Price discovery and market microstructure noise

Gustavo Fruet Dias

School of Economics, University of East Anglia and CREATES

Marcelo Fernandes

Sao Paulo School of Economics, FGV

Cristina Mabel Scherrer

Norwich Business School, University of East Anglia and CREATES

This version: November 12, 2020

Abstract: Using a continuous-time price discovery model, we show that the standard econometric framework typically yields inconsistent estimates of price discovery measures in the presence of market microstructure noise. We address this errors-in-variable issue using instrumental variables. We devise valid instruments for two alternative microstructure noise settings, and then establish the asymptotic behavior (infill and standard asymptotic) of the corresponding price discovery measures. We illustrate our findings by investigating price discovery for Alcoa, showing that market leadership conclusions depend heavily on whether we account or not for market microstructure noise.

JEL classification numbers: G15, G12, G32, C32

Keywords: high-frequency data, component share, information share, market microstructure noise, instrumental variables, weak instruments, realized measures

Acknowledgments: We are grateful to Tim Bollerslev, Elise Gourier, Emmanuel Guerre, Andrew Harvey, Nikolaus Hautsch, Oliver Linton, Albert J. Menkveld, Roberto Renò, Kevin Sheppard, George Tauchen, and Viktor Todorov, as well as to many seminar/workshop participants for valuable comments. Fernandes thanks financial support from FAPESP (2013/22930-0 and 2019/05798-7) and CNPq (302278/2018-4). The usual disclaimer applies.

1 Introduction

Regulatory changes implemented in 2007 set out the vision of multiple trading venues competing for order flow and liquidity (see Menkveld, 2014; O'Hara, 2015; Menkveld, 2016; among others).[1] Additionally, high-frequency trading has been contributing to scatter quotes across the different exchanges, making markets much faster with time scales dropping to microseconds or even nanoseconds (O'Hara, 2015; Hasbrouck, 2018).

In the context of fragmented markets that operate at extremely fast time frames, answering questions on the joint dynamics of markets' quoting activity calls for the use data in high resolution. However, it is well known that market microstructure noise contaminates ultra-high frequency data (see Hansen and Lunde, 2006; Ait-Sahalia, 2007; Bandi and Russell, 2008; Barndorff-Nielsen, Hansen, Lunde and Shephard, 2008; among others), so that errors-in-variable issues arise. This paper shows how the usual estimation procedures deliver inconsistent estimates of the price discovery measures in the presence of market microstructure noise. In addition, we show how to obtain consistent and asymptotic normally distributed estimates of the price discovery measures under alternative assumptions on the market microstructure noise.

Market design and characteristics clearly affect the mechanism and timing of price formation. Price discovery analyses aims at distinguishing leading from satellite trading platforms by studying how quickly they impound new information into prices. There are essentially two standard price discovery measures. The first comprises any variant of Hasbrouck (1995) information share (IS) 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, 2018). The second relies on the component share (CS) measure resulting from the permanent-transitory decomposition of Gonzalo and Granger (1995) and Gonzalo and Ng (2001).

If the idea is to gauge how quickly markets react to price innovations, it is paramount to focus on price behavior at a very high frequency. However, microstructure noise effects typically increase with the sampling frequency, contaminating both transaction and quotes data. The resulting error-in-variable problem disrupts traditional price discovery analyses through two channels. First,

[1] Specifically, Regulation Alternative Trading Systems (RegATS) in 2000 and Regulation National Market System (RegNMS) in 2007 in the U.S., and Markets in Financial Instruments Directive (MiFiDin) in 2007 in Europe lay the foundation to the existence of multiple trading venues linked together and competing for liquidity and trades.

measurement error precludes the consistent estimation by least squares (LS) of the parameters that regulate how fast prices respond to changes in the efficient price. Second, the usual (realized) covariance matrix estimator is inconsistent, converging to the (integrated) covariance matrix of price innovations plus the covariance matrix of the microstructure errors. In this paper, we show how to compute price discovery measures that are robust to market microstructure effects.

At first glance, it seems that reducing the sampling frequency would suffice to alleviate market microstructure effects. This is indeed the usual fix in the realized measure literature (see, for instance, Liu, Patton and Sheppard, 2015). Bandi and Russell (2005, 2008) provide some theoretical justification to this practice by deriving the optimal choice of sampling frequency in a mean squared error sense. Zhang (2011) extends their results to deal with tradeoffs between the bias and various sources of stochastic error terms, including nonsynchronous trading, microstructure noise, and time discretization. It turns out that, even if sampling at a more moderate frequency helps with the second channel, it does not solve the first channel through which microstructure noise affects price discovery analyses. We indeed show that the bias in the least-squares (LS) estimation does not shrink to zero, even if it reduces as the sampling frequency decreases.

Exact discretization of the continuous-time vector equilibrium-correction (VEC) model shows how sampling frequency affects discrete price discovery measures. Dias, Fernandes and Scherrer (2019) show that CS measure is invariant to the sampling frequency, whereas the IS measure converges to the uninformative value of $1/M$ as sampling frequency decreases, where M is the number of markets. This allows us to assess the common practice of estimating the CS measure by sampling at a low enough frequency (say, 5 minutes) to eliminate market microstructure effects. However the classical errors-in-variable problem persists even at low frequency. To obtain consistent estimates of the price discovery measures in the presence of microstructure noise, we propose the use of instrumental variables (IV).

Instrument choice obviously depends on how market microstructure noises interact over time, and across assets and trading platforms. We propose valid instruments in two settings. The first allows for a more flexible cross-correlation structure of the microstructure noises across markets, but limits the order of their autocorrelations. The second accommodates more persistence in the microstructure noises, but constrains dependence across markets. The additional cost of allowing more persistent microstructure noise is that instruments in this setting are likely to be rather weak. As a result, we move away from the standard IV estimator, adopting a continuous-updating GMM estimator (CU).

We establish consistency and asymptotic normality under both time-span and infill asymptotics. The former treats the number of intraday observations as fixed, while allowing the number of days in the sample to diverge. The latter fixes the number of days (say, to one), but allows the number of intraday observations to grow to infinity. Monte Carlo simulations confirm not only the inconsistency of the usual price discovery measures in the presence of market microstructure noise, but also that the IV-based measures perform really well for different noise-to-signal ratios and contemporaneous correlations.

Apart from IV, we also employ Barndorff-Nielsen, Hansen, Lunde and Shephard's (2011) realized kernel (RK) estimator of the covariance matrix, which is consistent in the presence of market microstructure noise. This allows us to accommodate stochastic changes in the covariance matrix over time, as well. This is very convenient given the recent evidence that the price discovery mechanism changes over time. For instance, Hasbrouck (2003), Chakravarty, Gulen and Mayhew (2004), Hansen and Lunde (2006), and Mizrach and Neely (2008) estimate time-varying price discovery measures by employing daily VEC model, whereas Ozturk, van der Wel and van Dijk (2014) focus on intraday variations within a state space approach.

In our empirical illustration, we investigate price discovery for Alcoa (AA) on both New York Stock Exchange (NYSE) and Nasdaq Stock Market from June 2012 to May 2013. We use mid-quotes at the one-second frequency, yielding an average of almost 18,000 observations per day. We compute CS and IS estimates (and their robust standard errors) using both LS and IV methods. The traditional price discovery analysis based on LS estimates indicates that Nasdaq leads the price discovery for most of the sample period. Once we account for microstructure noise, the picture changes dramatically, with the NYSE leading the price discovery for the entire period. This finding may raise a few questions about previous results in the literature using high-frequency data.

The remainder of this paper is as follows. Section 2 describes the continuous-time setting we use for the price discovery mechanism. Section 3 shows how to estimate price discovery measures in a consistent manner accounting for the presence of market microstructure noise. Section 4 reports an extensive Monte Carlo study showing that our IV-based price discovery measures outperform the standard measures based on the daily VEC approach. Section 5 investigates how the price informativeness of the NYSE relative to the Nasdaq changes over time for the Alcoa stock share. Section 6 offers some concluding remarks.

2 A continuous-time setting for price discovery

As in Dias et al. (2019), we assume that the log-prices of a given asset that trades on multiple venues follow, on any given day, a continuous-time process given by

$$\mathrm{d}P_t = \Pi P_t \,\mathrm{d}t + C \,\mathrm{d}W_t, \tag{1}$$

where $P_t = (P_{1,t}, \ldots, P_{M,t})'$ is an $M \times 1$ vector collecting the log-prices in each of the M trading venues, $\Pi = \alpha\beta'$ is an $M \times M$ reduced-rank matrix, α and β are $M \times R$ full-rank matrices, W is an $M \times 1$ vector of Brownian motions, and C is an $M \times M$ matrix such that the covariance matrix $\Sigma = CC'$ is positive definite.

Although not every discrete-time VEC model results from the exact discretization of a continuous-time reduced-rank Ornstein-Uhlenbeck process, (1) provides a very convenient framework to study the dynamics of a single asset traded at multiple venues. In particular, it implies that the efficient price is a martingale, while allowing returns to follow VMA(∞) processes (Dias et al., 2019). 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 $R = M - 1$ cointegrating relationships, with log-prices sharing the asset's efficient price as the single common stochastic trend. We assume without loss of generality that β is known, taking the form of $\beta = (I_R, -\iota_R)'$, where ι_R denotes a $R \times 1$ unit vector.

In turn, α determines how quickly each market reacts to deviations from the long-run equilibria given by $\beta' P_t$. Note that α reflects not only institutional characteristics of each market, such as cost structure, technological infrastructure, and traders' composition, but also liquidity aspects, such as trading activity, volume and market share (Eun and Sabherwal, 2003; Figuerola-Ferretti and Gonzalo, 2010; Frijns, Gilbert and Tourani-Rad, 2015). Similarly, the price innovations in each trading platform also reflect the market microstructure effects that may feedback into the efficient price dynamics

The solution to (1) is an homogenous Gaussian Markov process given by

$$P_t = \exp(t\Pi)\left[P_0 + \int_0^t \exp(-u\Pi)\, C\, \mathrm{d}W_u\right], \tag{2}$$

whose exact discretization boils down to an homoskedastic Gaussian VAR(1) process that admits a Gaussian VEC(0) representation.[2] To formalize the properties of (1), Kessler and Rahbek (2004) restrict the autoregressive matrix Π so as to ensure identification.

Assumption RROU The reduced-rank OU process in (1) is such that all eigenvalues of Π are real and no elementary divisor of Π occurs more than once. In addition, α, β and $\beta'\alpha$ have full column rank R, with every eigenvalue of $\beta'\alpha$ having negative real parts.

Assumption RROU follows directly from Kessler and Rahbek (2004) and establishes the necessary conditions for (1) the uniqueness of the mapping $\theta = (\Pi, \Sigma) \overset{\psi}{\mapsto} \psi(\theta) = (\Pi_\delta, \Sigma_\delta)$ from continuous-time to discrete-time parameters at sampling interval δ; (2) the injectivity of the mapping ψ and identifiability of θ; (3) the existence of a Granger representation theorem (GRT) in continuous time; and (4) the exact discretization of the reduced-rank OU process in (1).

To work out the exact discretization of (1), we assume prices are observed regularly and equidistantly over the unit interval $[0, 1]$ that represents a trading day (calendar-time sampling, as discussed in Hansen and Lunde, 2006). Denote each interval contained in $[0, 1]$ as $[t_i, t_{i-1}]$, where $i = 1, 2, \ldots, n$ and n is the total number of intervals such that $0 = t_0 < t_1 < \ldots < t_n = 1$. The length of each interval is $\delta = t_i - t_{i-1} = 1/n$ in $[0, 1]$. For instance, the usual trading day in the U.S. market lasts for 6.5 hours (23,400 seconds), and hence sampling 1 observation per minute yields $n = 390$ and $\delta = 1/390$. Next, denote by $\exp(A)$ the matrix exponential of an $M \times M$ matrix A such that $\exp(A) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} A^\ell$. The exact discretization of (1) at interval length δ reads

$$\Delta P_{t_i} = \Pi_\delta P_{t_{i-1}} + \varepsilon_{t_i}, \tag{3}$$

where $\Pi_\delta = \alpha_\delta \beta'$ and $\alpha_\delta = \alpha\,(\beta'\alpha)^{-1}\,[\exp(\delta\beta'\alpha) - I_R]$, with I_R denoting a R-dimensional identity matrix, P_{t_i} is an $M \times 1$ vector of log-prices observed at discrete time, and the innovation ε_{t_i} is iid Gaussian with zero mean and covariance matrix given by $\Sigma_\delta = \int_0^\delta \exp(u\Pi)\,\Sigma \exp(u\Pi')\,\mathrm{d}u$. Note

[2] To contemplate reduced-rank continuous-time processes with more general lag structures in their discrete-time counterparts, it suffices to apply the Laplace transform function to the lag operator (Nguenang, 2016).

that $\alpha_\delta \equiv \alpha_\delta(\alpha)$ and $\Sigma_\delta \equiv \Sigma_\delta(\alpha, \Sigma)$ are known functions of the continuous-time parameters in (1). We next summarize the properties of the discrete-time VEC model in (3) implied by the exact discretization of (1) as derived by Kessler and Rahbek (2004).

Lemma 1 Let Assumption RROU holds, then the discrete-time VEC process in (3) is such that prices are not explosive in that the roots of the characteristic polynomial $|I_k - (\Pi_\delta + I_k)z| = 0$ are either outside the unit circle or equal to one; rank $(\alpha\beta') =$ rank $(\alpha_\delta\beta')$, the column space of α_δ is the same as the column space of α in that the number of unit roots equals $M - R$; and the estimation of the continuous-time parameters readily follows form the discrete-time log-likelihood in that $\ell_n^c(\alpha, \Sigma) = \ell_n^d(\alpha_\delta, \Sigma_\delta)$, where ℓ_n^c and ℓ_n^d denote the continuous- and discrete-time log-likelihoods, respectively.

Computing price discovery measures requires to decompose the price vector in a permanent I(1) component and a transitory I(0) component. In what follows, we consider the GRT decomposition (Stock and Watson, 1988; Johansen, 1991; Hansen, 2005) because it ensures that the stochastic trend given by the efficient price is a martingale. As any VEC decomposition, it depends exclusively on the orthogonal complements of the speed-of-adjustment parameters and of the cointegrating vector given that they relate to the nonstationary directions of the processes.[3] In particular, the orthogonal complements of the speed-of-adjustment parameters indicate how the stochastic trend relates to the efficient price innovations, and hence they play a major role in any price discovery analysis.

Under Assumption RROU, Theorem 1 in Kessler and Rahbek (2001) shows that the GRT also holds in continuous time:

$$P_t = \Xi\left(CW_t + P_0\right) + \alpha\left(\beta'\alpha\right)^{-1} H_t \tag{4}$$

where $\Xi = \beta_\perp\left(\alpha_\perp'\beta_\perp\right)^{-1}\alpha_\perp'$, P_0 contains initial values, and $H_t = \beta' P_t$ denotes a stationary Ornstein-Uhlenbeck process given by $\mathrm{d}H_t = \beta'\alpha H_t\,\mathrm{d}t + \beta' C\,\mathrm{d}W_t$. This result follows directly from plugging the identity

$$I_k = \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp' + \alpha(\beta'\alpha)^{-1}\beta' \tag{5}$$

[3] If A is an $M \times R$ matrix with full column rank R and $sp(A)$ is the subspace in $\mathbb{R}^M$ spanned by the columns of A, then $A_\perp$ is any matrix with dimension $M \times (M - R)$ such that $sp(A_\perp) = sp(A)_\perp$. In particular, we respectively denote the orthogonal complements of the speed-of-adjustment parameters in continuous and discrete times by $\alpha_\perp$ and $\alpha_{\delta,\perp}$, whereas $\beta_\perp = \iota_M$ corresponds to the orthogonal complement of the cointegrating vector in the price discovery context.

into the solution of (1).

