Improving on daily measures of price discovery

Gustavo Fruet Dias, Marcelo Fernandes and Cristina M. Scherrer

CREATES Research Paper 2016-25
Smoothing price discovery measures over time in a continuous-time setting

Gustavo Fruet Dias, Marcelo Fernandes and Cristina M. Scherrer

CREATEES Research Paper 2016-25
Smoothing price discovery measures over time in a continuous-time setting

Gustavo Fruet Dias
Department of Economics and Business Economics, Aarhus University and CREATEES

Marcelo Fernandes
Queen Mary University of London and Sao Paulo School of Economics, FGV

Cristina Mabel Scherrer*
Department of Economics and Business Economics, Aarhus University and CREATEES

This version: April 20, 2017

Abstract: We formulate a continuous-time price discovery model and investigate how the standard price discovery measures vary with respect to the sampling frequency. We estimate daily measures of price discovery using a kernel-based estimator instead of running separate daily VECM regressions as is standard in the existing literature. We illustrate our theoretical findings by studying price discovery and its relationship with trading volume for 10 actively traded stocks in the U.S. from 2007 to 2013.

JEL classification numbers: C13, C32, C51, G14

Keywords: high-frequency data, price discovery, continuous-time model, sampling frequency, time-varying coefficients

*Corresponding author at Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. E-mail: cscherrer@econ.au.dk
1 Introduction

Price discovery analyses try to understand how markets impound new information in prices. There are essentially two standard price discovery measures. The first comprises any variant of Hasbrouck’s (1995) information share that gauges the contribution of each market/venue to the total variation of the efficient price innovation (see, for instance, Grammig, Melvin, and Schlag, 2005; Lien and Shrestha, 2009; Fernandes and Scherrer, 2014). The second relies on the permanent-transitory decomposition of Gonzalo and Granger (1995) and Gonzalo and Ng (2001). Applications of this measure to the price discovery analysis are usually named component share and include, among others, Booth, So, and Tseh (1999), Chu, Hsieh, and Tse (1999) and Figuerola-Ferretti and Gonzalo (2010). The common feature between the information share and the component shares is that they both rely on the estimation of the vector error-correction model (VECM). The speed-of-adjustment parameter from the VECM is specifically important for the construction of both measures. Baillie, Booth, Tse, and Zabotina (2002), Jong (2002) and Yan and Zivot (2010) provide a formal comparison between component and information shares in a discrete-time setting, showing that they render the same result if market innovations are contemporaneously uncorrelated.

This paper has two contributions to this literature. First, we examine both component share and information share measures of price discovery in a continuous-time setting. In particular, we put forward a continuous-time VECM with parameters that change in discrete time (say, at the daily frequency). Using the exact discretization from our continuous-time price discovery model, we investigate how the component and information shares vary with respect to the sampling frequency. We find that component share is invariant to the sampling frequency, implying that one may learn about the continuous-time price discovery at a lower frequency, say, one minute. This is in stark contrast with the information share measure, which converges as the sampling frequency decreases to the uninformative value of \(1/k\), where \(k\) denotes the number of markets in the analysis.

The second contribution concerns the estimation of time-varying price discovery measures. The standard in the price discovery literature is to estimate daily VEC models (Hasbrouck, 2003; Chakravarty, Gulen, and Mayhew, 2004; Hansen and Lunde, 2006; Mizrach and Neely, 2008), which is consistent with the continuous-time VECM with time-varying parameters we have in mind. Estimating individual daily VECMs essentially boils down to assuming that they are independent across
days. In this paper, we propose a framework to estimate time-varying price discovery measures that evolve smoothly over time and exploits the inter-daily information to obtain better finite-sample performance. In particular, we estimate daily speed-of-adjustment parameters in the VECM using Giraitis, Kapetanios, and Yates’s (2013) kernel methods. This means that we keep the estimates as nonparametric as possible in that our method allows for deterministic or stochastic variation of unknown form in the VECM parameters. This is in contrast with the parametric nature of Ozturk, van der Welv, and Dijk’s (2017) interesting state-space approach for the estimation of intraday price discovery measures, for instance.

We compare our smoothed price discovery measures with the standard measures based on daily VECM estimates. Through an extensive Monte Carlo exercise we show that smoothing indeed pays off. We then carry out an empirical application that examines the price informativeness of the New York Stock Exchange (NYSE) relative to the Nasdaq over a considerable large time span (namely, from 2007 to 2013). This is in contrast to the current studies on price discovery at the high frequency level that typically consider only one year of data. Using high-frequency midquotes of 10 of the most liquid stocks at the NYSE and Nasdaq, we find that there is indeed significant daily variation in the price discovery mechanism. By assuming that the latter changes smoothly from one day to the other, we are able to alleviate most of the noise in the daily VECM estimation, and hence offer a more precise picture of the relative informativeness of each market. Additionally, to better understand the daily variation in the price discovery mechanism, we study the long-run relationship between price discovery and trading volume.

The remainder of this paper is organized as follows. Section 2 describes the continuous-time setting for the price discovery mechanism and discusses how the information share and component share measures are affected by sampling frequency. Section 3 shows how to jointly estimate price discovery measures in a consistent manner accounting for daily stochastic changes in the covariance matrix and speed-of-adjustment parameter. Section 4 investigates how the price informativeness of the NYSE relative to the Nasdaq changes over time and the long-run relationship between price discovery and volume. Finally, we summarize our contributions in Section 5. The appendix presents the Monte Carlo study.
2 A continuous-time setting for price discovery

In this section, we propose a continuous-time model for price discovery and show how the component and information shares change with the sampling frequency. Let prices for a given asset that trades on multiple venues follow in day \(d\) the process

\[
dP_t = \Pi_0^{(d)} P_t \, dt + C_0^{(d)} \, dW_t, \quad \text{with} \quad P_0 = p_0^{(d)},
\]

where \(P_t = (p_{1,t}, \ldots, p_{k,t})'\) is a \((k \times 1)\) vector of log-prices with \(k\) denoting the number of trading venues, \(\Pi_0^{(d)} = \alpha_0^{(d)} \beta'\) is a \((k \times k)\) reduced-rank matrix with rank equal to \(0 < r < k\), \(\alpha_0^{(d)}\) and \(\beta\) are \((k \times r)\) full rank matrices, \(W\) is a \(k \times 1\) vector of Brownian motions, and \(C_0^{(d)}\) is a \(k \times k\) matrix, such that the covariance matrix \(\Sigma_0^{(d)} = C_0^{(d)} C_0^{(d)'}\) is assumed to be positive definite. Finally, the superscript \((d)\) indicates that \(\Pi_0^{(d)}\) and \(C_0^{(d)}\) may vary on a daily basis.

The reduced-rank Ornstein-Uhlenbeck process in (1) is the continuous-time counterpart of the discrete-time VECM in Hasbrouck (1995). Prices at the different markets should not drift much apart, oscillating around the (latent) efficient price, as they refer to the same asset. Accordingly, there is only one cointegrating relationship \((r = k - 1)\), with log-prices sharing the asset’s efficient price as a common stochastic trend. We assume without loss of generality that \(\beta\) is known and constant across days. In turn, \(\alpha_0^{(d)}\) determines how quickly each market reacts to deviations from the long-run equilibria given by \(\beta' P_t\).

