | 1.95 | 2.21 | 2.07 | 7.98 |
| Lag 1 | ISLD | -1.31 | 0.08 | 0.02 | 0.02 | 0.01 |
| ($\Gamma_1$) | INCA | 0.22 | -0.36 | 0.12 | 0.13 | 0.11 |
| | SLKC | 0.64 | 0.48 | -1.21 | 0.42 | 0.29 |
| | MASH | 0.09 | 0.07 | 0.04 | -2.06 | -0.13 |
| | NITE | 0.52 | 0.41 | 0.31 | 0.44 | -1.75 |
| Lag 2 | ISLD | -0.03 | 0.05 | 0.08 | -0.01 | 0.01 |
| ($\Gamma_2$) | INCA | 0.02 | -0.04 | -0.03 | 0.02 | -0.05 |
| | SLKC | -0.11 | -0.11 | -0.44 | -0.14 | -0.14 |
| | MASH | 0.04 | 0.02 | 0.02 | -0.09 | 0.12 |
| | NITE | 0.00 | 0.02 | -0.02 | -0.02 | -0.58 |
| Long run | ISLD | 2.70 | *1.01* | *0.94* | *0.94* | *0.80* |
| ($\bar{\Gamma}$) | INCA | 2.73 | 2.71 | *0.94* | *0.94* | *0.81* |
| | SLKC | 2.99 | 2.95 | 3.79 | *0.87* | *0.75* |
| | MASH | 2.64 | 2.65 | 2.90 | 2.93 | *0.79* |
| | NITE | 2.39 | 2.44 | 2.65 | 2.48 | 3.32 |
| Information | ISLD | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $D(\Gamma)$ | INCA | -0.04 | -0.04 | -0.06 | -0.06 | -0.05 |
| | SLKC | -0.17 | -0.16 | -0.17 | -0.15 | -0.14 |
| | MASH | -0.05 | -0.03 | -0.07 | -0.05 | -0.03 |
| | NITE | -0.20 | -0.19 | -0.23 | -0.21 | -0.20 |
| $\sigma^2$ | 2.54 | | | | | |

Table 2: Vector Error Correction

The table reports results obtained from the vector error correction model

$$\Delta p_t = c + As_{t-1} + D\Delta p_{t-1} + \epsilon_t.$$

The vector $s_t$ contains the difference between the quotes of ISLD and each of the other four dealers. Parameters are estimated by OLS. The table reports estimates of the long-run impact matrix of the VECM,

$$C(1) = \iota\theta'.$$

Also reported are the regression parameters from the errors on the random walk component,

$$\epsilon_t = \beta r_t + e_t,$$

with $r_t = \theta'\epsilon_t$. The "Info shares" are the minimum and maximum information shares (percentage) for each of the dealers, estimated using the methodology of Hasbrouck (1995). The residual covariance matrix is $\Sigma = \mathsf{E}[\epsilon_t\epsilon_t']$. Correlations are in *italics*. The last entry in the table is the variance of the random walk component, $\sigma^2 = \theta'\Sigma\theta$.

| | | | Info shares | | residual covariances (*corelations*) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dealer | $\theta$ | $\beta$ | min | max | ISLD | INCA | SLKC | MASH | NITE |
| ISLD | 0.21 | 1.05 | 0.03 | 0.70 | 4.38 | *0.70* | *0.52* | *0.55* | *0.37* |
| INCA | 0.53 | 1.03 | 0.12 | 0.91 | 2.66 | 3.28 | *0.61* | *0.61* | *0.44* |
| SLKC | 0.10 | 1.04 | 0.01 | 0.52 | 2.64 | 2.66 | 5.82 | *0.47* | *0.31* |
| MASH | 0.09 | 1.00 | 0.01 | 0.51 | 2.71 | 2.60 | 2.67 | 5.55 | *0.35* |
| NITE | 0.04 | 0.78 | 0.01 | 0.26 | 2.01 | 2.04 | 2.12 | 2.15 | 6.61 |
| $\sigma^2$ | 2.80 | | | | | | | | |

Table 3: Vector Moving Average

The table reports results obtaind from the vector moving average model

$$\Delta p_t = B_2\epsilon_{t-2} + B_1\epsilon_{t-1} + \epsilon_t$$

under the cointegration restriction

$$C(1) = I + B_1 + B_2 = \iota\theta'.$$

Parameters are estimated by GMM using the moment conditions for $\Gamma_0$, $\Gamma_1$ and $\Gamma_2$. Also reported are the regression parameters from the errors on the random walk component,

$$\epsilon_t = \beta r_t + e_t,$$

with $r_t = \theta'\epsilon_t$. The "Info shares" are the minimum and maximum information shares for each of the dealers, estimated using the methodology of Hasbrouck (1995). The residual covariance matrix is $\Sigma = \mathsf{E}[\epsilon_t\epsilon_t']$. Correlations are in *italics*. The last part of the table shows the variance of the random walk component, $\sigma^2 = \theta'\Sigma\theta$ and the criterion value of the GMM estimator known as Hansen's J-statistic.