Similarly, as the exact discretization in (3) yields the exact same data generation process, it follows that the GRT also holds in discrete time

$$P_{t_i} = \Xi_\delta \sum_{h=1}^{i} \varepsilon_{t_h} + \sum_{h=0}^{\infty} \Upsilon_{\delta,h} \varepsilon_{t_{i-h}} + \Xi_\delta P_0, \tag{6}$$

where $\Xi_\delta = \beta_\perp \left(\alpha'_{\delta\perp}\beta_\perp\right)^{-1} \alpha'_{\delta\perp}$, P_0 is a vector of initial values such that $P_{m,0} = P_{m',0}$ with $m, m' = 1, \ldots, M$, and $\Upsilon_{\delta,h}$ can be computed recursively from the parameters in (3) such that $\Upsilon_{\delta,h} = (I_M - \Xi_\delta)\left(I_M + \alpha_\delta \beta'\right)^h$ for $h = 0, 1, 2, ...$ (Corollary 2 in Hansen, 2005). The stochastic common trend given by the first term on the right-hand side of (6) reflects the efficient price of the asset. This stems from the fact $\beta_\perp = \iota_M$ implies that Ξ_δ has common rows. In particular, it is reassuring to observe that the stochastic trend $m_{t_i} = \left(\alpha'_{\delta\perp}\beta_\perp\right)^{-1} \alpha'_{\delta\perp} \sum_{h=1}^{i} \varepsilon_{t_h}$ is a martingale (Hansen and Lunde, 2006).

2.1 Price discovery measures

The component share (CS) measure captures how the efficient price relates to market innovations and hence depends exclusively on the orthogonal complement of the speed-of-adjustment parameters (see, among others, Booth, So and Tseh, 1999; Chu, Hsieh and Tse, 1999; Harris, McInish and Wood, 2002; Hansen and Lunde, 2006). Specifically, because Ξ_δ has common rows, the contribution of each market innovation to the efficient price is given by the columns of the $1 \times M$ vector $\left(\alpha'_{\delta,\perp}\beta_\perp\right)^{-1} \alpha'_{\delta,\perp}$. The orthogonal complement of α_δ is not unique and hence, without loss of generality, we normalize the sum of its elements to one, i.e., $\sum_{m=1}^{M} \alpha_{\delta,\perp,m} = 1$. Under this normalization, the common row of Ξ_δ reduces to $\alpha'_{\delta,\perp}$.

Dias et al. (2019) show that discretization does not affect $\alpha_{\delta,\perp}$ in that $\alpha_{0,\perp} = \alpha_{\delta,\perp}$ for any $0 < \delta < 1$. This implies that one can estimate the continuous-time orthogonal complements of the speed-of-adjustment parameters directly from discrete-sampled prices. It then follows that the continuous-time CS measure of the mth market reads

$$\mathrm{CS}_m = \alpha_{\delta,\perp_m}, \quad m = 1, \ldots, M. \tag{7}$$

Hasbrouck's (1995) information share (IS) is another prominent price discovery measure, gauging the contribution of each market/venue to the total variation in the efficient price innovation.

Apart from $\Xi_{\delta,\perp}$, the IS measure also takes into consideration the contemporaneous covariance between market innovations. Normalizing $\alpha_{\delta,\perp}$ such that $\sum_{m=1}^{M} \alpha_{\delta,\perp,m} = 1$ and using the fact that $\alpha_{0,\perp} = \alpha_{\delta,\perp}$ for any $0 \leq \delta < 1$, the discrete-time IS measure reads

$$IS_{\delta,m} = \frac{[\alpha'_{\perp} C_{\delta}]^2_m}{\alpha'_{\perp} \Sigma_{\delta} \alpha'_{\perp}}, \quad m = 1, \ldots, M \tag{8}$$

where $\Sigma_{\delta} = C_{\delta} C'_{\delta} = \int_0^{\delta} \exp(u\Pi) \Sigma \exp(u\Pi')\, \mathrm{d}u$ and $[\alpha'_{\perp} C_{\delta}]_m$ denotes the mth element of the $1 \times M$ vector $\alpha'_{\perp} C_{\delta}$.

It turns out that the information share becomes uninformative, converging to $1/M$, as δ increases. This happens because the contemporaneous correlation among markets increases with δ even if Σ is diagonal. This is the one of the main reasons why Hasbrouck (2018) advocates the use of prices sampled at high resolutions to obtain meaningful IS measures. Dias et al. (2019) formally address this issue by defining the continuous-time analogue of (8):

$$IS_m = \frac{[\alpha'_{\perp} C]^2_m}{\alpha'_{\perp} \Sigma \alpha'_{\perp}}, \quad m = 1, \ldots, M \tag{9}$$

which depends exclusively on the continuous-time parameters in (1). Altogether, it turns out that the estimation of the continuous-time CS measure and of the continuous- and discrete-time IS measures is down to estimating α, α_{δ}, Σ, and Σ_{δ} from discrete sampled prices.

Finally, one may also employ impulse response functions (IRFs) to build dynamic measures of price discovery in discrete time (Yan and Zivot, 2010; Scherrer, 2013; Nguenang, 2016). IRFs depict how each market responds over time to a change in the efficient price. Because the exact discretization in (3) implies that the innovations are Gaussian and homoskedastic, it is possible to obtain the IRF from the efficient price implied by the GRT in (6). In particular, it reads

$$\frac{\mathrm{d}p_{t_{i+h}}}{\mathrm{d}m_{t_i}} = \iota_M + \frac{\alpha'_{\perp} \beta_{\perp}}{\alpha'_{\perp} \Sigma_{\delta} \alpha_{\perp}} \Upsilon_{\delta,h} \Sigma_{\delta} \alpha_{\perp}, \tag{10}$$

with $\frac{\mathrm{d}p_{t_{i+h}}}{\mathrm{d}m_{t_i}}$ converging to one in every market because $\Upsilon_{\delta,h} \to 0$ as $h \to \infty$ (see, e.g., Hansen, 2005; Hansen and Lunde, 2006).

3 Price discovery analysis robust to microstructure noise

Most price discovery analyses compute CS and IS measures not only at different frequencies (e.g., using tick data, 30-second, 1-minute, and 5-minute returns) but also over different time spans (e.g., day, month, quarter, and year). These combinations between frequency and time span imply different sampling schemes and hence different asymptotic theory. Price discovery studies with long time spans (see, among others, Baillie, Booth, Tse and Zabotina, 2002; Grammig et al., 2005; Menkveld, Koopman and Lucas, 2007; Lien and Shrestha, 2009; Frijns et al., 2015; Fernandes and Scherrer, 2018) require fixing the sampling interval or, equivalently, the number of intraday observations, while letting the number of days grow. To estimate daily measures of price discovery (see, for instance, Hansen and Lunde, 2006; Hasbrouck, 2018), it is more appropriate to consider a large number of intraday observations, while fixing the number of days to one. In what follows, we provide asymptotic results under both settings.

In the absence of market microstructure contamination, the LS estimator is consistent for α_δ, with the sample covariance matrix converging to Σ_δ under standard regularity conditions. Additionally, one could consistently estimate α by nonlinear least squares and the continuous-time covariance matrix directly from the intraday returns at their highest frequency using a realized approach. However, in the presence of microstructure noise, it is no longer trivial to obtain consistent estimates of α_δ, Σ_δ, and Σ. Contamination biases the estimation of α_δ and, as such, also of Σ_δ. In fact, the measurement error affects both terms on the right-hand side of (4), leading to the misleading conclusion that market microstructure noise has long-run effects on the efficient price and its volatility. Moreover, the realized variance estimator is also biased and inconsistent in the presence of market microstructure noise. In fact, the variance of the latter dominates the probability limit of the realized variance estimator even under constant volatility and independent microstructure noise.

Let p_{m,t_i} denote the observed discrete-time log-price in market m at time t_i. It differs from the latent log-price P_{m,t_i} due to the market microstructure noise $u_{m,t_i} = p_{m,t_i} - P_{m,t_i}$. It follows that the observed prices is also VEC model with the same autoregressive polynomial than the discrete-time model in (3)

$$\Delta p_{t_i} = \alpha_\delta \beta' p_{t-1} + \mathrm{v}_{t_i}, \tag{11}$$

where Δp_{t_i} is an $M \times 1$ vector of observed price changes, $\mathrm{v}_{t_i} = \varepsilon_{t_i} + [I_M - (\alpha_\delta \beta' + I_M)L]\, u_{t_i}$ is an M-dimensional vector, with u_{t_i} denoting the $M \times 1$ vector of market microstructure noise and L

the usual lag operator.[4] It is apparent from (11) that $\beta' p_{t-1}$ correlates with v_{t_i} because both terms contains the past microstructure noise.

To appreciate why, consider the case with $M = 2$ markets for ease of exposition. After factoring the market microstructure noise, the first equation of (3) becomes

$$\Delta p_{1,t_i} = \alpha_{\delta,1}\left(p_{1,t_{i-1}} - p_{2,t_{i-1}}\right) + \mathrm{v}_{1,t_i}, \tag{12}$$

where $\mathrm{v}_{1,t_i} = \Delta u_{1,t_i} + \varepsilon_{1,t_i} - \alpha_{\delta,1}(u_{1,t_{i-1}} - u_{2,t_{i-1}})$. The LS orthogonality condition fails even in the very unrealistic case that market microstructure noises are mutually orthogonal white noises independent of the efficient price given that $\mathbb{E}\left[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\,\mathrm{v}_{1,t_i}\right] = -(1+\alpha_{\delta,1})\omega_1^2 - \alpha_{\delta,1}\omega_2^2$, with $\omega_m^2 \equiv \mathbb{E}(u_{m,t_i}^2) < \infty$ for $m = 1, 2$.[5] Interestingly, Appendix A.2 shows that the usual fix of including lags of the price changes in the regression only aggravates the problem. For instance, if $\alpha_{\delta,1} = 0$, the negative bias overestimates the reaction of the leading market, underestimating by how much the first market leads the price discovery.[6]

To address the error-in-variable problem that originates from the market microstructure noise, we employ instrumental variables (IV). In the presence of valid and relevant instruments, this approach yields consistent estimates even in the absence of market microstructure noise. It also makes use of the parametric structure implied by (1), being robust to serial correlation and heteroskedasticity in the residuals. Finally, it is a feasible estimator for both α and α_δ, and hence for $\alpha_\perp$ under both the infill and standard asymptotic settings. As a simple illustration, Figure 2 compares the box plots of 1,000 replications from estimates of the CS measures obtained from prices simulated from the exact discretization of (1) with $\alpha_{\delta=1/23400} = (0, 0.05)'$. We consider two scenarios, where prices are either free or contaminated with market microstructure noise. We consider iid microstructure noise with $\omega^2 = (0.001, 0.0005)'$. Noise plays a major role in the performance of the LS-based measures. The LS estimates of $\alpha_\perp$ falsely assign price discovery leadership to the second market regardless of the microstructure noise type. In contrast, the IV-based estimators correctly identify the first market as the leading one in the price discovery process. Estimating CS

[4] The result in (11) follows directly from pre-multiplying p_{t_i} by the autoregressive polynomial of the VAR(1) representation of (3) and then rearranging terms to obtain the VEC(0) model. See Hansen and Lunde (2014) for a similar result in the context of ARMA models.

[5] It is straightforward to entertain daily variation in the second moment of the market microstructure noises at the cost of very heavy notation. See Appendix A for a full characterization of the LS estimator bias.

[6] Appendix A.3 provides more details on the LS bias under different assumptions on the market microstructure noises.

from prices sample at lower frequencies help, but do not solve the consistency problem of the LS estimator.

Next, we show how to construct valid instruments to consistently estimate the standard price discovery measures (CS and IS) in the presence of market microstructure noise. The key is to constrain the correlation structure of the market microstructure noises over time and/or across markets.

Assumption MMN(TS) The market microstructure noise $u_{t_i} = \sum_{h=0}^{\bar{q}} \varpi_h \vartheta_{t_{i-h}}$ is a zero-mean invertible VMA($\bar{q}$) process, where $u_{t_i} = (u_{1,t_i}, \ldots, u_{M,t_i})'$, $\varpi_h = \text{diag}\,(\varpi_{1,h}, \ldots, \varpi_{M,h})$, $\sum_{h=0}^{\infty} \|\varpi_h\| < \infty$, and ϑ_{t_i} is a Gaussian white noise process. It then follows that u_{m,t_i} satisfies for $m = 1, \ldots, M$:

(a) $\rho_{m,m'}(q) \equiv \mathbb{E}\left(\varepsilon_{m,t_{i-q}} u_{m',t_i}\right) < \infty$ for $m, m' = 1, \ldots, M$, with $\rho_{m,m'}(q) = 0$ for $q > \bar{q}$;

(b) $\gamma_{m,m'}(q) \equiv \mathbb{E}(u_{m,t_i} u_{m',t_{i-q}}) < \infty$ for $m, m' = 1, \ldots, M$, with $\gamma_{m,m'}(q) = 0$ for all $q > \bar{q}$.

Condition (a) allows for endogenous microstructure noises in that u_{m,t_i} may correlate with P_{m,t_i} through ε_{m,t_i}. Notably, this feature is in line with the empirical stylized facts reported in Hansen and Lunde (2006). Note also the market microstructure noise remains endogenous even if $\bar{q} = 0$. Condition (b) requires the noises to exhibit neither much persistence nor much correlation across markets. Assumption MMN(TS) complies with Hansen and Lunde's (2006) Assumption 4, but it is slightly more restrictive than the conditions in Barndorff-Nielsen et al. (2011). In particular, microstructure noises that satisfy Assumption MMN(TS) automatically conform with Assumption U in Barndorff-Nielsen et al. (2011). This is convenient because it enables us to employ the realized kernel machinery to compute continuous-time IS measures in the infill asymptotic setting.

To obtain valid instruments under Assumption MMN(TS), we note that the observed price differences across markets in $\beta' p_{t_{i-1}}$ are orthogonal to their lagged values if sufficiently in the past. This is true because Assumption MMN(TS) imposes a finite dependence on the noise process. Lemma 2 summarizes the properties of the instruments based on the lagged values of $\beta' p_{t_{i-1}}$.