Consider, for instance, Hasbrouck’s (1995) simple example, but in continuous time. Suppose that a homogenous asset trades on two markets. Market 1 is the leading trading venue, with prices fully reflecting the efficient price, whereas the price on Market 2 reacts to deviations with respect to the (efficient) price on Market 1. Prices in these markets cointegrate with \(\beta = (1, -1)'\):

\[
d\begin{pmatrix} p_{1,t} \\ p_{2,t} \end{pmatrix} = \begin{pmatrix} 0 & (1, -1) \\ \alpha_0^{(d)} & \beta' \end{pmatrix} \begin{pmatrix} p_{1,t} \\ p_{2,t} \end{pmatrix} \, dt + C_0^{(d)} \, dW(t) \quad (2)
\]

where \(C_0^{(d)}\) and \(W\) are defined as in (1). The absence of adjustment in Market 1 \((\alpha_{0,1}^{(d)} = 0)\) implies that \(p_{1,t}\) coincides with the stochastic trend, and hence a coherent price discovery measure should identify Market 1 as the sole contributor to the price discovery process, regardless of the sampling frequency.
Now, denote by $\exp(A)$ the matrix exponential of a $(k \times k)$ matrix $A$ such that $\exp(A) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} A^\ell$. The exact discretization of (1) at frequency $\delta$ in day $d$ reads

$$\Delta P_t = \Pi_\delta^{(d)} P_{t-1} + \varepsilon_{t_i},$$

(3)

where $\Pi_\delta^{(d)} = \alpha_\delta^{(d)} \beta'$ and $\alpha_\delta^{(d)} = \alpha_0^{(d)} \left( \beta' \alpha_0^{(d)} \right)^{-1} \left[ \exp(\delta \beta' \alpha_0^{(d)}) - I_k \right]$, with $I_k$ denoting a $k$-dimensional identity matrix. The sampling frequency $\delta$ is such that $t_i = i\delta$ with $i = 1, 2, \ldots, n$ and with $n$ denoting the number of intraday observations within a day. Lastly, $\varepsilon_{t_i}^{(d)}$ is iid Gaussian with zero mean and covariance matrix $\Sigma_\delta^{(d)} = \int_0^\delta \exp(u \Pi_\delta^{(d)}) \Sigma_0^{(d)} \exp(u \Pi_\delta^{(d)'}) du$. It is important to note that temporal aggregation preserves the cointegration rank (rank $\Pi_\delta^{(d)} = \text{rank} \Pi_0^{(d)}$) and the definition of (co)integration for Ornstein-Uhlenbeck processes in continuous time is consistent with the definition in discrete time (Kessler and Rahbek (2004)). This means that one may conduct inference about rank and cointegrating space using discrete-time procedures and then interpret the results in the continuous-time setting.

Computing price discovery measures requires the Granger representation of (1) and (3). Kessler and Rahbek’s (2001) Theorem 1 shows that the Granger representation indeed holds in continuous time, namely,

$$P_t = \Xi_0^{(d)} \left( C_0^{(d)} W_t + P_0^{(d)} \right) + \eta_t^{(d)},$$

(4)

where $\Xi_0^{(d)} = \beta_\perp \left( \alpha_{0,\perp}^{(d)} \beta_\perp \right)^{-1} \alpha_{0,\perp}^{(d)\prime}$, and $\eta_t^{(d)}$ is a stationary Ornstein-Uhlenbeck process. In turn, the Granger representation in discrete time reads

$$P_{t_i} = \Xi_\delta^{(d)} \sum_{h=1}^{i} \varepsilon_{t_h}^{(d)} + \sum_{h=0}^{\infty} Y_{\delta,h}^{(d)} \varepsilon_{t_i-h} + P_0^{(d)},$$

(5)

where $\Xi_\delta^{(d)} = \beta_\perp \left( \alpha_{\delta,\perp}^{(d)} \beta_\perp \right)^{-1} \alpha_{\delta,\perp}^{(d)\prime}$, and $P_0^{(d)}$ is a vector of initial values. The stochastic common trend given by the first term on the right-hand side of (5) reflects the efficient price of the asset and follows from $\beta_\perp$ being a vector of ones, which implies that $\Xi_\delta^{(d)}$ has common rows. In particular, it is reassuring to observe that the stochastic trend, $\Xi_\delta^{(d)} \sum_{h=1}^{i} \varepsilon_{t_h}^{(d)}$, is a martingale and thus is consistent with the asset pricing theory with regard to non-arbitrage requirements (see discussion in Hansen and Lunde (2006)).

The speed-of-adjustment matrix $\alpha_\delta^{(d)}$ plays a major role in any measure of price discovery. The matrix $\alpha_\delta^{(d)}$ reflects the adjustment that each market implements such that their prices do not
deviate from the efficient latent price. Hence, the closer the $\alpha^{(d)}_\delta$ of a given market is to zero, the less it has to adjust to the efficient price. In the limit, $\alpha^{(d)}_\delta = 0$ means that the price of market $j$ coincides with the efficient price, as in the above example. Accordingly, $\alpha^{(d)}_\delta$ shows that satellite markets have to adjust more strongly to deviations from the long-run equilibrium than leading markets.

The component share relies on the orthogonal projection of $\alpha^{(d)}_\delta$, namely, $\alpha^{(d)}_{\delta, \perp}$ such that $\alpha^{(d)}_{\delta, \perp} \alpha^{(d)}_\delta = 0$ (see, among others, Booth, So, and Tseh, 1999; Chu, Hsieh, and Tse, 1999; Harris, McInish, and Wood, 2002; Hansen and Lunde, 2006). Note that $\alpha^{(d)}_{\delta, \perp}$ is not unique, and hence one typically impose $\sum_{m=1}^{k} \alpha^{(d)}_{\delta, \perp, m} = 1$. While $\alpha^{(d)}_\delta$ corresponds to the stationary direction of the process in (3), $\alpha^{(d)}_{\delta, \perp}$ relates to the nonstationary direction, gauging the amount of information that is incorporated in the common stochastic trend (i.e., the efficient price). This makes $\alpha^{(d)}_{\delta, \perp}$ a natural quantity to assess how the efficient price relates to each market innovation. Unlike $\alpha^{(d)}_\delta$, $\alpha^{(d)}_{\delta, \perp}$ is increasing with price informativeness (see, among others, Harris, McInish, and Wood, 2002; Jong, 2002; Hansen and Lunde, 2006). The market with the highest $\alpha^{(d)}_{\delta, \perp}$ has the least need of adjustment towards the latent price and, hence, it is the one that leads the price discovery process.