| | | | Info shares | | Residual covariances (*correlations*) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dealer | $\theta$ | $\beta$ | min | max | ISLD | INCA | SLKC | MASH | NITE |
| ISLD | 0.25 | 1.11 | 0.07 | 0.75 | 4.31 | *0.71* | *0.55* | *0.59* | *0.46* |
| INCA | 0.49 | 1.05 | 0.14 | 0.90 | 2.66 | 3.23 | *0.63* | *0.64* | *0.53* |
| SLKC | 0.02 | 0.94 | 0.05 | 0.47 | 2.57 | 2.56 | 5.04 | *0.51* | *0.44* |
| MASH | 0.10 | 1.06 | 0.05 | 0.56 | 2.79 | 2.66 | 2.64 | 5.29 | *0.42* |
| NITE | 0.08 | 0.91 | 0.05 | 0.39 | 2.25 | 2.25 | 2.31 | 2.25 | 5.54 |
| | | | | | $\sigma^2 = 2.64$ | | | $J(20) = 103.21$ | |

Table 4: Unobserved Components Model

The table reports results for the unobserved components model

$$p_t = \iota p_t^* + u_t,$$
$$p_t^* = p_{t-1}^* + r_t.$$

In panels A and B the dealer shocks are defined as

$$u_t = \alpha r_t + \Psi e_{t-1} + e_t.$$

Panel A is the "Watson" representation with $\sum_i \alpha_i = 0$, obtained as a reparameterization of the VMA in table 3. In panel C the specification is

$$u_t = \alpha r_t + \phi \Delta r_t + e_t.$$

In panels B and C the covariance matrix $\Omega = \mathsf{E}[e_t e_t']$ is diagonal. Entries report GMM estimates for $\sigma^2$, $\alpha$, $\Omega$ and the GMM criterion function. The $R^2$ refers to the squared correlation beteen the innovation of a quote update and $r_t$.

### A) "Watson" representation

| | | | Residual covariances (*correlations*) | | | | |
|---|---|---|---|---|---|---|---|
| Dealer | $\alpha$ | $R^2$ | ISLD | INCA | SLKC | MASH | NITE |
| ISLD | 0.07 | 0.72 | 1.16 | *-0.47* | *-0.06* | *-0.13* | *-0.16* |
| INCA | 0.03 | 0.87 | -0.32 | 0.40 | *-0.05* | *-0.20* | *-0.17* |
| SLKC | -0.01 | 0.55 | -0.09 | -0.05 | 2.09 | *0.05* | *0.03* |
| MASH | 0.02 | 0.54 | -0.22 | -0.19 | 0.11 | 2.40 | *-0.07* |
| NITE | -0.12 | 0.38 | -0.32 | -0.19 | 0.07 | -0.21 | 3.43 |
| | | | | $\sigma^2 = 2.64$ | | $J(20) = 103.21$ | |

### B) Diagonal covariance matrix

| Dealer | $\alpha$ | $R^2$ | ISLD | INCA | SLKC | MASH | NITE |
|---|---|---|---|---|---|---|---|
| ISLD | 0.00 | 0.62 | 1.52 | | | | |
| INCA | -0.01 | 0.79 | | 0.62 | | | |
| SLKC | 0.09 | 0.57 | | | 2.17 | | |
| MASH | -0.03 | 0.51 | | | | 2.49 | |
| NITE | -0.14 | 0.33 | | | | | 3.59 |
| | | | | $\sigma^2 = 2.54$ | | $J(29) = 162.03$ | |

### C) Lagged efficient price effects

| Dealer | $\alpha$ | $R^2$ | ISLD | INCA | SLKC | MASH | NITE |
|---|---|---|---|---|---|---|---|
| ISLD | -0.05 | 0.60 | 1.03 | | | | |
| INCA | -0.04 | 0.80 | | 0.59 | | | |
| SLKC | -0.01 | 0.53 | | | 2.14 | | |
| MASH | -0.00 | 0.52 | | | | 2.27 | |
| NITE | -0.23 | 0.36 | | | | | 2.62 |
| | | | | $\sigma^2 = 2.54$ | | $J(49) = 508.13$ | |