Lemma 2 Let $Z_{\bar{k},t_{i-\bar{q}-k}} = \beta' p_{t_{i-\bar{q}-k}}$ be a $R \times 1$ vector of instruments. Without loss of generality, let $M = 2$, it then follows under Assumptions RROU and MMN(TS) that for any $2 \leq k < \infty$

$$(i) \quad \mathbb{E}\left[Z_{\bar{k},t_{i-\bar{q}-k}}\,\mathrm{v}_{m,t_i}\right] = 0 \text{ for } m = 1, \ldots, M$$

$$(ii) \quad \mathbb{E}\left[Z_{\bar{k},t_{i-\bar{q}-k}}\,\beta' p_{t_{i-1}}\right] = \textstyle\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta \bar{\Upsilon}'_{\delta,\bar{q}+k-1+h} < \infty,$$

where $\bar{\Upsilon}_{\delta,h} = \beta'\left(I_M + \alpha_\delta\beta'\right)^h$ and in the context of $M = 2$, it is simply the difference between the first and second rows of $\Upsilon_{\delta,h}$ in (6). Because $\beta'\Xi_\delta = 0$, both results in Lemma 2 rest on the I(0) component of the GRT in (6), and one can express the endogenous regressor as an infinite summation of lagged disturbances, i.e., $\beta' p_{t_{i-1}} = \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-1-h}} + \beta' u_{t_i}$. Result (i) states that $Z_{\bar{k},t_{i-\bar{q}-k}}$ is a vector of valid instruments, which follows directly from $\rho_{m,m'}(q) = \gamma_{m,m'}(q) = 0$ for all $q > \bar{q}$. Result (ii) ensures that $Z_{\bar{k},t_{i-\bar{q}-k}}$ is a vector of relevant instruments.[7] Notably, the strength of the instruments decays as $\bar{q}$ and k increase given that

$$\lim_{h\to\infty} \Upsilon_{\delta,h} = \lim_{h\to\infty} \beta'\left(I_M + \alpha_\delta\beta'\right)^h = 0,$$

and hence $\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,\bar{q}+k-1+h}$ also shrinks to zero at the exponential rate.

It is usual in the realized measure literature to assume that the market microstructure noise is iid, even if shallow empirically. Assumption MMN(TS) is much more general, allowing for some persistence in the market microstructure noises. However, the instruments we define in Lemma 2 are not suitable to market microstructure noises that follow MA(∞) processes as in Barndorff-Nielsen et al. (2011). To deal with the latter, we have to look for instruments in the cross-section dimension—namely, other assets traded at alternative trading platforms—and hence impose some constraints to the correlation between microstructure noises across markets and assets.

To come up with valid and relevant instruments, we explicitly use the fact that the microstructure noise is asset-specific (trading frictions, informational effects, and data recording errors) and/or exchange specific (trading frictions and data recording errors), as argued by Ait-Sahalia (2007). Formally, we use a cross-section of different assets traded at alternative trading platforms as instruments. Let $S_{t_i} = (S_{1,t_i}, \ldots, S_{J,t_i})'$ denote a J-dimensional vector of discrete-time log-prices of another asset generated from a continuous-time reduced rank OU process.

Assumption OTP Let an asset trades at J alternative trading platforms and follows a reduced-

[7] See Appendix B for more details.

rank OU process

$$\mathrm{d}S_t = \widetilde{\Pi} S_t \,\mathrm{d}t + \widetilde{C}\,\mathrm{d}B_t, \tag{13}$$

where $\widetilde{\Pi} = \widetilde{\alpha}_\delta \widetilde{\beta}'$ is a $J \times J$ reduced-rank matrix with rank $J-1$, B is a $J \times 1$ vector of Brownian motions, and $\widetilde{C}$ is a $J \times J$ matrix such that the covariance matrix $\widetilde{\Sigma} = \widetilde{C}\widetilde{C}'$ is positive definite. Furthermore, the process in (13) is such that all eigenvalues of $\widetilde{\Pi}$ are real and no elementary divisor of $\widetilde{\Pi}$ occurs more than once; $\widetilde{\alpha}$ and $\widetilde{\beta}$ have full column ranks $J-1$; $\widetilde{\beta}'\widetilde{\alpha}$ has full rank $J-1$; and all eigenvalues of $\widetilde{\beta}'\widetilde{\alpha}$ have negative real parts.

The exact discretization of the reduced-rank OU process in Assumption OTP is

$$\Delta S_{t_i} = \widetilde{\alpha}_\delta \widetilde{\beta}' S_{t_{i-1}} + \epsilon_{t_i}, \tag{14}$$

where $\widetilde{\Pi}_\delta = \widetilde{\alpha}_\delta \widetilde{\beta}'$ and $\widetilde{\alpha}_\delta = \widetilde{\alpha}(\widetilde{\beta}'\widetilde{\alpha})^{-1}\left[\exp(\delta\widetilde{\beta}'\widetilde{\alpha}) - I_{J-1}\right]$, S_{t_i} is a $J \times 1$ vector of log-prices observed at time t_i, and ϵ_{t_i} is a Gaussian white noise with zero mean and covariance matrix $\widetilde{\Sigma}_\delta = \int_0^\delta \exp(u\widetilde{\Pi}t)\widetilde{\Sigma}\exp(u\widetilde{\Pi}')\,\mathrm{d}u$. It follows from Assumption OTP that the results in Lemma 1 still hold. Additionally, it is straightforward to augment the vector ΔS_{t_i} to contain $V > 1$ different assets that are traded at these J alternative trading platforms at a cost of heavy notation. In such cases, the data generation process in (14) becomes a VJ-dimensional reduced-rank OU such that there is $V(J-1)$ cointegrating vectors and V stochastic trends. This indeed is the specification we adopt in both our Monte Carlo study and empirical application.

As before, we observe only $s_{j,t_i} = S_{j,t_i} + \upsilon_{j,t_i}$, where υ_{j,t_i} is the market microstructure noise at market j, with $j = 1\ldots,J$. In particular, we assume that the market microstructure noises that plague a given asset in markets m and m', with $1 \le m \ne m' \le M$, are orthogonal to the market microstructure noises that affect the prices of another asset in markets j and j', with $1 \le j \ne j' \le J$.

Assumption MMN(CS) The microstructure noises $u_{t_i} = (u_{1,t_i},\ldots,u_{M,t_i})'$ and $\upsilon_{t_i} = (\upsilon_{1,t_i},\ldots,\upsilon_{J,t_i})'$ are zero-mean MA(∞) processes given by $u_{t_i} = \sum_{h=0}^{\infty}\varpi_h\vartheta_{t_{i-h}}$ and $\upsilon_{t_i} = \sum_{h=0}^{\infty}\varrho_h\varphi_{t_{i-h}}$, where $\varpi_h = \mathrm{diag}\,(\varpi_{1,h},\ldots,\varpi_{M,h})$, $\varrho_h = \mathrm{diag}\,(\varrho_{1,h},\ldots,\varrho_{J,h})$ $\sum_{h=0}^{\infty}\|\varpi_h\| < \infty$, $\sum_{h=0}^{\infty}\|\varrho_h\| < \infty$, and both ϑ_{m,t_i} and φ_{m,t_i} are Gaussian white noise processes. The microstructure noises satisfy for $m = 1,\ldots,M$ and $j = 1,\ldots,J$:

(a) $\mathbb{E}\left(u_{m,t_i}\upsilon_{j,t_{i-q}}\right) = 0$ for $q \ge 0$ with $m = 1,2,\ldots,M$ and $j = 1,\ldots,J$;

(b) $\mathbb{E}\left(\varepsilon_{m,t_{i-q}} u_{m',t_i}\right) < \infty$ for $q \geq 0$ with $m, m' = 1, 2, \ldots, M$;

(c) $\mathbb{E}\left(\varepsilon_{m,t_{i-q}} \upsilon_{j,t_i}\right) = 0$ for $q \geq 0$ with $m = 1, 2, \ldots, M$ and $j = 1, \ldots, J$.

Condition (a) rules out correlation between market microstructure noises of different assets in different trading platforms. Condition (b) allows the market microstructure noise of the asset of interest to correlate with the efficient log-price through the idiosyncratic component in (3), whereas (c) dictates that the log-price of the asset of interest in market m, with $1 \leq m \leq M$, is orthogonal to the market microstructure of the other asset in markets $j = 1, \ldots, J$. We are now ready to define a set of valid and relevant instruments.[8]

Lemma 3 Let Assumptions RROU, OTP and MMN(CS) hold. Let $Z_{t_{i-1}} \equiv \widetilde{\beta}' s_{t_{i-1}}$ correspond to the vector of instruments. Without loss of generality, let $M = J = 2$, it then follows that,

(i) $\mathbb{E}\left[Z_{t_{i-1}} \mathrm{v}_{m,t_i}\right] = 0$ for $m = 1, \ldots, M$

(ii) If $\mathbb{E}\left[\epsilon_{j,t_i} \varepsilon_{m,t_i}\right] \neq \mathbb{E}\left[\epsilon_{j',t_i} \varepsilon_{m',t_i}\right]$ for some $j, j' = 1, \ldots, J$ with $j \neq j'$ and $m, m' = 1, \ldots, M$ with $m \neq m'$, then $\mathbb{E}\left[Z_{t_{i-1}} \beta' p_{t_{i-1}}\right] = \sum_{h=0}^{\infty} \widetilde{\beta}' \widetilde{\Upsilon}_{\delta,h} \mathbf{\Sigma}_{\delta}(0) \Upsilon'_{\delta,h} \beta < \infty$,

where $\mathbf{\Sigma}_{\delta}(0) = \mathbb{E}\left[\epsilon'_{t_i} \otimes \varepsilon_{t_i}\right]$ and $\widetilde{\Upsilon}_{\delta,h}$ denotes the parameter matrices of the I(0) component of the GRT of (14).

Result *(i)* in Lemma 3 states $\widetilde{\beta}' s_{t_{i-1}}$ is orthogonal to $\beta' P_{t_{i-1}}$ and hence $\widetilde{\beta}' s_{t_{i-1}}$ is a valid instrument under Assumption MMN(CS). Similarly to Lemma 2, relevance of the instruments rests on the I(0) component of the GRT and some degree of heterogeneity on the correlation among assets innovations. Specifically, we expect assets innovations to correlate because they should respond to market wide news that are impounded differently across assets and trading platforms.

Chief to the estimation and limiting theory discussed in the following sections is the type of dependence associated with the moment conditions stated in Lemmas 2 and 3. Specifically, both $Z_{\bar{k},t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}$ and $Z_{t_{i-1}} \mathrm{v}_{m,t_i}$ sequences are not martingale differences sequences, but (potentially infinitely) correlated processes. This implies that the central limit (CLT) theorems for martingale differences sequences or iid MA(∞) processes no longer hold. Alternatively, we exploit the feature

[8] See Appendix B for more details.

that these moment conditions are asymptotic unpredictable, meaning that the c-period-ahead-forecast converge in absolute expected value to zero as $c \to \infty$. Sequences that satisfy this property are called L^1-mixingale sequences.[9] Lemma 4 formalizes these results.

Lemma 4 Let $\beta' p_{t_{i-\bar{q}-\underline{k}}}$ with $\underline{k}$ satisfying $2 \leq k \leq \underline{k} \leq \bar{k}$ and $\widetilde{\beta}' s_{t_{i-1}}$ be vectors of instruments under Assumptions MMN(TS) and MMN(CS), respectively.

(i) $\left\{\beta' p_{t_{i-\bar{q}-\underline{k}}}, \widetilde{\beta}' s_{t_{i-1}}, \mathrm{v}_{m,t_i}\right\}$ is a stationary ergodic sequence;

(ii) $\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\}$ and $\left\{\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ are stationary ergodic adapted L^1-mixingale sequences, where $\beta'_{\underline{m}}$, $\widetilde{\beta}'_{\underline{j}}$ and v_{m,t_i} denote the $\underline{m}$-, $\underline{j}$-, and m-th rows of β', $\widetilde{\beta}'$, and v_{t_i}, respectively, with $\underline{m}, m = 1, \ldots, M$, $\underline{j} = 1, \ldots, J$, and $\mathcal{F}_{t_i}$ is the σ-algebra generated by the entire current and past history of p_{t_i} and s_{t_i};

(iii) $\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\}$ and $\left\{\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ are uniformly integrable.

Lemma 4 essentially shows that the moment conditions in both set of assumptions are stationary ergodic uniformly integrable L^1-mixingale sequences, which in turn allows us to derive the limiting distribution of the IV-based price discovery measures.

3.1 Asymptotic theory with fixed sampling intervals

We first focus on the case in which we estimate the price discovery measures over long time spans, fixing the sampling frequency. This means that we treat the number of intraday observations n as fixed, while allowing the number of days D in the sample to diverge. Because our sample consists of prices observed intra-daily over different days, we denote the overall number of observations by $T = nD$, which grows without bound as $D \to \infty$.

It follows from the discrete-time process in (3) that, once we account for the market microstructure noise, the dynamics of the observed log-prices at a fixed sampling interval δ is given by

$$\Delta p_\tau = \alpha_\delta \beta' p_{\tau-1} + \mathrm{v}_\tau, \qquad \text{with } \tau = 1, \ldots, T = nD \tag{15}$$

or, in a more compact notation,

$$\Delta \boldsymbol{p} = \bar{\boldsymbol{p}} \boldsymbol{\alpha}_\delta + \mathbf{v}, \tag{16}$$

[9] A sequence $\{\mathcal{Z}_t, \mathcal{F}_{t_i}\}_{-\infty}^{\infty}$ is said to be a L^1-mixingale with respect to $\mathcal{F}_{t_i}$ if $\mathbb{E}\left\|\mathbb{E}\left(\mathcal{Z}_t | \mathcal{F}_{t_{i-c}}\right)\right\| \leq f_{t_i} \xi_c$, where $\{f_{t_i}, \xi_c\}$ are deterministic sequences and $\lim_{c\to\infty} \xi_c = 0$ (see, for instance, Definition 16.1 in Davidson, 1994)

where $\Delta \boldsymbol{p} = \text{vec}\,(\Delta p_1, \ldots, \Delta p_M)$ is an $M(T-1) \times 1$ matrix with $\Delta p_m = (\Delta p_{m,2}, \ldots, \Delta p_{m,T})'$ for $m = 1, 2, \ldots, M$; $\bar{\boldsymbol{p}}$ is a $M(T-1) \times MR$ block diagonal matrix with M diagonal blocks given by the $(T-1) \times R$ matrix $\boldsymbol{p}_{-1}'\beta$, where $\boldsymbol{p}_{-1} = (p_1, \ldots, p_{T-1})$ is a $M \times T-1$ matrix; $\boldsymbol{\alpha}_\delta$ is a $MR \times 1$ vector that stacks the M rows of α_δ, such that $\boldsymbol{\alpha}_\delta = \text{vec}\,(\alpha_\delta')$; and $\mathbf{v} = \text{vec}\,(\mathrm{v}_1, \ldots, \mathrm{v}_M)$ denotes a $M(T-1) \times 1$ matrix collecting the disturbance terms with $\mathrm{v}_m = (\mathrm{v}_{m,2}, \ldots, \mathrm{v}_{m,T})'$ for $m = 1, 2, \ldots, M$. Furthermore, from the exact discretization of of the reduced-rank OU process in Assumption OTP and market microstructure noise contamination, we rewrite the J-dimensional price process as

$$\Delta s_\tau = \widetilde{\alpha}_\delta \widetilde{\beta}' s_{\tau-1} + \bar{\epsilon}_\tau, \qquad \text{with} \;\; \tau = 1, \ldots, T = nD. \tag{17}$$

Let $\boldsymbol{Z}$ denote a $(T-1) \times \bar{r}$ matrix of valid and relevant instruments with $\bar{r} \geq R$. The IV estimator is equivalent to the two-stage least-squares (2SLS) estimator that replace the endogenous regressor by its projection on the subspace spanned by the instruments. Because instruments are valid, it is straightforward to use GMM framework to derive the IV-based estimator, with sample moment conditions and weighing matrix given by $(I_M \otimes \boldsymbol{Z})' \, \mathbf{v}$ and $\boldsymbol{W} = \left[(I_M \otimes \boldsymbol{Z})' \, (I_M \otimes \boldsymbol{Z})\right]^{-1}$, respectively. It then follows that the IV-based estimator reads

$$\widehat{\boldsymbol{\alpha}}_{\delta,IV} = \left[\bar{\boldsymbol{p}}' \, (I_M \otimes \boldsymbol{Z}) \, \boldsymbol{W} \, (I_M \otimes \boldsymbol{Z})' \, \bar{\boldsymbol{p}}\right]^{-1} \bar{\boldsymbol{p}}' \, (I_M \otimes \boldsymbol{Z}) \, \boldsymbol{W} \, (I_M \otimes \boldsymbol{Z})' \, \Delta \boldsymbol{p}, \tag{18}$$

with a direct mapping $\boldsymbol{\alpha}_\delta \mapsto \alpha_\delta$ given by $\widehat{\alpha}_{\delta,IV} = \left(I_M \otimes \widehat{\boldsymbol{\alpha}}_{\delta,IV}'\right) (\text{vec}\,(I_M) \otimes I_R)$. Although (18) is valid for instruments with either Assumption MMN(TS) or MMN(CS) in mind, the dimensions of $\Delta \boldsymbol{p}$ and $\boldsymbol{Z}$ differ from one case to the other. Under Assumption MMN(TS), for instance, instruments correspond to lagged values of $\beta' p_{\tau-1}$ and hence $\Delta \boldsymbol{p}$ and $\boldsymbol{Z} = \left(Z_{\bar{k},1}, \ldots, Z_{\bar{k},T-\bar{q}-\bar{k}}\right)'$ have dimensions $M(T - \bar{q} - \bar{k}) \times 1$ and $(T - \bar{q} - \bar{k}) \times (\bar{k} - k + 1)R$, respectively.