Another popular price discovery measure in the literature is Hasbrouck’s (1995) information share (IS). In short, the IS measure gives the share of each market contribution to the total variance of the efficient price with $\sum_{m=1}^{k} IS_{\delta, m} = 1$. Using the exact discretization of (1), the IS measure of a given market $m \in \{1, \ldots, k\}$ for $0 \leq \delta < 1$ is

$$IS^{(d)}_{\delta, m} = \frac{\left[\xi^{(d)}_\delta C^{(d)}_\delta\right]_m}{\xi^{(d)}_\delta \Sigma^{(d)}_\delta \xi^{(d)}_\delta'},$$

where $\Sigma^{(d)}_\delta = C^{(d)}_\delta C^{(d)'}_\delta = \int_0^\delta \exp\left(u\Pi^{(d)}_0\right)\Sigma^{(d)}_0 \exp\left(u\Pi^{(d)'}_0\right) du$ and $\xi^{(d)}_\delta$ is the common row of $\Xi^{(d)}_\delta$ in (5). Because $C^{(d)}_\delta$ is not an unique decomposition of $\Sigma^{(d)}_\delta$, the standard practise in the literature consists of decomposing the covariance matrix using the Cholesky decomposition. The problem is that the latter decomposition depends on the ordering of the variables. The usual fix is to compute IS measures considering every possible ordering and then averaging them up to obtain the overall IS measure. It is clear from (6) that the information share of market $m \in \{1, \ldots, k\}$ depends on the sampling frequency $\delta$ essentially through the contemporaneous correlation across markets.

Interestingly, the information share for any market converges to the uninformative value of $1/k$ as
Figure 1 illustrates this inconvenient property. In particular, we entertain the exact discretization of (2), with $\delta$ ranging from zero to $1/390$ (30 minutes) and with different contemporaneous correlation across markets, $\rho_{0,12} \in \{0, 0.3, 0.5, 0.7, 0.9\}$. The left panel displays $IS_{\delta,1}^{(d)}$ and $\alpha_{\delta,\perp,1}^{(d)}$ for the alternative sampling frequencies, whereas the right panel plots the contemporaneous correlation $\rho_{\delta,12}$ across markets at frequency $\delta$. It is interesting to observe how discretization affects the price discovery measures in very different manners. The component share measure based on $\alpha_{\delta,\perp}^{(d)}$ is completely immune in that $\alpha_{0,\perp}^{(d)} = \alpha_{\delta,\perp}^{(d)}$ for any $0 \leq \delta < 1$. In contrast, the IS measure becomes totally uninformative as $\delta$ increases essentially because $\rho_{\delta,12}$ converges to one. The higher the correlation across markets and/or the absolute value of the speed-of-adjustment parameters, the faster the IS measure converges to $1/k$. This means the IS measure fails to identify Market 1 in (2) as the unique contributor to the price discovery process for any $\rho_{0,12} > 0$. The component share and information share measures yield the same result only in continuous time and in the absence of contemporaneous correlation between markets, i.e, for $\delta = 0$ and $\rho_{0,12} = 0$.

The continuous-time setting allows a better understanding of the price discovery measures at different sampling frequencies than earlier comparisons in discrete time (Baillie, Booth, Tse, and Zabotina, 2002; Jong, 2002; Yan and Zivot, 2010). In particular, in light of the above discussion, it seems much more convenient to use $\alpha_{\delta,\perp}^{(d)}$ as a measure of price discovery, unless one prefers to back out the continuous-time IS measure implied in $IS_\delta$.

3 Estimating daily measures of price discovery

From the notation in (5), the price discovery mechanism is constant within the day, but may change across days. The idea is that the ability of a trading venue to impound new information mostly depends on market features (e.g., cost structure, market design, technological infrastructure, and relative presence of high-frequency traders) and market characteristics (e.g., trading intensity, trading volume and volatility). Both market features and characteristics definitely change over time, but neither in a continuous nor in a brusk fashion. This is why we only allow for smooth discrete-time variation in the price discovery mechanism. This is consistent with (Eun and Sabherwal, 2003) and (Frijns, Gilbert, and Tourani-Rad, 2015), who show that price discovery drivers, such as volume
and volatility, are highly persistent. Accordingly, we henceforth assume that $\alpha_{\delta}^{(d)}$ changes in discrete time, at the daily frequency, as a bounded local stable stochastic process.

Consistent estimation of $\alpha_{\delta,\delta,\ldots}^{(d)}$ requires only a consistent estimator of $\alpha_{\delta}^{(d)}$ for any frequency $0 \leq \delta < 1$, given that we assume $\beta$ known. Because our sample consists of prices observed intradaily over different days and our estimation method uses the entire sample, it is convenient to adapt our notation as follows. Denote by $T = nD$ the total number of observations, where $n$ is the number of intraday observations and $D$ is number of trading days. As it is standard in the daily VECM approach, we augment the lag structure in (3) so that the autoregressive parameter matrices, $\Gamma_{\delta,j}^{(d)}$ with $j = 1, \ldots, \ell - 1$, are also allowed to vary over time:

$$\Delta P_{\tau} = \alpha_{\delta}^{(d)} \beta' P_{\tau-1} + \sum_{j=1}^{\ell-1} \Gamma_{\delta,j}^{(d)} \Delta P_{\tau-j} + \varepsilon_{\tau}, \quad d = 1, 2, \ldots, D \quad \tau = 1, 2, \ldots, T. \quad (7)$$

The literature usually captures the daily variation in price discovery by estimating daily VECM (see, for instance, Hasbrouck, 2003; Chakravarty, Gulen, and Mayhew, 2004; Mizrach and Neely, 2008).

In contrast, we employ Giraitis, Kapetanios, and Yates’s (2013) kernel-based estimator to retrieve daily estimates of $\alpha_{\delta}^{(d)}$ and $\Gamma_{\delta,j}^{(d)}$. This means that, as opposed to the daily VECM least-square estimates, our $\alpha_{\delta}^{(d)}$ estimates are not independent across days. The kernel approach exploits the assumption that the daily variations in the alpha and Gamma matrices are smooth in order to obtain more efficient estimators. Because we assume that $\beta$ is known, consistent estimation of the time-varying parameters in (7) follows directly from the results in Giraitis, Kapetanios, and Yates (2013). The kernel-based least-square (KLS) estimator exploits the assumption that the time-varying parameters are persistent processes (either deterministic or stochastic). Rewriting (7) in a more compact notation yields

$$\Delta P_{\tau} = B^{(d)} X_{\tau} + \varepsilon_{\tau}, \quad (8)$$

where $B^{(d)}$ is a $(k \times r + k(\ell - 1))$ matrix collecting the free parameters in (7) and $X_{\tau} = (P_{1,\tau-1} - P_{k,\tau-1}, \ldots, P_{k-1,\tau-1} - P_{k,\tau-1}, \Delta P_{t_{\tau-1}}, \ldots, \Delta P_{t_{\ell-1}})'$ with dimension $(r + k(\ell - 1) \times 1)$. In the event that the parameters are driven by stochastic processes, Giraitis, Kapetanios, and Yates (2013) requires that $\sup_{d \leq D} \|B^{(d)}\| = O_p(1)$ and a local stability condition in the form of $\sup_{d \leq D} \|B^{(d)} - B^{(d')}\|^2 = O_p(h/d)$. For instance, the bounded random walk process we consider in our
Monte Carlo study meets these conditions. Alternatively, in the case of deterministic variation in $B(d)$, Robinson (1989) shows that asymptotic normality of the kernel-based least-square estimator requires the parameters to satisfy a Lipschitz condition.

From Giraitis, Kapetanios, and Yates (2013), the kernel-based least-square estimator reads

$$\hat{B}(d) = \left( \sum_{\tau=1}^{T} b_{\tau,d} \Delta P_{\tau} X_{\tau}' \right)^{-1} \left( \sum_{\tau=1}^{T} b_{\tau,d} X_{\tau} X_{\tau}' \right),$$  \hspace{1cm} (9)$$

where $b_{\tau,d} := K ((nd - \tau) / H)$, with $H$ denoting a bandwidth such that, for fixed $n$, $H \to \infty$ and $H = o(nD / \ln (nD))$ as $D \to \infty$. The kernel function $K(x)$ is positive for any $x \in \mathbb{R}$ as well as a continuous and bounded, with a bounded first derivative such that $\int K(x) dx = 1$. Because the regressors in (8) are covariance stationary processes, it readily follows from Giraitis, Kapetanios, and Yates (2013, 2015) that, as $D \to \infty$,

$$\sqrt{H} \left( \text{vec}(\hat{B}(d)) - \text{vec} \left( B(d) \right) \right) \overset{d}{\to} N \left( 0, \left[ I_k \otimes Q(d)^{-1} \right] S(d) \left[ I_k \otimes Q(d)^{-1} \right] \right),$$  \hspace{1cm} (10)$$

where $\otimes$ denotes the Kronecker product, $Q(d) = \text{plim}_{D \to \infty} \frac{1}{H} \sum_{\tau=1}^{T} b_{\tau,d} X_{\tau} X_{\tau}'$, and $S(d) = \text{plim}_{D \to \infty} \frac{1}{H} \sum_{\tau=1}^{T} \zeta_t^{(d)} \zeta_t^{(d)'}$ for $\zeta_t^{(d)} = (b_{\tau,d} X_{\tau,1}, \ldots, b_{\tau,d} X_{\tau,k})'$.

To establish the asymptotic behavior of the price discovery measures, it suffices to employ the delta method. For instance, in the case of $k = 2$ markets, the orthogonal projections of the speed-of-adjustment parameters are

$$\alpha_{\delta,1}^{(d)} = \left( \begin{array}{c} -\alpha_{\delta,2}^{(d)} \\ \alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)} \\ \alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)} \end{array} \right),$$  \hspace{1cm} (11)$$

where $\alpha_{\delta,1}^{(d)} \alpha_{\delta,2}^{(d)} = 0$. The asymptotic distribution of the KLS estimator $\hat{\alpha}_{\delta,1}^{(d)}$ of $\alpha_{\delta,1}^{(d)}$ then reads

$$\sqrt{H} \left( \hat{\alpha}_{\delta,1}^{(d)} - \alpha_{\delta,1}^{(d)} \right) \overset{d}{\to} N \left( 0, \Lambda_\perp R_\alpha \left[ I_k \otimes Q(d)^{-1} \right] S(d) \left[ I_k \otimes Q(d)^{-1} \right] R_\alpha' \Lambda_\perp' \right),$$  \hspace{1cm} (12)$$

where $R_\alpha = (I_2, 0_{2 \times (d-1)})$ is a deterministic matrix that selects the elements of $\alpha_{\delta}^{(d)}$ from vec$(\hat{B}(d))$, and $\Lambda_\perp$ is the matrix of partial derivatives given by

$$\Lambda_\perp \equiv \frac{\partial \alpha_{\delta,1}^{(d)'}}{\partial \alpha_{\delta}^{(d)}} = \left( \begin{array}{cc} \frac{\alpha_{\delta,2}^{(d)}}{(\alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)})^2} & -\frac{\alpha_{\delta,1}^{(d)}}{(\alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)})^2} \\ \frac{\alpha_{\delta,1}^{(d)}}{(\alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)})^2} & \frac{\alpha_{\delta,2}^{(d)}}{(\alpha_{\delta,1}^{(d)} - \alpha_{\delta,2}^{(d)})^2} \end{array} \right).$$  \hspace{1cm} (13)$$
In a Monte Carlo study (see supporting material), we assess the relative performance of our estimation strategy to retrieve daily measures of price discovery. Smoothing the daily estimates of alpha entail much lower root mean squared errors than if we carry out the VECM estimation individually for each day. This holds for every instance we entertain, that is to say, regardless of the contemporaneous correlation between markets, sampling frequency, and bandwidth choice. The results heavily support the use of the KLS estimator for price discovery analyses.

4 Price informativeness: NYSE versus Nasdaq

In this section, we take to data our estimation strategy based on the KLS estimation of the speed-of-adjustment matrix. Given that we allow for time variation in the VECM, we use a massive high-frequency dataset that ranges from January 2007 to December 2013. This is in stark contrast with the current studies in price discovery that consider sample periods ranging from 4 to 12 months (see, for instance, de Jong and Schotman, 2010; Riordan and Storkenmaier, 2012; Ozturk, van der Welv, and Dijk, 2014; Benos and Sagade, 2016; Otsubo, 2017). We focus on 10 very actively traded stocks, namely, Bank of America (BAC), General Electric (GE), Hewlett-Packard (HPQ), International Business Machines (IBM), J.C. Penney Company (JCP), JP Morgan Chase (JPM), Coca-Cola Company (KO), Altria Group (MO), Verizon Communications (VZ), and ExxonMobil (XOM). We extract quotes data from TAQ for the two most active trading venues, NYSE and Nasdaq. We implement the same cleaning filters as in (Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2009), discarding any observation with a zero quote, negative bid-ask spread, or outside the main trading hours (9:30 to 16:00). We also discard any data point either with a bid-ask spread higher than 50 times the median spread on that day or with a midquote deviating by more than 10 mean absolute deviations from a rolling centered median of 50 observations. Finally, we take the median bid and ask quotes at each second in the event that there are multiple ticks taking place at the same second. We then synchronize the midquotes from both trading venues by sampling at regularly spaced intervals of 1 minute. Although the sampling frequency does not affect the component share in principle, Hupperets and Menkveld (2002), among others, advocate that a bit of time aggregation helps alleviate market microstructure noise. Table 1 details the cleaning process and provides the final number of observations for the 10 stocks we entertain.
We set the bandwidth to $H = n\sqrt{D}$ as in Giraitis, Kapetanios, and Yates (2013). The Monte Carlo study (see supporting material) confirms that this bandwidth choice indeed performs very well. We determine the lag structure by obtaining the most parsimonious specification in which we cannot reject the absence of residual autocorrelation at the 5% significance level (Hansen and Lunde, 2006).\footnote{Choosing the lag length that minimizes the Bayesian information criterion does not yield any qualitative changes. Results are available upon request.} As expected, we find only one cointegrating vector for every pair of stock prices using Johansen’s maximum eigenvalue and trace tests at the 1% significance level.\footnote{Daily test results are available upon request.}