In what follows, we derive the consistency and asymptotic normality of the IV-based price discovery measures under either Assumption MMN(TS) or MMN(CS). Consistent estimation of CS requires $\widehat{\alpha}_{\delta,IV} \xrightarrow{p} \alpha_{\delta,IV}$ for any δ given that $\text{CS} = \alpha_{\delta,\perp} = \alpha_\perp$. In addition, asymptotic normality of the CS estimator readily follows from the limiting distribution of $\widehat{\boldsymbol{\alpha}}_{\delta,IV}$. As for the consistent estimation of the IS measure in discrete time, it calls for $\widehat{\Sigma}_{\delta,IV} \xrightarrow{p} \Sigma_\delta$ in addition to $\widehat{\alpha}_{\perp,\mathrm{IV}} \xrightarrow{p} \alpha_\perp$. Because of the serial correlation in v_{t_i} due to the market microstructure noise, we estimate Σ_δ using a heteroskedasticity and autocorrelation consistent (HAC) estimator. It follows that the IV-based estimate of the discrete-time IS measure reads

$$\widehat{IS}_{\delta,IV,m} = \frac{[\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{C}_\delta]^2_m}{\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{\Sigma}_\delta\ \widehat{\alpha}'_{\perp,\mathrm{IV}}}, \qquad m = 1, \ldots, M. \tag{19}$$

Consistency of the IV-based residuals, $\widehat{\mathbf{v}} = \boldsymbol{\Delta p} - \bar{\boldsymbol{p}}\widehat{\boldsymbol{\alpha}}_{\delta,IV}$, implies that the HAC estimation should suffice because the long-run covariance matrix to which it converges already accounts for the spurious autocorrelation patterns that may arise from microstructure effects.

As for the estimation of the continuous-time IS measure, we can make use of the exact discretization of (1) to map an estimate of Σ from $\widehat{\Sigma}_\delta$ computed from the IV-based residuals $\widehat{\mathbf{v}}$. Specifically, we also estimate $\widehat{\Sigma}_\delta$ using the HAC estimator as in the case of the discrete-time IS measure, but we now use $\widehat{\Sigma}_\delta$ to solve $\widehat{\Sigma}_\delta = \int_0^\delta \exp\left(u\widehat{\Pi}_{\mathrm{IV}}\right)\widehat{\Sigma}\exp\left(u\widehat{\Pi}'_{\mathrm{IV}}\right)\mathrm{d}u$ for $\widehat{\Sigma}$ with $\widehat{\Pi} = \Pi\delta^{-1}\log\left(\widehat{\alpha}_{\delta,\mathrm{IV}}\,\beta' + I_M\right)$. Because $\alpha_{\delta,\perp} = \alpha_\perp$ for any $0 < \delta < 1$, it follows that an estimate of the continuous-time IS measure based in discrete sampled prices reads

$$\widehat{IS}_{IV,m} = \frac{[\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{C}]^2_m}{\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{\Sigma}\ \widehat{\alpha}'_{\perp,\mathrm{IV}}}, \qquad m = 1, \ldots, M, \tag{20}$$

where $\widehat{\Sigma}$ is the HAC-based estimator of Σ that is mapped from $\widehat{\Sigma}_\delta$, and $\widehat{C}$ denotes a decomposition of $\widehat{\Sigma}$ such that $\widehat{C}\widehat{C}' = \widehat{\Sigma}$.

Theorem 1 formalizes these results for both noise settings.

Theorem 1 Let the conditions in Lemma 2 or 3 hold, and consider their respective vector of instruments. It then follows that, as $T = nD \to \infty$ with fixed n,

(i) $\widehat{\alpha}_{\perp,\mathrm{IV}} \xrightarrow{p} \alpha_\perp$;

(ii) $\sqrt{T}\left(\widehat{\alpha}_{\perp,\mathrm{IV}} - \alpha_\perp\right) \xrightarrow{d} N\left(0, \boldsymbol{C}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}\boldsymbol{C}'\right)$,

where $\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}} = \sum_{j=-\bar{q}-1}^{\bar{q}+1}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j)$ and $\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}} = \sum_{j=-\infty}^{\infty}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j)$ under Assumptions MMN(TS) and MMN(CS), respectively, with $\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j) = \mathbb{E}\left[\mathrm{v}_\tau\mathrm{v}'_{\tau-j}\otimes Z_{\tau-k^*}Z'_{\tau-k^*-j}\right]$, $\mathrm{v}_\tau = (\mathrm{v}_{1,\tau}, \ldots, \mathrm{v}_{M,\tau})'$, and $1 \le k^* < \infty$ denoting the appropriate lag order for the instruments; $\boldsymbol{C} = \boldsymbol{C_1}\boldsymbol{C_2}\boldsymbol{C_3}$, with $\boldsymbol{C_1} = \mathcal{S}\left(\psi'\otimes\Xi\right)\boldsymbol{K}_{M,R}$, $\boldsymbol{C_2} = \left[\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}\boldsymbol{Q}'_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\right]^{-1}$, and $\boldsymbol{C_3} = \boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}$; $\mathcal{S}$ is an $M \times M^2$ deterministic matrix that selects the first row of Ξ from $\mathrm{vec}(\Xi)$; $\psi = \bar{\alpha}_\delta\left(\Xi - I_M\right)$ with $\bar{\alpha}_\delta = -\alpha_\delta\left(\alpha'_\delta\alpha_\delta\right)^{-1}$; $\boldsymbol{K}_{M,R}$ is the commutation matrix such that $\boldsymbol{K}_{M,R}\mathrm{vec}(\alpha_\delta) = \boldsymbol{\alpha}_\delta$; and $\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}} = \operatorname*{plim}_{T\to\infty}\frac{1}{T}\bar{\boldsymbol{p}}'\left(I_M\otimes\boldsymbol{Z}\right)$ and $\boldsymbol{Q}_{\boldsymbol{W}} = \operatorname*{plim}_{T\to\infty}\frac{1}{T}\boldsymbol{W}$;

(iii) $\widehat{IS}_{\delta,IV,m} \xrightarrow{p} IS_{\delta,m}$;

(iv) $\widehat{IS}_{IV,m} \xrightarrow{p} IS_{m}$;

Under Assumption MMN(TS), the vector of instruments $Z_{\bar{k},\tau-\bar{q}-k}$ is strong for small $\bar{q}$ and k. Furthermore, simulations show that the empirical densities of the t-statistics based on Theorem 1(ii) are close to normal even when we sample one observation per second within one trading day ($D = 1$ and $n = 23,400$). This is true regardless of the the variance of the market microstructure noise (see upper panel in Figure 1).

Under Assumptions MMN(CS) and OTP, the vector of instruments $Z_{\tau-1} \equiv \widetilde{\beta}' s_{\tau-1}$ is likely not very strong in practise, so that $\widehat{\alpha}_{\delta,IV}$ might become biased for moderate-to-high values of ω^2 (see simulations and discussion in Section 4.2). If instruments are indeed weak, the standard asymptotic theory for the IV estimator fails (see, for instance, Bound, Jaeger and Baker, 1995; Staiger and Stock, 1997; Stock and Yogo, 2002, 2005a). One possible solution is to add more instruments, that is to say, to consider a large cross-section of different assets traded at alternative exchanges, hoping to improve the instruments' relevance: $Z_{\tau-1} \equiv \beta_{VJ}' s_{\tau-1}$, where β_{VJ} denotes a $VJ \times V(J-1)$ cointegration matrix and $s_{\tau-1}$ is now a VJ-dimensional vector collecting the observed prices from V different assets at J trading platforms. Although the 2SLS estimator remains inconsistent under many weak instruments, with asymptotic bias increasing with the number of instruments (Stock and Yogo, 2002, 2005b), one may employ the GMM-based continuous updating (CU-GMM) estimator of Hansen, Heaton and Yaron (1996), which is robust to (many) weak instruments.

In the context of weak instruments, CU-GMM estimators are not useful for inference given that their asymptotic distribution depends on nuisance parameters. Stock and Wright (2000) nonetheless show how to carry out hypothesis testing and construct asymptotically valid confidence sets directly from the objective function. Specifically, at the true parameter values, the objective function has a standard χ^2 distribution under mild conditions. Accordingly, one may obtain asymptotically valid confidence sets by inverting the test of the null $\alpha_{\delta,m} = 0$ for $m = 1, \ldots, M$. Null hypotheses of this kind are informative in the price discovery context of (3) with $M = 2$, because they are equivalent to testing for the strong exogeneity of one market with respect to the cointegrating vector β. Both hypothesis tests and confidence sets have asymptotic size and coverage equal to their nominal levels uniformly over the entire parameter space.

In the case of many weak instruments, the CU-GMM estimator is consistent under heteroskedastic and serially correlated moment conditions for properly chosen weighting matrices (Han and Phillips, 2006; Newey and Windmeijer, 2009; Hausman, Lewis, Menzel and Newey, 2011). In particular, consistency requires $V^2(J-1)^2/T \to 0$, whereas asymptotic normality requires the number of instruments $V(J-1)$ to satisfy $V^3(J-1)^3/T \to 0$. The many weak instruments setting fits well our framework because it is straightforward to let V grow arbitrarily large by using a cross-section of (actively) traded stocks at other liquid trading platforms and by augmenting the set of instruments with lagged values of $\widetilde{\beta}' s_{\tau-1}$, such that $Z_{\bar{k},\tau-1} = \text{vec}\left(\widetilde{\beta}' s_{\tau-1}, \ldots, \widetilde{\beta}' s_{\tau-1-\bar{k}}\right)$ with $0 \leq \bar{k} < \infty$. The latter is not completely irrelevant in view that assets presumably impound information at different paces (Scherrer, 2013). The price to pay for employing the CU-GMM estimator is that we cannot guarantee that it has finite moments of any order. This may affect finite-sample performance, leading to wide dispersion (Hausman et al., 2011).

Denote by $\widehat{\boldsymbol{\alpha}}_{\delta,\text{CU-GMM}}$ the CU-GMM estimator that minimizes the objective function $\mathcal{C}_{\text{CU-GMM},T}(\widehat{\boldsymbol{\alpha}}_\delta) = -\boldsymbol{g}_T(\widehat{\boldsymbol{\alpha}}_\delta)' \widehat{\Psi}_\delta^{-1}(\widehat{\boldsymbol{\alpha}}_\delta) \boldsymbol{g}_T(\widehat{\boldsymbol{\alpha}}_\delta)$, where $\boldsymbol{g}_T(\widehat{\boldsymbol{\alpha}}_\delta) = \frac{1}{T-1-\bar{k}} \left(I_M \otimes \boldsymbol{Z}\right)' \widehat{\mathbf{v}}$ is a $M\big(V(J-1)\big) \times 1$ vector containing the average moment conditions over time with $\boldsymbol{Z} = (Z_1, \ldots, Z_{T-1-\bar{\kappa}})'$, and $\widehat{\Psi}_\delta(\widehat{\boldsymbol{\alpha}}_\delta)$ is an long-run variance estimator of the the $MV(J-1) \times MV(J-1)$ asymptotic variance of $\sqrt{T}\boldsymbol{g}_T(\boldsymbol{\alpha}_\delta)$: $\Psi_\delta(\boldsymbol{\alpha}_\delta) = \sum_{j=-\infty}^{\infty} \Gamma_{\boldsymbol{g}_T}(j)$. Note that the CU-GMM objective function parametrizes the long-run variance as a function of $\widehat{\boldsymbol{\alpha}}_\delta$ via $\widehat{\mathbf{v}} = \Delta \boldsymbol{p} - \bar{\boldsymbol{p}} \widehat{\boldsymbol{\alpha}}_\delta$.

For the estimation of $\widehat{\Psi}_\delta(\widehat{\boldsymbol{\alpha}}_\delta)$, we use Andrews's (1991) HAC estimator with a Parzen kernel to control for the serial correlation in v_τ. Similarly to Theorem 1, consistency of the CS and IS estimators follows directly from $\widehat{\boldsymbol{\alpha}}_{\delta,\text{CU-GMM}} \xrightarrow{p} \boldsymbol{\alpha}_\delta$. Section 4 provides evidence that the CU-GMM estimator indeed solves the weak instruments issue under Assumption MMN(CS), with median biases very close to zero and decreasing with V for different levels of market microstructure noise.

3.2 Infill asymptotic theory

This section addresses the case in which $n \to \infty$, while D is fixed (say, one trading day). As before, we back out CS and IS measures from the estimates of the speed-of-adjustment parameters of the daily VEC models. Differently from Section 3.1, we now let $\delta \to 0$ to ensure that $n \to \infty$. Although $\alpha_\perp = \alpha_{\delta,\perp}$ for any $0 < \delta < 1$, consistency of $\alpha_\perp$ now requires consistent estimation of α

given that α_δ is no longer a fixed quantity. Similarly, consistency of the continuous-time IS calls for the consistent estimation of the quadratic covariation.

As we let $\delta \to 0$, it is necessary to account for nonsynchronous trading. To this end, we synchronize tick data using the refresh-time scale. Suppose we observe the price in market m at times $t_{m,1}, t_{m,2}, \ldots$ and let $N_m(t)$ denote the counting process with the number of observations up to time t in market m, with $m = 1, \ldots, M$. The first refresh time is $\tau_1 = \max_{1\leq m\leq M} t_{m,1}$, whereas the subsequent refresh times are $\tau_{i+1} = \max_{1\leq m\leq M} t_{m,N_j(\tau_i)+1}$ (see more details in Barndorff-Nielsen et al., 2011). To comply with the necessary conditions for the consistent estimation of the quadratic covariation, we also carry out some jittering to the first and last prices of the day as in Jacod, Li, Mykland, Podolskij and Vetter (2009) and Barndorff-Nielsen et al. (2011).