Figure 2 plots the KLS estimates of $\hat{\alpha}_{(d)}^{(d)}$ and their respective 95% confidence intervals. As aforementioned, we normalize the orthogonal projections such that $\hat{\alpha}_{(d)}^{(d)} + \hat{\alpha}_{(d),T,N} = 1$, with subscripts $N$ and $T$ denoting NYSE and Nasdaq, respectively. We obtain confidence intervals for $\hat{\alpha}_{(d)}^{(d)}$ using the asymptotic distribution in (12). The daily variation in $\hat{\alpha}_{(d)}^{(d)}$ is very strong, with relative market informativeness alternating a lot over the sample. In particular, for the majority of the stocks, NYSE seems to contribute the most to price discovery as from mid 2011. Price discovery contribution typically moves in tandem with relative liquidity and NYSE experiences on average much more volume than Nasdaq.\footnote{The Nasdaq-NYSE volume ratio is a popular technical indicator of speculative activity, with high levels of Nasdaq trading relative to NYSE suggesting excessive speculation by investors.}

To better understand the daily variation in the price discovery mechanism, we estimate a VEC model that links daily estimates of the orthogonal projection of alpha and relative volume at the NYSE.\footnote{Depending on the stock, we need between 5 and 9 lags in the VAR specification to cope with the persistence in the orthogonal projection of alpha and relative volume. To simplify matters, we carry out the cointegration analysis using a VAR(9) specification for every stock, which leads to VEC representations with 8 lags.} In particular, we measure the latter by the logarithm of the NYSE-Nasdaq volume ratio. Both time series are very persistent, indicating the presence of unit roots. This is not surprising given that we allow the orthogonal projection of alpha to follow a bounded random walk. In addition, standard cointegration tests strongly suggest a long-run positive relationship between price discovery and relative volume.

Table 2 shows that every cointegrating vector is such that the NYSE contribution to the price discovery increases with the NYSE-Nasdaq volume ratio. It is interesting that the latter reacts significantly to deviations from the long-run equilibrium for every stock, whereas the same does not apply to the daily estimates of the orthogonal projection of alpha. In particular, the price
discovery measures of the BAC, GE, HPQ and JCP stocks do not respond to deviations from the long-run equilibrium. In addition, changes in the relative volume do not significantly affect future changes in the price discovery measures of BAC, HPQ, JCP, KO, MO, and VZ. Accordingly, we cannot reject at the usual significance levels the absence of Granger causality running from relative volume to the price discovery measures for BAC, HPQ and JCP. Altogether, it seems that relative volume chases more price discovery than the other way around.

5 Conclusion

We consider a different approach to address the estimation of daily measures of price discovery. First, we entertain a continuous-time price discovery model in which the speed-of-adjustment parameters evolve stochastically across days. We document that their orthogonal projection is invariant to the discretization frequency, whereas the information share converges to $1/k$ as the sampling frequency decreases. Second, by assuming that the VECM parameters are persistent over time, we show how to estimate smoothed price discovery measures by kernel-based least squares (Giraitis, Kapetanios, and Yates, 2013). By exploiting the inter-dependence across days, the KLS approach yields more efficient estimates than we would otherwise obtain by treating the daily variation in the VECM parameters as independent over time. Monte Carlo simulations indeed confirm that the KLS estimator easily outperforms the standard least-square daily VECM estimator under different scenarios. We assess the informativeness of the NYSE relative to Nasdaq for 10 actively traded stocks. We find strong evidence that market leadership alternates over time, with NYSE leading most of the price discovery only as from mid-2011. Finally, with an analysis of the long-run relationship between price discovery and volume, we find that relative volume chases more price discovery than the other way around.

6 Acknowledgments

Dias and Scherrer acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Fernandes thanks financial support from FAPESP (2013/22930-0) and CNPq (302272/2014-3).
References


Table 1: Data description

We report summary statistics for raw and cleaned data for Nasdaq and NYSE. The first two columns present the number of quotes (in millions) for each stock on the two trading venues before any cleaning filter (raw data). The following two columns display the total number of quotes (in millions) after the implementation of the cleaning procedure. The following two columns (obs per day) report the daily average number of quotes (in thousands). The last two columns report the total number of days we have for each stock in the sample period (January 2007 to December 2013).

<table>
<thead>
<tr>
<th></th>
<th>raw ('000,000)</th>
<th>clean ('000,000)</th>
<th>obs per day ('000)</th>
<th>number of days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nasdaq</td>
<td>NYSE</td>
<td>Nasdaq</td>
<td>NYSE</td>
</tr>
<tr>
<td>bac</td>
<td>523</td>
<td>503</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>ge</td>
<td>363</td>
<td>427</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>hpq</td>
<td>326</td>
<td>277</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>ibm</td>
<td>122</td>
<td>149</td>
<td>21</td>
<td>25</td>
</tr>
<tr>
<td>jcp</td>
<td>175</td>
<td>149</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>jpm</td>
<td>696</td>
<td>542</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>ko</td>
<td>244</td>
<td>205</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>mo</td>
<td>178</td>
<td>204</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>vz</td>
<td>264</td>
<td>257</td>
<td>25</td>
<td>29</td>
</tr>
<tr>
<td>xom</td>
<td>503</td>
<td>417</td>
<td>31</td>
<td>33</td>
</tr>
</tbody>
</table>
Table 2: Long-run relationship between price discovery and volume

We report the results of a VECM(8) for the daily price discovery measure and relative volume at the NYSE. We report the cointegrating vector estimates and their standard errors (note that we omit the one related to the price discovery measure as we force its loading to one) as well as the corresponding speeds of adjustment for the price discovery measure and for the relative volume. Finally, the last column displays the adjusted $R^2$ to gauge how much the VECM explains of the overall variation of the daily changes in the price discovery measure.