Assumption RT Define the durations between observation times as $\Delta_i \equiv \tau_i - \tau_{i-1} = \mathcal{A}_i/n$ for every i. We assume that $\max_{i'+1\leq i\leq i'+K} D_i = o_p(\sqrt{K})$ for any i' and that $\tau_0 \leq 0$ and $\tau_{m+1} \geq 1$. In addition, $\mathbb{E}\left(\mathcal{A}^r_{\lfloor nt\rfloor} \,\middle|\, \mathcal{F}_{\tau_{\lfloor nt\rfloor-1}}\right) \xrightarrow{p} \chi_r(t)$ for $0 < r \leq 2$ as $n \to \infty$, where $\chi_r(t)$ is a strictly positive, càdlàg process adapted to $\mathcal{F}_t$.

Assumption RT essentially allows us to treat irregularly space prices as equally spaced in refresh time. The rate on $\mathcal{A}_i$ is the same as in Phillips and Yu (2008). It is more restrictive that the $O_p(1)$ rate from earlier works, such as Jacod (n.d.), Mykland and Zhang (2006) and Barndorff-Nielsen, Hansen, Lunde and Shephard (2008), because it allows for random times. Empirically, the refresh-time scheme is heavily dependent on the least liquid asset/market given that one forms a new price tuple only after every price refreshes. This means that effective sample size decreases with the number of prices and with non-synchronicity. Accordingly, this is especially a problem under Assumption MMN(CS). In this case, we must synchronize VJ prices, where V is large and the J alternative exchanges are usually less liquid than the M main markets.

To illustrate this issue, let time stamps come from independent Bernoulli trials, with $V = 25$ and $J = 3$, so that each price process contains 9,000 observations per day, on average. Applying refresh time to synchronize these prices yields, on average, about 2,200 observations per day, with an average price duration of 10.6 seconds. The low levels of data retention in this setting affects negatively the performance of the CU-GMM estimator as we necessitate very large sample sizes to extract enough information from weak instruments and the noise-to-signal ratio increases with the average duration. For this reason, we restrict attention to the infill asymptotic theory under Assumption MMN(TS).

Next, we focus on constructing an estimator for the continuous-time speed-of-adjustment parameter α. In order to reflect Assumption RT, it is convenient to adapt our notation and rewrite (15) and (16) as follows

$$\Delta p_{\tau_i} = \alpha_\delta\left(\boldsymbol{\alpha}\right)\beta' p_{\tau_{i-1}} + \mathrm{v}_{\tau_i}, \qquad \text{with} \;\; i = 1, \ldots, n \tag{21}$$

$$\Delta \boldsymbol{p} = \bar{\boldsymbol{p}}\boldsymbol{\alpha}_\delta\left(\boldsymbol{\alpha}\right) + \mathbf{v}, \tag{22}$$

where the subscript τ now refers to the refresh-time clock and n is the resulting refresh-time sample size; $\alpha_\delta\left(\boldsymbol{\alpha}\right)$ is a function of the continuous-time speed of adjustment parameter α, such that $\alpha_\delta\left(\alpha\right) = \alpha\left(\beta'\alpha\right)^{-1}\left[\exp(\delta\beta'\alpha) - I_R\right]$, $\boldsymbol{\alpha}_\delta\left(\boldsymbol{\alpha}\right) = \mathrm{vec}\left(\alpha'_\delta\left(\boldsymbol{\alpha}\right)\right)$, and $\boldsymbol{\alpha} = \mathrm{vec}\left(\alpha'\right)$; $\Delta p_{\tau_i} = \left(\Delta p_{1,\tau_i}, \ldots, \Delta p_{M,\tau_i}\right)'$ is a $M \times 1$ vector and $\Delta \boldsymbol{p} = \mathrm{vec}\left(\Delta p_1, \ldots, \Delta p_M\right)$ is a $M(n-1) \times 1$ matrix with $\Delta p_m = \left(\Delta p_{m,2}, \ldots, \Delta p_{m,n}\right)'$ being a $(n-1) \times 1$ vector for $m = 1, 2, \ldots, M$; $p_{\tau_{i-1}} = \left(p_{1,\tau_{i-1}}, \ldots, p_{M,\tau_{i-1}}\right)'$ is a $M \times 1$ vector and $\bar{\boldsymbol{p}}$ is a $M(n-1) \times MR$ block diagonal matrix with M diagonal blocks given by the $(n-1) \times R$ matrix $\boldsymbol{p}'_{-1}\beta$, where $\boldsymbol{p}_{-1} = (p_1, \ldots, p_{n-1})$ is a $m \times n-1$ matrix; and $\mathrm{v}_{\tau_i} = \left(\mathrm{v}_{1,\tau_i}, \ldots, \mathrm{v}_{M,\tau_i}\right)'$ is a $M \times 1$ vector and $\mathbf{v} = \mathrm{vec}\left(\mathrm{v}_1, \ldots, \mathrm{v}_M\right)$ denotes a $M(n-1) \times 1$ matrix collecting the disturbance terms with $\mathrm{v}_m = \left(\mathrm{v}_{m,2}, \ldots, \mathrm{v}_{m,n}\right)'$ being a $(n-1) \times 1$ vector for $m = 1, 2, \ldots, M$. Recall from Lemma 1 that, in the absence of market microstructure noise, the moment condition implied by $\ell^c_n(\alpha, \Sigma)$ is $\mathbb{E}\Big(\left(\Delta P_{\tau_i} - \alpha_\delta\left(\boldsymbol{\alpha}\right)\beta' P_{\tau_{i-1}}\right) \otimes \left(\beta' P_{\tau_{i-1}}\right)\Big) = 0$ with ΔP_{τ_i} and P_{τ_i} respectively denoting the noiseless counterparts of Δp_{τ_i} and $p_{\tau_{i-1}}$ in (21). It then follows that $\mathbb{E}\Big(\left(\Delta P_{\tau_i} - \alpha_\delta\left(\boldsymbol{\alpha}\right)\beta' P_{\tau_{i-1}}\right) \otimes \left(\beta' P_{\tau_{i-1-\bar{\kappa}}}\right)\Big) = 0$ for all $\bar{\kappa} > 0$, so that we may estimate the continuous-time speed-of-adjustment parameters using a set of valid moment conditions with $\bar{\kappa} > 0$ that matches the requirements in Assumption MMN(TS).

To formally define the GMM criterion function, let $Z_{\bar{k},\tau_{i-\bar{q}-k}} \equiv \mathrm{vec}\left(\beta' p_{\tau_{i-\bar{q}-k}}, \ldots, \beta' p_{\tau_{i-\bar{q}-\bar{k}}}\right)$ with integers k and $\bar{k}$ such that $2 \leq k \leq \bar{k} < \infty$ and then define $\boldsymbol{Z} = \left(Z_{\bar{k},\tau_1}, \ldots, Z_{\bar{k},\tau_{n-\bar{q}-\bar{k}}}\right)'$ as the $(n-\bar{q}-\bar{k}) \times (\bar{k}-k+1)R$ matrix of instruments. It then follows that the $M(\bar{k}-k+1)R$ sample moment conditions read

$$\boldsymbol{g}_{n_\tau}(\widehat{\boldsymbol{\alpha}}) = \frac{1}{(n-\bar{q}-\bar{k})}\left(I_M \otimes \boldsymbol{Z}\right)'\widehat{\mathbf{v}} = \frac{1}{(n-\bar{q}-\bar{k})}\sum_{i=\bar{q}+k+(\bar{k}-k)+1}^{n}\left(\widehat{\mathrm{v}}_{\tau_i} \otimes Z_{\bar{k},\tau_{i-\bar{q}-k}}\right), \tag{23}$$

where $\widehat{\mathbf{v}} = \Delta\boldsymbol{p} - \bar{\boldsymbol{p}}\boldsymbol{\alpha}_\delta\left(\widehat{\boldsymbol{\alpha}}\right)$ and $\widehat{\mathrm{v}}_{\tau_i} = \Delta p_{\tau_i} - \alpha_\delta\left(\widehat{\boldsymbol{\alpha}}\right)\beta' p_{\tau_{i-1}}$. As for the optimal weighting matrix, we choose $\widehat{\Psi}$ to be the long-run variance estimator of the the $M(\bar{k}-k+1)R\times M(\bar{k}-k+1)R)$ asymptotic variance of $\sqrt{n}\boldsymbol{g}_{n_\tau}(\boldsymbol{\alpha})$: $\Psi_\tau = \sum_{j=-\bar{q}-1}^{\bar{q}+1}\Psi_\tau(j)$, where $\Psi_\tau(j) = \mathbb{E}\left[\mathrm{v}_{\tau_i}\mathrm{v}'_{\tau_{i-j}} \otimes Z_{\bar{k},\tau_{i-\bar{q}-k}} Z'_{\bar{k},\tau_{i-\bar{q}-k-j}}\right]$. The GMM criterion function then reads $\mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}}) = -\boldsymbol{g}_{n_\tau}(\widehat{\boldsymbol{\alpha}})'\widehat{\Psi}^{-1}\boldsymbol{g}_{n_\tau}(\widehat{\boldsymbol{\alpha}})$. As a result, the IV estimator of $\boldsymbol{\alpha}$ is the maximizer that solves the nonlinear optimization problem

$$\widehat{\boldsymbol{\alpha}}_{\mathrm{IV}} = \underset{\widehat{\boldsymbol{\alpha}}}{\operatorname{argmax}}\ \mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}}), \tag{24}$$

with $\widehat{\alpha}_{IV} = \left[\left(\mathrm{vec}\left(I_M\right)' \otimes I_R\right)\left(I_M \otimes \widehat{\boldsymbol{\alpha}}_{\mathrm{IV}}\right)\right]'$.

To establish the large sample properties of the CS measure, it is necessary to show the consistency and asymptotic normality of $\widehat{\boldsymbol{\alpha}}_{\mathrm{IV}}$. Assumption GMM provides primitive conditions.

Assumption GMM Consider $\widehat{\boldsymbol{\alpha}}$ an arbitrary vector of parameter satisfying Assumption RROU. We assume that

(a) $\widehat{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathbf{A}$, where $\mathbf{A} \subset \mathbb{R}^{MR}$ is a compact set;

(b) $\mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}})$ is continuous in $\widehat{\boldsymbol{\alpha}} \in \mathbf{A}$ for all possible samples and is a measurable function of the data for all $\widehat{\boldsymbol{\alpha}} \in \mathbf{A}$;

(c) $\sup_{\widehat{\boldsymbol{\alpha}}\in\mathbf{A}}\left|\mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}}) - \mathcal{C}(\widehat{\boldsymbol{\alpha}})\right| \xrightarrow{p} 0$ and $\sup_{\widehat{\boldsymbol{\alpha}}\in\mathbf{A}}\left|\mathcal{C}_{\mathrm{GMM},n_t}(\widehat{\boldsymbol{\alpha}}) - \mathcal{C}(\widehat{\boldsymbol{\alpha}})\right| \xrightarrow{p} 0$, where $\mathcal{C}_{\mathrm{GMM},n_t}(\widehat{\boldsymbol{\alpha}})$ denotes the GMM criterion function that is based on the latent equally spaced clock with sampling interval given by $\delta = 1/n$ and $\mathcal{C}(\widehat{\boldsymbol{\alpha}})$ is a nonstochastic function;

(d) $\mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}})$ $\mathcal{C}_{\mathrm{GMM},n_t}(\widehat{\boldsymbol{\alpha}})$ are continuously differentiable on the interior of $\mathbf{A}$. Additionally, suppose that $\frac{\partial^2\mathcal{C}_{\mathrm{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}})}{\partial\widehat{\boldsymbol{\alpha}}\partial\widehat{\boldsymbol{\alpha}}'}$ and $\frac{\partial^2\mathcal{C}_{\mathrm{GMM},n_t}(\widehat{\boldsymbol{\alpha}})}{\partial\widehat{\boldsymbol{\alpha}}\partial\widehat{\boldsymbol{\alpha}}'}$ exist in a continuous and open neighborhood of $\boldsymbol{\alpha}$;

We require conditions (a)–(c) to obtain uniform convergence of the sample moments to their population counterparts and ultimately to $\widehat{\boldsymbol{\alpha}}_{\mathrm{IV}} \xrightarrow{p} \boldsymbol{\alpha}$. Alternatively, we could adapt results in Theorem 6.9 in White (2000) to obtain $\widehat{\Psi} \xrightarrow{p} \Psi$ and Theorem 2.2 in Phillips and Yu (2008) to show pointwise convergence of $\mathcal{C}_{\mathrm{GMM},n_\tau}$ to $\mathcal{C}_{\mathrm{GMM},n_t}$. Assuming the existence of Lipschitz-type conditions on $\mathcal{C}_{\mathrm{GMM},n_\tau}$ and $\mathcal{C}_{\mathrm{GMM},n_t}$ would suffice to deliver stochastic equicontinuity (see, for instance, Theorem 21.10 in Davidson, 1994) and hence the uniform convergence assumed Assumption in GMM (c). Condition (d) is standard in the nonlinear GMM framework (see, e.g., Newey and McFadden, 1994; Amemiya, 1985). Note that Lemma 1, $\boldsymbol{\alpha}$ is the unique maximizer of $\ell_n^c(\alpha, \Sigma)$, implying that $\boldsymbol{\alpha}$ is also the unique solution to $\mathcal{C}(\widehat{\boldsymbol{\alpha}})$. Finally, Lemma 4 complements the set of primitive conditions and provides the sufficient conditions to apply the central limit theorem (CLT) for L^1-mixingales in White (2000).

As for the consistent estimation of the continuous-time IS measure, one may estimate the daily quadratic variation directly from the intraday returns at their highest frequency using a realized kernel (RK) approach, namely,

$$\widehat{IS}_{RK,m} = \frac{[\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{C}_{RK}]^2_m}{\widehat{\alpha}'_{\perp,\mathrm{IV}}\widehat{\Sigma}_{RK}\;\widehat{\alpha}'_{\perp,\mathrm{IV}}}, \qquad m = 1,\ldots,M, \tag{25}$$

where $\widehat{\alpha}_{\perp,\mathrm{IV}}$ is the orthogonal complement of $\widehat{\alpha}_{\mathrm{IV}}$ recovered from (24), $\widehat{\Sigma}_{RK}$ is the RK estimator of Σ, and $\widehat{C}_{RK}$ denotes a decomposition of $\widehat{\Sigma}_{RK}$ such that $\widehat{C}_{RK}\widehat{C}'_{RK} = \widehat{\Sigma}_{RK}$.

We compute the RK estimator of Σ as $RK \equiv \sum_{h=-\tau}^{\tau}\kappa(h/H)\Gamma_h$, where $H \propto \tau^{\zeta}$ with $1/2 < \zeta < 1$, Γ_h is the hth realized autocovariance estimator $\Gamma_h = \sum_{i=h+1}^{m}\Delta p_{\tau_i}\Delta p'_{\tau_{i-h}}$ for $h \geq 0$ and $\Gamma_h = \Gamma'_{-h}$ for $h < 0$, and κ is a nonstochastic kernel function. For the kernel function, we assume the same conditions as in Barndorff-Nielsen et al.'s (2011) Assumptions K.

Assumption K The nonstochastic kernel function κ is such that

(a) $\kappa(0) = 1$ and $\kappa'(0) = 0$;

(b) κ is twice differentiable with continuous derivatives;

(c) $\int_0^{\infty}[\kappa(x)]^2\,\mathrm{d}x < \infty$, $\int_0^{\infty}[\kappa'(x)]^2\,\mathrm{d}x < \infty$, and $\int_0^{\infty}[\kappa''(x)]^2\,\mathrm{d}x < \infty$;

(d) $\int_{-\infty}^{\infty}\kappa(x)\exp(ixs)\,\mathrm{d}x \geq 0$ for all $s \in \mathbb{R}$.

Condition (a) implies not only that Γ_0 gets unit weight but also that the kernel gives close-to-unit weight to Γ_h for values of h that are near the origin. Conditions (b) and (c) are purely technical, whereas (d) ensures that the resulting estimator is positive semi-definite by Bochner's theorem. The multivariate realized kernel is very similar to the standard HAC covariance matrix estimator, but with the additional assumption that $\kappa'(0) = 0$. The next result establishes the infill asymptotic theory for the estimates of CS and IS measures.

Theorem 2 Let the conditions in Lemma 2 hold and consider the corresponding vector of instruments. It then follows under Assumptions RT, GMM and K that, as $n \to \infty$, (*i*) $\widehat{\alpha}_{\perp,\mathrm{IV}} \xrightarrow{p} \alpha_{\perp}$;

(*ii*) $\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}_{\perp,\mathrm{IV}} - \boldsymbol{\alpha}_{\perp}\right) \xrightarrow{d} N\left(0, \mathcal{S}\left(\psi' \otimes \Xi\right)\left(\boldsymbol{G}'\Psi^{-1}\boldsymbol{G}\right)^{-1}\left(\psi' \otimes \Xi\right)' \mathcal{S}'\right)$,

where $\psi = \bar{\alpha}(\Xi - I_M)$, $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$, $\boldsymbol{G} = \mathbb{E}\left[\nabla_{\boldsymbol{\alpha}}\, \boldsymbol{g}_{n_{t_i}}(\boldsymbol{\alpha})\right]$ with $\boldsymbol{g}_{n_{t_i}}(\boldsymbol{\alpha}) = \left(\Delta p_{\tau_i} - \alpha_\delta(\boldsymbol{\alpha})\,\beta' p_{\tau_{i-1}}\right) \otimes Z_{\bar{k},\tau_{i-\bar{q}-k}}$, and $\Psi = \sum_{j=-\bar{q}-1}^{\bar{q}+1} \Psi(j)$ with $\Psi(j) = \mathbb{E}\left[\mathrm{v}_{t_i}\mathrm{v}'_{t_{i-j}} \otimes Z_{\bar{k},t_{i-\bar{q}-k}} Z'_{\bar{k},t_{i-\bar{q}-k-j}}\right]$;

(*iii*) $\widehat{IS}_{RK,m} \xrightarrow{p} IS$.

In particular, the empirical densities of the t-statistics based on Theorem 2(*ii*) are close to Gaussian for samples that mimic one trading day, i.e., prices recorded at random times and aggregated using the refresh-time scheme with approximately $n = 17,980$ observations. As before, this holds regardless of the variance of the market microstructure noise (see lower panel in Figure 1). In the next section, we show that the asymptotic results in Theorems 1 and 2 offer in general very good finite-sample approximations and compare favourably to the standard price discovery framework, i.e., computing CS and IS measures without taking into consideration the market microstructure noise.

4 Monte Carlo simulations

We next assess the relative performance of the IV-based approach under different scenarios. The characterization of the least-squares bias in Appendix A establishes that there are four sources of bias. The first is due to any dependence between efficient price and market microstructure noises. The second stems from the nonzero magnitude of the market microstructure noises, whereas the third comes from any correlation between microstructure noises across markets. Finally, any persistence in the microstructure noises adds another layer of bias in the LS estimation. In what follows, we aim to examine how the IV estimator accommodates each source of bias.

We focus on the situation in which one asset trades at two different platforms from 09:30 to 16:00. We assume without loss of generality that the first market leads the price discovery process. We simulate prices at the 1-second sampling interval from the exact discretization of (1) with $\alpha_\delta = (\alpha_{\delta,1}, \alpha_{\delta,2})'$, $\beta = (1, -1)'$, $\Sigma = \mathrm{diag}\,(1/23400, 1/23400)$, and correlation among markets equal to 0.5. The elements of α_δ drive the price discovery for they determine how fast each market adjusts.

We report results for two alternative specifications: $\alpha_\delta = (0.000, 0.050)'$ and $\alpha_\delta = (-0.025, 0.050)'$, corresponding to $\alpha_\perp = (1, 0)'$ and $\alpha_\perp = (2/3, 1/3)'$, respectively.[10]

The market microstructure noises follow possibly correlated MA(q) with lag orders $q \in \{0, 1, \infty\}$:

$$u_{t_i} = e_{t_i} + \sum_{\ell=1}^{q} \Lambda_{e,\ell}\, e_{t_{i-\ell}},$$

where $\Lambda_{e,\ell} = \operatorname{diag}(\lambda_{e,1,\ell}, \lambda_{e,2,\ell})$, e_{t_i} is normally distributed with zero mean and covariance matrix Ω_e, and $\operatorname{corr}(u_{t_i}, \varepsilon_{t_i})$ is set to either zero (exogenous noise) or -0.2 (endogenous noise). A negative correlation between market innovations and the market microstructure noise is consistent with Hansen and Lunde's (2006) empirical findings.

The type of persistence we allow in the microstructure noise depends on the assumptions in Section 3 and hence on the type of instruments we consider. Under Assumption MMN(TS), we contemplate either iid (exogenous) noises or serially dependent market microstructure noise coming from endogenous MA(1) processes with $\rho_{m,m'}(0) = -0.2$ and $\gamma_{m,m'}(q) \neq 0$ for $q = 0, 1$. Under Assumption MMN(CS), we consider a similar noise to the latter, but with $q = \infty$. In practice, we model the market microstructure noises as AR(1) processes with autoregressive parameter set to 0.50, so that $\lambda_{e,1,\ell} = \lambda_{e,2,\ell} = -0.5^\ell$ for $\ell = 0, 1, \ldots$.

The variance of the microstructure noise plays a key role in our Monte Carlo experiments. Empirical evidence suggests that the size of the noise is relatively small, about $\mathbb{V}(u_{m,t_i}) = \omega_m^2 < (1/1000)\Sigma_m$, with Σ_m denoting the mth market integrated variance (Hansen and Lunde, 2006). It has also been decreasing over the years (Ait-Sahalia and Xiu, forthcoming). We entertain different magnitudes for the microstructure noises: $100\,\omega^2 \in \{(0.01, 0.01)', (0.05, 0.05)', (0.10, 0.10)', (0.10, 0.05)'\}$. We account for non-synchronous quoting/trading activity by randomly selecting about 18,000 observations for each stock using independent Bernoulli trials. We the synchronize log-prices using the refresh-time procedure. Finally, we aggregate the resulting data at the 1- and 5-minute sampling intervals.

4.1 Finite-sample properties under Assumption MMN(TS)

This section assesses the finite sample performance of the LS- and IV-based estimates of the CS and of the continuous- and discrete-time IS measures using tick data. We report results for exoge-

[10] The online appendix also entertains different correlation among markets (namely, 0 and 0.9) and alternative speed-of-adjustment parameters, $(0, 0.01)'$ and $(-0.005, 0.010)'$.

nous iid and endogenous MA(1) market microstructure noises using instruments consistent with Assumption MMN(TS) in Section 3. We contemplate LS-based price discovery measures based on both

$$\text{VEC}(0) \qquad \Delta p_{t_i} = \alpha_\delta (p_{1,t_{i-1}} - p_{2,t_{i-1}}) + \mathrm{v}_{t_i}$$

$$\text{VEC}(1) \qquad \Delta p_{t_i} = \alpha_\delta (p_{1,t_{i-1}} - p_{2,t_{i-1}}) + \Gamma_\delta \Delta p_{t_{i-1}} + \mathrm{v}_{t_i},$$

whereas the IV-based estimates rest exclusively on the VEC(0) specification, bearing in mind that $\mathrm{v}_{t_i} = \Delta u_{t_i} + \varepsilon_{t_i} - \alpha_\delta (u_{1,t_{i-1}} - u_{2,t_{i-1}})$. Lastly, we also report the results for the LS estimates of the CS measures at the 1- and 5-minute sampling intervals.

As for the IS measures, we start with a standard VEC framework, employing least-squares estimates of the VEC parameters and the realized variance estimator of the covariance matrix for every frequency (i.e., tick-by-tick, 1-minute, and 5-minute). In contrast, the IV-based measures rest obviously on the IV estimates of $\alpha_\perp$, but consider two different estimators of the covariance matrix. In particular, we either estimate $\Sigma = CC'$ using a realized kernel approach or estimate its exact discretization $\Sigma_\delta = C_\delta C_\delta'$ using Andrews's (1991) HAC estimator with a Parzen kernel.

Table 1 documents the performance of each estimator for the CS measure of the first market $\alpha_{\perp,1}$ in the case of iid microstructure noises. By construction, it suffices to focus exclusively on one market given that the CS measures must sum up to one across markets. The first two panels respectively display the bias analyses for $\alpha \in \{(0, 0.05)', (-0.025, 0.050)'\}$, whereas the third and fourth panels reveal their relative root median squared error (RRMSE) with respect to the LS-based estimates from VEC(0) at the tick-by-tick frequency. We report different estimates of the CS measures under iid microstructure noises with variances ω_m^2 ranging from 0.0001 to 0.001 ($m = 1, 2$). The columns VEC($\jmath$), VEC$_1$($\jmath$), VEC$_5$($\jmath$) refer to the LS estimates of $\alpha_\perp$ from the VEC($\jmath$) specification ($\jmath = 0, 1$) at the tick-by-tick, 1-minute and 5-minute sampling intervals, respectively.

When estimating $\alpha_\perp$ by 2SLS using $(p_{1,t_{i-1-\bar{q}-\kappa}} - p_{2,t_{i-1-\bar{q}-\kappa}})'$ as instruments, where κ and $\bar{\kappa}$ are integers such that $2 \le \kappa \le \bar{\kappa} \le \infty$, we select two set of valid and relevant instruments under Assumption MMN(TS). In particular, the columns 2SLS and 2SLS$_1$ concern the 2SLS estimates of $\alpha_\perp$ from the VEC(0) specification using tick data, with $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ as instruments for $2 \le \kappa \le 6$ and $3 \le \kappa \le 7$, respectively. Finally, we also report the F-statistic from a joint significance test on the parameter estimates from the first-stage equation of the 2SLS estimator as a measure of the relevance of the instruments.

Altogether, the 2SLS estimator provides the best performance in terms of both bias and RRMSE. This holds irrespectively of the microstructure noise level and speed-of-adjustment parameters. Notably, as our theoretical results suggest, using VEC(1) at the tick-by-tick frequency does not help reducing the bias and root median square error (RMSE), whereas fitting both a VEC(0) and VEC(1) models at 1-minute sampling interval alleviates the bias and RMSE for low levels of microstructure noises. The $2SLS_1$-based CS estimates perform worst than the 2SLS-based measures because the latter employs more relevant instruments.

We next turn our attention to endogenous MA(∞) microstructure noise. Table 2 reports the results for the CS measure in the case of endogenous MA(1) microstructure noises. This is a much more complex and interesting specification, with market microstructure noises featuring both serial dependence and endogeneity. These features affect not only the estimates of Σ, but also the asymptotic bias of α_δ. Surprisingly, most of the Monte Carlo experiments in the realized measure literature does not account for both serial dependence and endogeneity. As before, the IV approach we propose easily outclasses the extant price discovery measures. The gains in both bias and RRMSE are considerable, highlighting the importance of accounting for the microstructure noise.

Tables 3 and 4 display the results for the continuous-time IS measures in the cases of iid and endogenous MA(1) microstructure noises, respectively. In particular, we report bias and RRMSE for ten estimators. As before, the columns VEC($\jmath$), $\text{VEC}_1(\jmath)$, $\text{VEC}_5(\jmath)$ refer to the LS-based estimates of the continuous-time IS measures from the VEC($\jmath$) specification ($\jmath = 0, 1$) at the tick-by-tick, 1-minute and 5-minute sampling intervals, respectively. We estimate the integrated covariance matrix using the realized covariance estimator for the LS-based measures, whereas we employ either RK or HAC estimators in the case of IV-based measures. Note however that these estimators are very different in nature. While the RK approach estimates directly the continuous-time covariance matrix, the HAC estimator has to use the exact discretization to back out the continuous-time unconditional covariance matrix.

As in the case of the CS measures, the bias of the LS-based estimates of the continuous-time IS is large for all levels of microstructure noises. The $2SLS_{\text{RK}}$ and $2SLS_{\text{RK},1}$ estimators have in general smaller biases and RMSE, but still quite off target regardless of the magnitude of the microstructure noise. These biases are due to the upward bias on the off-diagonal elements of the RK estimates of the integrated covariance matrix. While the finite-sample biases of the RK estimates of the diagonal elements of Σ_0 are close to zero, those of the off-diagonal elements are usually large and positive. This happens mainly because the speed-of-adjustment parameters we consider in our data generation process are relatively fast, implying that at the tick-by-tick frequency (about one observation every 1.6 seconds), the signal required to consistently estimate the correlation between markets is already too weak. In turn, the 2SLS_{HAC} and $2\text{SLS}_{\text{HAC},1}$ estimators perform best in terms of both finite-sample bias and RMSE, largely outperforming the other estimators.

Table 5 documents the results for the discrete-time IS measures. Because the discrete-time IS measures converge to $1/M$ as the sampling frequency decreases, we restrict our analyses to the tick-by-tick frequency using the same estimators as before. Once more, the IV-based measures outperform the competitors irrespectively of the speed of adjustment in each market and market microstructure noise we use.

LS inconsistency affects not only the usual price discovery measures, but also the impulse response functions. To formally address the finite-sample performance of both LS- and IV-based IRFs, Figures 3 and 4 plot the 5%, 50% and 95% empirical quantiles of the corresponding IRFs based on 1,000 replications. The rows contain the different estimators, while the columns relate to the magnitudes of the microstructure noise. The solid lines correspond to the median IRFs, with shades covering the region between the 5% and 95% quantiles, and the dotted lines are the true IRFs.

In line with our previous findings, the IV-based impulse response functions behave better than their LS counterparts. Regardless of whether we use a VEC(0) or VEC(1) specification, there is a large positive bias in the LS-based impulse responses in the presence of market microstructure noise, overestimating the speed at which markets adjust to a change in the efficient price. A different picture emerges if we examine IRFs based on 2SLS_{HAC} and $2\text{SLS}_{\text{R,HAC}}$ estimates. They exhibit little bias, correctly tracking the true IRF in all leads. As expected, the 90% confidence intervals increase with the magnitude of the market microstructure noise, implying that fragmented markets might appear more efficient than they actually are if one does not account for the market microstructure noise.