<table>
<thead>
<tr>
<th></th>
<th>cointegrating vector</th>
<th>speed of adjustment</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intercept</td>
<td>volume</td>
<td>price discovery measure</td>
<td>volume</td>
</tr>
<tr>
<td>bac</td>
<td>$-0.0749$ (0.1732)</td>
<td>$-1.6909$ (0.3827)</td>
<td>$-7.33 \times 10^{-5}$ (0.0002)</td>
<td>$0.0417$ (0.0097)</td>
</tr>
<tr>
<td>ge</td>
<td>$-1.0747$ (0.0825)</td>
<td></td>
<td>$-0.0011$ (0.0006)</td>
<td>$0.1242$ (0.0228)</td>
</tr>
<tr>
<td>hpq</td>
<td>$-1.9721$ (0.2823)</td>
<td></td>
<td>$-7.16 \times 10^{-5}$ (0.0003)</td>
<td>$0.0482$ (0.0102)</td>
</tr>
<tr>
<td>ibm</td>
<td>$-0.5962$ (0.0617)</td>
<td></td>
<td>$-0.0049$ (0.0014)</td>
<td>$0.1211$ (0.0282)</td>
</tr>
<tr>
<td>jcp</td>
<td>$-1.8361$ (0.2824)</td>
<td></td>
<td>$-5.33 \times 10^{-5}$ (0.0005)</td>
<td>$0.0403$ (0.0091)</td>
</tr>
<tr>
<td>jpm</td>
<td>$-0.3339$ (0.0926)</td>
<td>$-1.1122$ (0.2241)</td>
<td>$-0.0010$ (0.0004)</td>
<td>$0.0678$ (0.0140)</td>
</tr>
<tr>
<td>ko</td>
<td>$-0.6748$ (0.0501)</td>
<td></td>
<td>$-0.0054$ (0.0013)</td>
<td>$0.1499$ (0.0323)</td>
</tr>
<tr>
<td>mo</td>
<td>$-0.4463$ (0.0552)</td>
<td></td>
<td>$-0.0040$ (0.0014)</td>
<td>$0.1324$ (0.0362)</td>
</tr>
<tr>
<td>vz</td>
<td>$-0.9991$ (0.0864)</td>
<td></td>
<td>$-0.0032$ (0.0012)</td>
<td>$0.1224$ (0.0238)</td>
</tr>
<tr>
<td>xom</td>
<td>$-1.0343$ (0.0897)</td>
<td></td>
<td>$-0.0011$ (0.0007)</td>
<td>$0.1150$ (0.0214)</td>
</tr>
</tbody>
</table>
Figure 1: Information share and contemporaneous correlation as sampling frequency decreases

The first plot displays the average information share $IS_{\delta,1}$ of Market 1 (across orderings), as well as its component share $\alpha_{\delta,\perp,1}$, for $\delta$ ranging from 0 to 1/390 (30 minutes). The second plot depicts the exact correlation across markets at each sampling frequency $\delta$. We consider a range of values for the correlation across markets at $\delta = 0$ (continuous time), and then compute discrete-time counterparts using the exact discretization of the reduced-rank Ornstein-Uhlenbeck process as in (2).
The plots portray the kernel-based LS daily estimates of $\alpha^{(d)}_\perp$ for each stock, with their 95% confidence bands (in shades). We fix the bandwidth at $n\sqrt{D}$, where $n$ is the number of intraday observations (average of 390 observations per day) and $D$ is the number of trading days (1735 days). We normalize the orthogonal projections such that the element-wise estimates sum up to one, i.e., $\hat{\alpha}^{(d)}_1 \perp + \hat{\alpha}^{(d)}_2 \perp = 1$. 
Figure 3: Trading Volume

The plots portrait the volume shares of Nasdaq (T) and NYSE (N) for each stock. For example, the volume share of Nasdaq is computed as the trading volume at Nasdaq over the sum of the trading volume at Nasdaq and NYSE.
7 Appendix

7.1 Monte Carlo study

We assess the relative performance of the KLS estimator for the daily measures of price discovery. Our setup consists of an asset traded at \( k = 2 \) trading venues. We simulate price data from the exact discretization of the continuous-time process in (2.1) for \( D = 500 \) trading days (about 2 years of daily data), with a contemporaneous correlation \( \rho_{0,12} \in \{0, 0.3, 0.5, 0.7, 0.9\} \). We sample prices at fixed intervals of 0.5, 1, 3, and 5 minutes. As a usual trading day entails 23,400 seconds (6.5 hours), our daily samples have size ranging from 39,000 to 390,000 observations. The elements of the speed-of-adjustment matrix \( \alpha^{(d)}_{\delta} = (\alpha^{(d)}_1, \alpha^{(d)}_2)^t \) follow bounded random walk processes at the 5 minute frequency with upper and lower bounds given by \( \alpha^{(d)}_{1,\delta = 5} \in \{-0.49, -0.01\} \) and \( \alpha^{(d)}_{2,\delta = 5} \in \{0.01, 0.49\} \). We then compute the speed-of-adjustment parameters at the different frequencies, including \( \delta = 0 \), by means of

\[
\alpha^{(d)}_{\delta} = \alpha^{(d)}_0 \left( \beta^t \alpha^{(d)}_0 \right)^{-1} \left[ \exp(\delta \beta^t \alpha^{(d)}_0) - I_k \right],
\]

where \( I_k \) denotes a \( k \)-dimensional identity matrix.

The instantaneous covariance matrix \( \Sigma^{(d)}_0 \) is set to evolve daily in a stochastic manner. Specifically, the log-variances follow AR(1) processes:

\[
\ln \sigma^2_{m,t} = \varphi_0 + \varphi_1 \ln \sigma^2_{m,t-1} + \varsigma \vartheta_{m,t}, \quad \text{for } m = 1, 2
\]

with \( \sigma^2_{1,t} \) and \( \sigma^2_{2,t} \) denoting the diagonal elements of \( \Sigma^{(d)}_0 \). The volatility innovations \( \vartheta_{1,t} \) and \( \vartheta_{2,t} \) are standard Gaussian white noises with a correlation of 0.95 between themselves. The coefficient of variation is given by \( \text{var} \left( \sigma^2_{m,t} \right) / \mathbb{E} \left[ \sigma^2_{m,t} \right]^2 = \exp(\varsigma/(1 - \varphi^2_1)) - 1 \). We calibrate the stochastic volatility models as in Jacquier, Polson, and Rossi (1994), namely, we fix the autoregressive parameter to 0.98, the expected annual volatility to 20%, and the coefficient of variation to 0.5.

We assume a known cointegrating vector \( \beta = (1, -1)^t \), and estimate \( \alpha^{(d)}_{\delta} \) both by OLS as in the daily VECM approach and by KLS using an Epanechnikov kernel, with bandwidth \( H \in \{n^{8/10} \sqrt{D}, n^{9/10} \sqrt{D}, n \sqrt{D} \} \), where \( n \) and \( D \) account for the numbers of intraday and trading days. We report the root mean squared errors of the OLS estimates of the speed-of-adjustment parameters and of \( \alpha_{\delta,1,\perp} \) relative to their KLS counterparts. Given the normalization of the orthogonal projections, the relative (root) mean squared errors of the \( \alpha_{\delta,1,\perp} \) and \( \alpha_{\delta,1,\perp} \) estimates are identical
by construction.

Table 3 displays the root mean squared error of the KLS estimator relative to LS over 1,000 replications. The smoothed estimates entail much lower root mean squared errors than those based on a daily VECM approach. This holds for every instance we entertain, that is to say, regardless of the contemporaneous correlation between markets, sampling frequency, and bandwidth choice. It turns out, nonetheless, that the difference in performance increases with the correlation across markets, but declines with the sampling frequency and with the bandwidth value. Table 4 report the sample bias of both LS and KLS estimators. Biases are very small for both estimators. There is a clear decreasing pattern with the sampling interval and bandwidth value, whereas the magnitude of the bias does not seem to vary with the correlation across markets. In the absence of significant bias, we conclude that the lower mean squared error of the KLS estimator is essentially due to a drastic decline of the sampling variance.

All in all, smoothing the daily estimates of alpha entail much lower root mean squared errors than if we carry out the VECM estimation individually for each day. This holds for every instance we entertain, that is to say, regardless of the contemporaneous correlation between markets, sampling frequency, and bandwidth choice. The results heavily support the use of the KLS estimator for price discovery analyses.
We document the performance of the LS and the KLS estimators of $\alpha_{\delta,1}$ and $\alpha_{\delta,2}$ as well as of the orthogonal projection $\alpha_{\delta,1,\perp}$ for $D = 500$ days. In particular, we report the root mean squared error of the KLS estimator relative to LS, so that figures below one imply better performance of the KLS estimator. Note that we do not report the results for $\alpha_{\delta,2,\perp}$ because they are by construction identical to the ones for $\alpha_{\delta,1,\perp}$. The instantaneous correlation between markets $\rho_{0,12}$ ranges from 0 to 0.9., whereas the sampling frequency ranges from 30 seconds to 5 minutes. More specifically, we sample data at the $1/2$-, $1$-, $2$-, $3$- and $5$-minute frequencies, yielding $n = 780, 390, 195, 130, 78$ intraday observations, respectively. We compute the KLS estimator using three bandwidths: $H = n^{8/10}\sqrt{D}$, $H = n^{9/10}\sqrt{D}$ and $H = n\sqrt{D}$.

<table>
<thead>
<tr>
<th>$\rho_{0,12}$</th>
<th>$\delta = 1/2$</th>
<th>$\delta = 1$</th>
<th>$\delta = 2$</th>
<th>$\delta = 3$</th>
<th>$\delta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H = n^{8/10}\sqrt{D}$</td>
<td>0.00</td>
<td>0.72</td>
<td>0.72</td>
<td>0.68</td>
<td>0.43</td>
</tr>
<tr>
<td>0.30</td>
<td>0.62</td>
<td>0.61</td>
<td>0.40</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>0.50</td>
<td>0.54</td>
<td>0.54</td>
<td>0.45</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>0.70</td>
<td>0.46</td>
<td>0.46</td>
<td>0.42</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>0.90</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.32</td>
<td>0.32</td>
</tr>
</tbody>
</table>

| $H = n^{9/10}\sqrt{D}$ | 0.00 | 0.89 | 0.88 | 0.82 | 0.65 | 0.64 | 0.63 | 0.45 | 0.46 | 0.44 | 0.37 | 0.37 | 0.37 | 0.28 | 0.28 | 0.27 |
| 0.30 | 0.73 | 0.72 | 0.48 | 0.54 | 0.53 | 0.49 | 0.39 | 0.39 | 0.36 | 0.32 | 0.32 | 0.31 | 0.26 | 0.26 | 0.25 |
| 0.50 | 0.62 | 0.61 | 0.50 | 0.46 | 0.45 | 0.42 | 0.34 | 0.34 | 0.32 | 0.29 | 0.29 | 0.28 | 0.23 | 0.24 | 0.23 |
| 0.70 | 0.49 | 0.49 | 0.43 | 0.38 | 0.37 | 0.34 | 0.29 | 0.29 | 0.28 | 0.25 | 0.25 | 0.25 | 0.22 | 0.22 | 0.21 |
| 0.90 | 0.34 | 0.34 | 0.30 | 0.28 | 0.28 | 0.24 | 0.24 | 0.24 | 0.22 | 0.22 | 0.22 | 0.21 | 0.19 | 0.19 | 0.19 |

| $H = n\sqrt{D}$ | 0.00 | 1.16 | 1.14 | 1.02 | 0.79 | 0.78 | 0.74 | 0.51 | 0.52 | 0.49 | 0.39 | 0.39 | 0.39 | 0.28 | 0.28 | 0.27 |
| 0.30 | 0.93 | 0.91 | 0.57 | 0.63 | 0.62 | 0.59 | 0.41 | 0.41 | 0.40 | 0.32 | 0.32 | 0.32 | 0.24 | 0.24 | 0.23 |
| 0.50 | 0.77 | 0.75 | 0.60 | 0.52 | 0.52 | 0.48 | 0.35 | 0.35 | 0.34 | 0.28 | 0.28 | 0.28 | 0.22 | 0.22 | 0.21 |
| 0.70 | 0.60 | 0.59 | 0.50 | 0.41 | 0.41 | 0.37 | 0.29 | 0.29 | 0.28 | 0.24 | 0.24 | 0.23 | 0.19 | 0.19 | 0.19 |
| 0.90 | 0.37 | 0.36 | 0.31 | 0.28 | 0.27 | 0.23 | 0.21 | 0.21 | 0.19 | 0.19 | 0.19 | 0.18 | 0.16 | 0.16 | 0.16 |
Table 4: Bias of daily alpha and component share estimates

We document the bias ($\times 100$) of the LS and KLS estimators of $\alpha_{\delta,1}$ and $\alpha_{\delta,2}$ as well as of the orthogonal projection $\alpha_{\delta,1,\perp}$ for $D = 500$ days. As before, we do not report the results for $\alpha_{\delta,2,\perp}$ because they are symmetrical to the bias in the estimation of $\alpha_{\delta,1,\perp}$ by construction. The instantaneous correlation between markets $\rho_{0,12}$ ranges from 0 to 0.9, whereas the sampling frequency ranges from 30 seconds to 5 minutes. We compute the KLS estimator using a bandwidth $H = n^{8/10} \sqrt{D}$, with $b = 8, 9, 10$.

<table>
<thead>
<tr>
<th>$\rho_{0,12}$</th>
<th>$\delta = 1/2$</th>
<th>$\delta = 1$</th>
<th>$\delta = 2$</th>
<th>$\delta = 3$</th>
<th>$\delta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>least squares</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>0.21 0.23 -0.22 -0.24</td>
<td>0.35 0.38 -0.37 -0.38</td>
<td>0.60 0.66 -0.59 -0.66</td>
<td>0.84 0.86 -0.75 -0.76</td>
<td>1.35 1.41 -1.12 -1.15</td>
</tr>
<tr>
<td>0.30</td>
<td>-0.25 0.24 -0.23 -0.25</td>
<td>-0.31 0.30 -0.31 -0.31</td>
<td>-0.33 0.35 -0.36 -0.36</td>
<td>-0.33 0.34 -0.36 -0.36</td>
<td>-0.28 0.27 0.36 0.39</td>
</tr>
<tr>
<td>0.50</td>
<td>-0.25 0.24 -0.24 -0.24</td>
<td>-0.31 0.30 -0.30 -0.30</td>
<td>-0.32 0.32 -0.32 -0.32</td>
<td>-0.31 0.31 -0.31 -0.31</td>
<td>-0.25 0.27 0.39 0.39</td>
</tr>
<tr>
<td>0.70</td>
<td>-0.25 0.24 -0.25 -0.25</td>
<td>-0.31 0.30 -0.30 -0.30</td>
<td>-0.32 0.32 -0.32 -0.32</td>
<td>-0.31 0.31 -0.31 -0.31</td>
<td>-0.25 0.27 0.39 0.39</td>
</tr>
<tr>
<td>0.90</td>
<td>-0.26 0.23 -0.26 -0.26</td>
<td>-0.31 0.30 -0.30 -0.30</td>
<td>-0.25 0.25 -0.25 -0.25</td>
<td>-0.23 0.23 -0.23 -0.23</td>
<td>-0.13 0.13 0.16 0.16</td>
</tr>
</tbody>
</table>