All in all, our Monte Carlo simulations show that the IV-based price discovery measures are more reliable than their LS-based counterparts in every scenario we investigate, regardless of the speed-of-adjustment in each market, leadership pattern, and sort of microstructure noise (large vs small in magnitude, idd vs persistent, exogenous vs endogenous). We further document that the CS estimates have better finite-sample properties than both the continuous- and discrete-time IS measures under Assumption MMN(TS).

4.2 Finite-sample performance under Assumption MMN(CS)

The second set of Monte Carlo experiments examines the finite-sample performance of the LS- and IV-based price discovery measures in the presence of endogenous MA(∞) microstructure noises. Under these circumstances, we have to select instruments in line with Assumption MMN(CS): namely, current and lagged values of $\widetilde{\beta}' s_{\tau-1}$. Because we have the entire cross section of assets to chose from, it is natural to design a simulation study that allows for potentially many instruments. We consider IV-based estimators that use $V = 1, 5, 10, 20$ assets as instruments. The main challenge of Assumption MMN(CS) is that instruments are rather weak, and so we also report the results for the price discovery measures using the CU-GMM estimator for 10 and 20 instruments.

We simulate the data as a system of 42 log-prices at the 1-second sampling interval from the exact discretization of (1). They correspond to 21 different assets, each of them trading at two exchanges. The resulting VEC(0) model has 21 cointegrating vectors. We define the 42×21 matrix α in such way that there is no Granger causality among the different assets. The remaining steps are the same as the ones in the previous section: (1) we model the market microstructure noise as an endogenous MA(∞) process so that they satisfy Assumption MMN(CS); (2) we account for nonsynchronous quoting/trading activity by randomly selecting about 18,000 observations for each stock using independent Bernoulli trials; (3) we aggregate the prices using the refresh time technology; and then (4) aggregate data to the 1- and 5-minute sampling intervals.

To save space, we do not report results from the LS-based estimates from the VEC(1) models, as they are essentially the same as in Section 4.1. As such, we restrict attention to the LS-based measures from the VEC(0) models at the tick-by-tick, 1-minute, and 5-minute sampling intervals, whereas the IV- and CU-GMM estimates are only for the VEC(0) models at the tick-by-tick frequency. We organize the tables in the same fashion as in Section 4.1, reporting not only median bias but also RRMSE using the LS estimator at the tick-by-tick frequency as benchmark.

Table 6 shows the performance of each estimator of the first market's component share. As expected, the LS-based estimators perform very poorly irrespectively of the sampling frequency and the magnitudes of the microstructure noise and of the speed-of-adjustment parameters. For the lowest level of noise, $\omega^2 = (0.0001, 0.0001)'$, the 2SLS estimator with a single instrument performs very well, with virtually no finite-sample bias and with a low RRMSE. For larger magnitudes of microstructure noise, both the first-stage F-statistics and R^2 drop sharply, and hence the performance of the 2SLS estimator deteriorates. Adding more instruments is not of great help, as finite-sample bias remains moderate to high. These results are in line with the asymptotic theory in that the 2SLS estimator is inconsistent in the case of many weak instruments.

The picture dramatically changes once we focus on the CU-GMM estimators. The corresponding CS measures are virtually unbiased in all scenarios we entertain, apart from displaying the lowest RRMSEs. Taking into consideration both median bias and RRMSE criteria, the CU-GMM estimator with 20 instruments exhibits the overall best performance, supporting our rationale that using many weak instruments pays off when estimating the CS measures.

Table 7 presents the results for the continuous- and discrete-time IS measures. Overall, we find that all estimators exhibit some finite-sample bias for any magnitude of the microstructure noise apart from $\omega^2 = (0.0001, 0.0001)'$. Combining the HAC covariance matrix estimator with VEC parameter estimates coming either from the 2SLS with 5 instruments or from the CU-GMM with 20 instruments entails the best improvement relative to the standard LS approach, even if it still exhibits some finite-sample bias.

As for the impulse response functions, Figure 5 depicts their median values, as well as their 5% and the 95% empirical quantiles, based on 1,000 replications. In each plot, the rows display the IV-based estimates with 1 to 20 instruments and the CU-GMM estimates with 10 and 20 instruments, whereas the columns refer to the magnitude of the microstructure noise. As before, solid and dotted lines correspond respectively to the median and true impulse response functions.[11]

[11] To save space, we refrain to report the impulse response functions for the LS estimates, as they do not differ in any qualitative or quantitative aspect from the ones in Figure 4. Results are obviously available from the authors upon request.

In line with the previous discussion, all IV-based estimators deliver unbiased estimates of the IRFs in the case of $\omega^2 = (0.0001, 0.0001)'$. For larger microstructure noises, we find that the CU-GMM estimator with 20 instruments performs best, especially at what concerns the coverage of the 90% empirical confidence interval.

In summary, the IV-based estimators of α and $\alpha_\perp$ perform very well in both settings. As for the IS measures, our simulations show that it is preferable to estimate the VEC parameters by IV and then compute the HAC estimator of the residual covariance matrix using the exact discretization results. Impulse response functions based on LS residuals are generally downward biased, making the convergence to the long-run effect falsely too fast. Finally, the additional Monte Carlo experiments in the online appendix confirm our findings under Assumptions MMN(TS) and MMN(CS) for slower speeds of adjustment and different types of market microstructure noises (magnitude and dependence structure).

5 Price informativeness: NYSE versus Nasdaq

In this section, we compute price discovery measures for Alcoa (AA). Although Alcoa trades on multiple venues in the US, we focus exclusively on the NYSE and Nasdaq in view that they are the most active trading platforms. We extract quotes data from TAQ for the period ranging from June 2012 to May 2013. 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 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 (midquotes) at each second in the event that there are multiple ticks taking place at the same second.

We synchronize the midquotes from both trading venues by using the refresh-time approach and then sampling at regularly spaced intervals of 1 and 5 minutes. Apart from the entire sample, we also consider one subsample for each of the 12 months in the sample period. In the context of Assumption MMN(CS), we take as potential instruments a panel of 25 stocks that trade at ARCA, BATS and NASDAQ OMX BX Stock Exchange (formerly Boston Stock Exchange). Their ticker symbols are AAPL, BAC, BRKB, CSCO, DAL, GM, GOOG, HPQ, IBM, JCP, JNJ, JPM, KO, MO, MRK, MRVL, MSFT, NOK, ORCL, PFE, PG, VZ, WFC, XOM, and YHOO. We handle their data using exactly the same filters we use for AA. Finally, as a measure of liquidity and trading activity, Table 9 also reports the average duration between quotes for the entire sample period for all stocks in the different trading venues.

To test for cointegration in each month of the sample, we employ the procedure in Hansen and Lunde (2006). We first choose the lag length of the daily VEC models by obtaining the most parsimonious specification in which the LM test cannot reject the absence of residual autocorrelation at the 5% significance level,[12] and then apply Johansen's maximum eigenvalue and trace cointegration tests. We find strong evidence, at the 1% level of significance, that there is only one cointegrating vector in every day of the sample, regardless of the sampling interval. In the standard price discovery analysis, one must typically equip the VEC specification with a rich lag structure, ignoring the fact that the presence of market microstructure noise may hatch spurious autocorrelation in the least-squares residuals. In contrast, our IV-based price discovery analysis always consider a simple VEC(0) model, without adding any past price changes to the specification, to conform with the continuous-time model we have in mind.

The first set of results considers Assumption MMN(TS) with $0 \leq q < \infty$ and the past price differences between the NYSE and Nasdaq as instruments. Table 10 reports both monthly and yearly estimates of the NYSE's component share, which we denote by $\alpha_{\perp,N}$, at the tick-by-tick, 1-minute and 5-minute sampling intervals.[13] As in the previous section, we compare the 2SLS estimates with the LS estimates of the VEC(0) model at the different frequencies.

Interestingly, inference about market leadership depends on whether the estimator is robust to microstructure noise. This holds for every month in the sample period. The LS estimates at the tick-by-tick frequency assign Nasdaq as the market leader, whereas both 2SLS estimates attribute market leadership to the NYSE. The LS results are not robust however to the sampling frequency. Using 1- and 5-minute intervals, the LS-based analyses indicate that the NYSE is more important than Nasdaq for the price discovery. This is consistent with our theoretical and simulation results

[12] Minimizing the Bayesian information criterion yields very similar results.

[13] Recall that, by construction, the Nasdaq's component share is $1 - \alpha_{\perp,N}$, as we normalize the elements of the orthogonal complement to sum one.

in that the finite-sample bias of the LS estimator increases with the sampling frequency.

In line with our simulation study, the instruments under Assumption MMN(TS) are generally strong enough, with R^2 over 20%, for virtually every month. We also reject the null that $\beta' P_{t_{i-1}}$ is exogenous in all subsamples given the significant differences between the LS and IV estimates of the speed-of-adjustment parameters. These differences are larger for the NYSE, perhaps hinting that the microstructure noise is larger in magnitude at the NYSE.[14] This is exactly the situation in which LS-based market leadership analyses fail.[15] To confirm whether this is the case, we estimate the the realized variance in each market and use the fact that $RV/(2n) \xrightarrow{p} \omega^2$ under the assumption of iid microstructure noises. We indeed find that $\widehat{\omega}^2_N > \widehat{\omega}^2_T$, corroborating the indirect evidence in the difference between LS and IV estimates. The main take way is that, in the presence of microstructure noise, one should always consider IV-based price discovery measures to make inference about market leadership.

Table 11 reports the information shares we estimate by LS and IV. Once again, the continuous-time IS measures are markedly different according to the estimator we use. In particular, we find that their $2SLS_{\text{HAC}}$ and $2SLS_{\text{HAC},1}$ estimates are seemingly the best to identify the leading market in the price discovery process. They indicate that the continuous-time IS is about 0.60 across every subsample. This highlights that the speeds at which both NYSE and Nasdaq currently operate call for ultra high resolutions; otherwise, the contemporaneous correlation between markets gets very close to unit, implying uninformative information shares (Hasbrouck, 2018).

In the second set of empirical results, we assume that the microstructure noises are MA(∞) processes. Just as in the Monte Carlo study, we use other assets traded at different exchanges as instruments. Using the refresh-time scheme, we synchronize the AA stock midquotes with the entire set of instruments at the tick-by-tick frequency. Because refresh time depends heavily on the least liquid asset, we entertain two sets of instruments with different sampling frequencies. The first selects assets with at least 7,500 quotes on every trading day (about one quote every 3 seconds) within the month. This means that the tickers in this set of instruments may vary across months. The second set of instruments raises the bar to at least 5,000 quotes per day (about one quote

[14] The full set of results, including the speed-of-adjustment estimates for the NYSE and Nasdaq markets, are available upon request

[15] See, for instance, our simulation results for $\omega^2 = (0.0010, 0.0005)'$ in the previous section.

every 5 seconds).

Table 12 documents the monthly CU-GMM estimates of $\alpha_{\perp,N}$ for both set of instruments. The instruments in the top panel correspond to stocks that post at least $7,500$ quotes per day. It is reassuring that the CU-GMM component share estimates also indicate that the NYSE leads the price discovery mechanism in every month of the sample period. Moreover, we reject the null that $\alpha_{\perp,N} = 0$ for all months but two at the 10% significance level. As we cannot individually reject the null that $\alpha_{\perp,T} = 0$ for every month, it seems that the NYSE is the sole contributor of the price discovery process. As expected, the instruments are rather weak, with R^2 and F-statistics well below 0 and 10, respectively. The average duration is about 8 seconds, which is fourfold that under Assumption MMN(TS). This poses an extra challenge to the estimation of the CS measures, as the strength of the signal we use to identify the price discovery measures decreases with the duration.

The bottom panel of Table 12 displays the CS estimates using as instruments the stocks with at least $5,000$ quotes. By decreasing the liquidity requirement, the number of instruments increases from about 14 to 28 assets. Because we now consider less liquid assets, the average duration also increases to about 17 seconds, which contribute to make instruments even weaker. We nonetheless find very similar results, with the NYSE still leading the price discovery process. Altogether, the results under Assumption MMN(CS) corroborate our previous findings under Assumption MMN(TS) and, most importantly, highlights the importance of using methods that are robust to the presence of market microstructure noise.

Out last set of results rests on the infill asymptotic theory in Section 3.2. We compute daily measures of the CS measures and investigate how the price discovery evolves over time. We treat each trading day independently by running individual VEC models for each day in our sample. For the VEC(1) model, we estimate by LS the CS measures from tick data, whereas we estimate $\alpha_\perp$ in the VEC(0) models by LS and IV. We consider two alternative sets of instruments that meet the requirements in Assumption MMN(TS). The upper panel of Figure 6 displays the CS measures based on the LS estimator in both VEC(0) and VEC(1) specifications, whilst the lower panel exhibits the IV estimates of $\alpha_\perp$.

In short, the daily measures of price discovery confirm our previous results under the standard asymptotic theory. Estimating component shares by LS entail divergent results in terms of market leadership relative to the microstructure-robust IV-based measures. In line with our theoretical and simulation results, the LS-based component share estimates seem biased towards 1/2. As before, the daily IV estimates identify the NYSE as the leading market over the entire sample period. As for the IRFs, Figure 7 displays the median, 5%, and 95% percentiles of the daily IRF of Nasdaq to shocks in the efficient price. Given our simulations, we expect that LS-based impulse response functions to converge way faster towards one in the presence of market microstructure noise. This is indeed the case, with the IRFs of the VEC(0) models converging almost instantaneously to one for the LS estimates, whereas the robust-to-noise IV-based IRFs take, on average, five to ten periods to incorporate the information from the efficient price into market prices.

All in all, this section shows the importance of using estimators that are robust to market microstructure noise, as the non-robust methods are heavily inconsistent to the point that conclusion regarding market leadership in the price discovery process flips.

6 Conclusion

This paper examines the impact of market microstructure noise in the estimation of price discovery measures. In particular, we show that the presence of microstructure noise biases the standard least-squares VEC estimation that permeates the traditional price discovery analysis. We propose an IV-based approach and derive valid sets of instruments under two different assumptions about the market microstructure noises. We establish the large sample properties of the IV-based estimators under two asymptotic settings: the usual asymptotic theory that fixes the number of intraday observations, while letting the number of days diverge; the infill asymptotic theory that fixes the number of trading days to one whilst letting the number of intraday observations to grow to infinity. Monte Carlo simulations confirm that the standard LS estimates are inconsistent, whereas our IV approach performs very well, with virtually no bias in finite samples. Interestingly, in our empirical application, we find that market leadership between the NYSE and Nasdaq actually flips once we account for microstructure noise. This result holds for the two different assumptions about market microstructure noise and the alternative asymptotic settings.