For $H = n^{8/10} \sqrt{D}$:

<table>
<thead>
<tr>
<th>$\rho_{0,12}$</th>
<th>$\delta = 1/2$</th>
<th>$\delta = 1$</th>
<th>$\delta = 2$</th>
<th>$\delta = 3$</th>
<th>$\delta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>-0.44 0.42 -0.42 -0.42</td>
<td>-0.52 0.51 -0.52 -0.52</td>
<td>-0.52 0.53 -0.53 -0.53</td>
<td>-0.49 0.49 -0.49 -0.49</td>
<td>-0.41 0.41 0.48 0.48</td>
</tr>
<tr>
<td>0.30</td>
<td>-0.44 0.42 -0.43 -0.43</td>
<td>-0.53 0.52 -0.53 -0.53</td>
<td>-0.55 0.55 -0.55 -0.55</td>
<td>-0.49 0.49 -0.49 -0.49</td>
<td>-0.41 0.41 0.48 0.48</td>
</tr>
<tr>
<td>0.50</td>
<td>-0.44 0.42 -0.43 -0.43</td>
<td>-0.51 0.51 -0.51 -0.51</td>
<td>-0.50 0.50 -0.50 -0.50</td>
<td>-0.47 0.47 -0.47 -0.47</td>
<td>-0.37 0.37 0.52 0.52</td>
</tr>
<tr>
<td>0.70</td>
<td>-0.44 0.41 -0.45 -0.45</td>
<td>-0.52 0.50 -0.52 -0.52</td>
<td>-0.51 0.51 -0.51 -0.51</td>
<td>-0.48 0.48 -0.48 -0.48</td>
<td>-0.45 0.45 0.43 0.43</td>
</tr>
<tr>
<td>0.90</td>
<td>-0.44 0.41 -0.47 -0.47</td>
<td>-0.52 0.50 -0.52 -0.52</td>
<td>-0.45 0.45 -0.45 -0.45</td>
<td>-0.40 0.40 -0.40 -0.40</td>
<td>-0.29 0.29 0.60 0.60</td>
</tr>
</tbody>
</table>

For $H = n^{9/10} \sqrt{D}$:

<table>
<thead>
<tr>
<th>$\rho_{0,12}$</th>
<th>$\delta = 1/2$</th>
<th>$\delta = 1$</th>
<th>$\delta = 2$</th>
<th>$\delta = 3$</th>
<th>$\delta = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>-0.74 0.72 -0.74 -0.74</td>
<td>-0.84 0.83 -0.84 -0.84</td>
<td>-0.84 0.82 -0.84 -0.84</td>
<td>-0.75 0.78 -0.75 -0.78</td>
<td>-0.62 0.66 0.62 0.66</td>
</tr>
<tr>
<td>0.30</td>
<td>-0.73 0.72 -0.73 -0.73</td>
<td>-0.82 0.81 -0.82 -0.82</td>
<td>-0.79 0.80 -0.79 -0.80</td>
<td>-0.72 0.72 -0.72 -0.72</td>
<td>1.35 1.19 1.35 1.19</td>
</tr>
<tr>
<td>0.50</td>
<td>-0.73 0.71 -0.73 -0.73</td>
<td>-0.83 0.81 -0.83 -0.83</td>
<td>-0.78 0.82 -0.78 -0.82</td>
<td>-0.71 0.75 -0.71 0.75</td>
<td>-0.58 0.64 0.58 0.64</td>
</tr>
<tr>
<td>0.70</td>
<td>-0.73 0.71 -0.73 -0.73</td>
<td>-0.83 0.81 -0.83 -0.83</td>
<td>-0.78 0.82 -0.78 -0.82</td>
<td>-0.71 0.73 -0.71 0.73</td>
<td>-0.61 0.61 0.61 0.61</td>
</tr>
<tr>
<td>0.90</td>
<td>-0.74 0.71 -0.74 -0.74</td>
<td>-0.82 0.82 -0.82 -0.82</td>
<td>-0.72 0.80 -0.72 0.80</td>
<td>-0.65 0.82 -0.65 0.82</td>
<td>-0.49 0.73 0.49 0.73</td>
</tr>
</tbody>
</table>
Research Papers

2016

2016-08  N. Haldrup, O. Knapik and T. Proietti: A generalized exponential time series regression model for electricity prices

2016-09:  Ole E. Barndorff-Nielsen: Assessing Gamma kernels and BSS/LSS processes

2016-10:  Tim Bollerslev, Andrew J. Patton and Rogier Quaedvlieg: Modeling and Forecasting (Un)Reliable Realized Covariances for More Reliable Financial Decisions

2016-11:  Tom Engsted and Thomas Q. Pedersen: The predictive power of dividend yields for future inflation: Money illusion or rational causes?

2016-12:  Federico A. Bugni, Mehmet Caner, Anders Bredahl Kock and Soumendra Lahiri: Inference in partially identified models with many moment inequalities using Lasso

2016-13:  Mikko S. Pakkanen and Jani Lukkarinen: Arbitrage without borrowing or short selling?

2016-14:  Andrew J.G. Cairns, Malene Kallestrup-Lamb, Carsten P.T. Rosenskjold, David Blake and Kevin Dowd: Modelling Socio-Economic Differences in the Mortality of Danish Males Using a New Affluence Index

2016-15:  Mikkel Bennedsen, Ulrich Hounyo, Asger Lunde and Mikko S. Pakkanen: The Local Fractional Bootstrap

2016-16:  Martin M. Andreasen and Kasper Jørgensen: Explaining Asset Prices with Low Risk Aversion and Low Intertemporal Substitution

2016-17:  Robinson Kruse, Christian Leschinski and Michael Will: Comparing Predictive Accuracy under Long Memory - With an Application to Volatility Forecasting

2016-18:  Søren Johansen and Bent Nielsen: Tightness of M-estimators for multiple linear regression in time series

2016-19:  Tim Bollerslev, Jia Li and Yuan Xue: Volume, Volatility and Public News Announcements

2016-20:  Andrea Barletta, Paolo Santucci de Magistris and Francesco Violante: Retrieving Risk-Neutral Densities Embedded in VIX Options: a Non-Structural Approach

2016-21:  Mikkel Bennedsen: The Local Fractional Bootstrap

2016-22:  Søren Johansen and Morten Ørregaard Nielsen: The cointegrated vector autoregressive model with general deterministic terms

2016-23:  Yunus Emre Ergemen and Carlos Vladimir Rodríguez-Caballero: A Dynamic Multi-Level Factor Model with Long-Range Dependence

2016-24:  Shin Kanaya: Convergence rates of sums of α-mixing triangular arrays: with an application to non-parametric drift function

2016-25:  Gustavo Fruet Dias, Marcelo Fernandes and Cristina M. Scherrrer: Smoothing price discovery measures over time in a continuous-time setting