Appendix A Proofs

Proof of Theorem 1 We start by proving (i). First note that because $\beta = (I_r, -\iota_r)'$, then $\beta_\perp = \iota_M$, meaning that Ξ has common rows. Furthermore, as (11) has no autoregressive terms and $\sum_{m=1}^{M} \alpha_{m,\perp} = 1$, $\alpha_\perp$ is equal to any row of Ξ. Let $\widehat{\Xi} = \beta_\perp \left[\widehat{\alpha}'_\perp \beta_\perp\right]^{-1} \widehat{\alpha}'_\perp$ and rewrite

$$
\begin{aligned}
\widehat{\Xi} &= \beta_\perp \left[\alpha'_\perp \beta_\perp + \left(\widehat{\alpha}_\perp - \alpha_\perp\right)' \beta_\perp\right]^{-1} \left(\alpha'_\perp + \left(\widehat{\alpha}'_\perp - \alpha'_\perp\right)\right), \\
&= \Xi - \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \left(\widehat{\alpha}_\perp - \alpha_\perp\right)' \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp + \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \left(\widehat{\alpha}_\perp - \alpha_\perp\right)'. \qquad (26)
\end{aligned}
$$

Without loss of generality, for any fixed $\delta \in [0,1]$, choose $\widehat{\alpha}_{\perp,\text{IV}} = \alpha_\perp - \widehat{\bar{\alpha}}_\delta \widehat{\alpha}'_{\delta,IV} \alpha_\perp$ with $\widehat{\bar{\alpha}}_\delta = -\alpha_\delta \left(\widehat{\alpha}'_{\delta,IV} \alpha_\delta\right)^{-1}$ such that $\widehat{\alpha}'_{\delta,IV} \widehat{\alpha}_{\perp,\text{IV}} = 0$. Substituting $\widehat{\alpha}_{\perp,IV} - \alpha_\perp$ into (26) and rearranging terms,

$$
\begin{aligned}
\widehat{\Xi} - \Xi &= \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \left[\alpha'_\perp \left(\widehat{\alpha}_{\delta,IV} - \alpha_\delta\right) \widehat{\bar{\alpha}}'_\delta\right] \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp - \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp \left(\widehat{\alpha}_{\delta,IV} - \alpha_\delta\right) \widehat{\bar{\alpha}}'_\delta, \\
&= \Xi \left(\widehat{\alpha}_{\delta,IV} - \alpha_\delta\right) \widehat{\bar{\alpha}}'_\delta \Xi - \Xi \left(\widehat{\alpha}_{\delta,IV} - \alpha_\delta\right) \widehat{\bar{\alpha}}'_\delta, \\
&= \Xi \left(\widehat{\alpha}_{\delta,IV} - \alpha_\delta\right) \widehat{\bar{\alpha}}'_\delta \left(\Xi - I_M\right). \qquad (27)
\end{aligned}
$$

Using the properties of the vec operator, it follows that

$$
\text{vec}\,(\widehat{\Xi}) - \text{vec}\,(\Xi) = \left(\widehat{\psi}' \otimes \Xi\right) \Big(\text{vec}\,(\widehat{\alpha}_{\delta,IV}) - \text{vec}\,(\alpha_\delta)\Big), \qquad (28)
$$

where $\widehat{\psi} = \widehat{\bar{\alpha}}_\delta \left(\Xi - I_M\right)$. Let $\boldsymbol{K}_{M,R}$ be the commutation matrix with dimensions $MR \times MR$ such that $\boldsymbol{K}_{M,R}\text{vec}\,(\alpha_\delta) = \text{vec}\,(\alpha'_\delta)$. Specifically, $\boldsymbol{K}_{M,R} = \sum_{h=1}^{M} \sum_{h'=1}^{R} K_{h,h'} \otimes K'_{h,h'}$, where $K_{h,h'}$ is a $M \times R$ matrix with $\left[K_{h,h'}\right] = 1$ and $\left[K_{\hbar,\hbar'}\right] = 0$ if $\hbar \neq h$ or $\hbar' \neq h'$. It then follows that (28) reads

$$
\text{vec}\,(\widehat{\Xi}) - \text{vec}\,(\Xi) = \left(\widehat{\psi}' \otimes \Xi\right) \boldsymbol{K}_{M,R} \Big(\boldsymbol{\alpha}_{\delta,IV} - \boldsymbol{\alpha}_\delta\Big). \qquad (29)
$$

Because Ξ has common rows, define a deterministic matrix $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_M)$ such that $\widehat{\alpha}_\perp = \mathcal{S}\text{vec}\,(\widehat{\Xi})$ and $\alpha_\perp = \mathcal{S}\text{vec}\,(\Xi)$, where $\mathcal{S}_j = (\iota(j), 0_{M\times R})$ with $j = 1, \ldots, M$, $\iota(j)$ denoting an $M \times 1$ vector that has its jth row equal to one and all the remaining entries equal to zero, and $0_{M\times R}$ is an $M \times R$ matrix of zeros. It the follows that (29) reads

$$
\widehat{\alpha}_\perp - \alpha_\perp = \mathcal{S} \left(\widehat{\psi}' \otimes \Xi\right) \boldsymbol{K}_{M,R} \Big(\boldsymbol{\alpha}_{\delta,IV} - \boldsymbol{\alpha}_\delta\Big). \qquad (30)
$$

To show $(\boldsymbol{\alpha}_{\delta,IV} - \boldsymbol{\alpha}_\delta) \xrightarrow{p} 0$, note that

$$
\begin{aligned}
\boldsymbol{\alpha}_{\delta,IV} - \boldsymbol{\alpha}_\delta = &\left[\left(\frac{1}{T} \bar{\boldsymbol{p}}' \left(I_M \otimes \boldsymbol{Z}\right)\right) \Big(T\boldsymbol{W}\Big) \left(\frac{1}{T} \left(I_M \otimes \boldsymbol{Z}\right)' \bar{\boldsymbol{p}}\right)\right]^{-1} \times \\
&\left(\frac{1}{T} \bar{\boldsymbol{p}}' \left(I_M \otimes \boldsymbol{Z}\right)\right) \Big(T\boldsymbol{W}\Big) \left(\frac{1}{T} \left(I_M \otimes \boldsymbol{Z}\right)' \mathbf{v}\right). \qquad (31)
\end{aligned}
$$

Assumption RROU and relevance of the instruments (item ($ii.$) in Lemma 2 or 3) imply the first term on the right-hand-side of (31) is $O_p(1)$, meaning that $\underset{T\to\infty}{\text{plim}}\,\frac{1}{T}\bar{\boldsymbol{p}}'(I_M\otimes\boldsymbol{Z}) = \boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}} < \infty$ and $\underset{T\to\infty}{\text{plim}}\,T\boldsymbol{W} = \boldsymbol{Q}_{\boldsymbol{W}} < \infty$ are $O_p(1)$ terms. Standard weak law of large number for stationary processes (see, for instance, Proposition 7.5 in Hamilton, 1994) and instruments validity (item ($i.$) in Lemma 2 or 3) render $\underset{T\to\infty}{\text{plim}}\left(\frac{1}{T}(I_M\otimes\boldsymbol{Z})'\,\mathbf{v}\right) = 0$, implying that $\boldsymbol{\alpha}_{\delta,IV}\xrightarrow{p}\boldsymbol{\alpha}_\delta$. Finally, consistency of $\boldsymbol{\alpha}_{\delta,IV}$ implies $\widehat{\alpha}_\delta\xrightarrow{p}\bar{\alpha}_\delta = -\alpha_\delta(\alpha_\delta'\alpha_\delta)^{-1}$ and $\widehat{\psi}\xrightarrow{p}\psi$. It then follows that $\widehat{\alpha}_\perp\xrightarrow{p}\alpha_\perp$, which proves (i).

To prove (ii), use the results in (i) and rewrite (30) such that as $T\to\infty$,

$$\sqrt{T}\Big(\widehat{\alpha}_{\perp,\text{IV}}-\alpha_\perp\Big) = \mathcal{S}\,(\psi'\otimes\Xi)\,\boldsymbol{K}_{M,R}\left[\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}'\right]^{-1}\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}\Big(\frac{1}{\sqrt{T}}(I_M\otimes\boldsymbol{Z})'\,\mathbf{v}\Big). \tag{32}$$

Note that $\frac{1}{\sqrt{T}}(I_M\otimes\boldsymbol{Z})'\,\mathbf{v} = \text{vec}\left(\frac{1}{\sqrt{T}}\sum_{\tau=1}^{T}Z_{\tau-k^*}\text{v}_{1,\tau},\ldots,\frac{1}{\sqrt{T}}\sum_{\tau=1}^{T}Z_{\tau-k^*}\text{v}_{M,\tau}\right)$ is not a vector of martingale difference sequences, but a serially correlated covariance stationary processes. Lemma 4 shows that $\left\{\beta_{\underline{m}}'p_{t_i-\bar{q}-\underline{k}}\text{v}_{m,t_i},\mathcal{F}_{t_i}\right\}$ and $\left\{\widetilde{\beta}_{\underline{j}}'s_{t_{i-1}}\text{v}_{m,t_i},\mathcal{F}_{t_i}\right\}$ are ergodic uniform integrable L^1- mixingale sequences under Assumptions MMN(TS) and MMN(CS), respectively, which satisfies the conditions of the CLT for L^1-mixingale sequences (see, for instance, White, 2000; Davidson, 1994). By applying the Cramér-Wold device, it then follows that

$$\frac{1}{\sqrt{T}}(I_M\otimes\boldsymbol{Z})'\,\mathbf{v}\xrightarrow{d}N\left(0,\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}\right), \tag{33}$$

where $\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}} = \sum_{j=-\bar{q}-1}^{\bar{q}+1}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j)$ and $\sum_{j=-\infty}^{\infty}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j)$ under Assumptions MMN(TS) and MMN(CS), respectively, and $\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}(j) = \mathbb{E}\left[\text{v}_\tau\text{v}_{\tau-j}'\otimes Z_{\tau-k^*}Z_{\tau-k^*-j}'\right]$ with k^* denoting the instruments lag order and $\text{v}_\tau = (\text{v}_{1,\tau},\ldots,\text{v}_{M,\tau})'$. The limiting distribution of (32) then reads

$$\sqrt{T}\Big(\widehat{\alpha}_{\perp,\text{IV}}-\alpha_\perp\Big)\;\xrightarrow{d}\;N\left(0,\boldsymbol{C_1C_2C_3}\Gamma_{\bar{\boldsymbol{p}}\boldsymbol{Z}}\boldsymbol{C_3'C_2'C_1'}\right), \tag{34}$$

where $\boldsymbol{C_1} = \mathcal{S}(\psi'\otimes\Xi)\boldsymbol{K}_{M,R}$, $\boldsymbol{C_2} = \left[\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}\boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}'\right]^{-1}$ and $\boldsymbol{C_3} = \boldsymbol{Q}_{\bar{\boldsymbol{p}},\boldsymbol{Z}}\boldsymbol{Q}_{\boldsymbol{W}}$, which proves item (ii).

In view that $\widehat{\text{v}}_{t_i}$ is consistent, $\widehat{\Sigma}_{\delta,IV}\xrightarrow{p}\Sigma_\delta$ follows from standard HAC asymptotic theory (see, for instance, White's (2000) Corollary 6.11). Combining this result with item (i) in this theorem and applying the Slutsky's theorem, it follows that $\widehat{IS}_{\delta,IV}\xrightarrow{p}IS_\delta$, which proves item (iii). Finally, using $\widehat{\Sigma}_{\delta,IV}\xrightarrow{p}\Sigma_\delta$ and the Slutsky's theorem, it follows that $\widehat{IS}_{IV}\xrightarrow{p}IS$, which proves ($iv.$). ■

Proof of Theorem 2 To prove (i), first recall that Lemma 1 states that instruments are valid and relevant, implying that the moment conditions are valid and identified. Next, by replicating the steps in Theorem 1(i), write

$$\widehat{\alpha}_{\perp,\text{IV}} - \alpha_{\perp} = \mathcal{S}\left(\widehat{\psi}' \otimes \Xi\right) \boldsymbol{K}_{M,R}\Big(\widehat{\boldsymbol{\alpha}}_{\text{IV}} - \boldsymbol{\alpha}\Big), \tag{35}$$

where $\widehat{\psi} = \widehat{\bar{\alpha}}\,(\Xi - I_M)$. Recall that $\widehat{\boldsymbol{\alpha}}_{\text{IV}} = \text{argmax}_{\widehat{\boldsymbol{\alpha}}}\ \mathcal{C}_{\text{GMM},n_\tau}(\widehat{\boldsymbol{\alpha}})$. It then follows that from Assumption GMM (c),

$$\widehat{\boldsymbol{\alpha}}_{\text{IV}} = \underset{\widehat{\boldsymbol{\alpha}}}{\text{argmax}}\ \mathcal{C}_{\text{GMM},n_t}(\widehat{\boldsymbol{\alpha}}) + o_p(1). \tag{36}$$

We now use the same arguments as in Newey and McFadden (1994). For any $\varsigma > 0$, we have that with probability approaching one

$$\mathcal{C}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right) > \mathcal{C}_{\text{GMM},n_t}(\widehat{\boldsymbol{\alpha}}_{\text{IV}}) - \varsigma/3 > \mathcal{C}_{\text{GMM},n_t}(\boldsymbol{\alpha}) - 2\varsigma/3 > \mathcal{C}\left(\widehat{\boldsymbol{\alpha}}\right) - \varsigma \tag{37}$$

Next, define $\mathbf{B}$ any open subset of $\mathbf{A}$ that contains $\boldsymbol{\alpha}$, and $\mathbf{B}^c$ the complement of $\mathbf{B}$ in $\mathbb{R}^{MR}$. It then follows that $\mathbf{A} \cap \mathbf{B}^c$ is compact and $\text{argmax}_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}\cap\mathbf{B}^c}\ \mathcal{C}(\widehat{\boldsymbol{\alpha}})$ exists. Assumption GMM (a) and (b) imply that $\mathcal{C}(\boldsymbol{\alpha}) - \sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}\cap\mathbf{B}^c}\ \mathcal{C}(\widehat{\boldsymbol{\alpha}}) > 0$. Finally, fix $\varsigma = \mathcal{C}(\boldsymbol{\alpha}) - \sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}\cap\mathbf{B}^c}\ \mathcal{C}(\widehat{\boldsymbol{\alpha}})$ and plug into the right-hand side of (38), so that $\mathcal{C}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right) > \sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}\cap\mathbf{B}^c}\ \mathcal{C}(\widehat{\boldsymbol{\alpha}})$ with probability approaching one, implying that $\widehat{\boldsymbol{\alpha}}_{\text{IV}} \in \mathbf{B}$ and $\widehat{\boldsymbol{\alpha}}_{\text{IV}} \xrightarrow{p} \boldsymbol{\alpha}$. Finally, consistency of $\widehat{\boldsymbol{\alpha}}_{\text{IV}}$ implies $\widehat{\bar{\alpha}} \xrightarrow{p} \bar{\alpha} = -\alpha\left(\alpha'\alpha\right)^{-1}$ and $\widehat{\psi} \xrightarrow{p} \psi$. It then follows that $\widehat{\alpha}_{\perp,\text{IV}} \xrightarrow{p} \alpha_{\perp}$, which proves (i).

As for the proof of (ii), multiply (35) through by $\sqrt{n}$

$$\sqrt{n}\left(\widehat{\alpha}_{\perp,\text{IV}} - \alpha_{\perp}\right) = \mathcal{S}\left(\widehat{\psi}' \otimes \Xi\right) \boldsymbol{K}_{M,R}\Big[\sqrt{n}\Big(\widehat{\boldsymbol{\alpha}}_{\text{IV}} - \boldsymbol{\alpha}\Big)\Big]. \tag{38}$$

From item (i) in this theorem, $\widehat{\boldsymbol{\alpha}}_{\text{IV}} \xrightarrow{p} \boldsymbol{\alpha}$, which implies that $\bar{\alpha} \xrightarrow{p} \alpha\left(\alpha'\alpha\right)^{-1}$ and hence $\widehat{\psi} \xrightarrow{p} \psi$. It therefore suffices to show the limiting distribution of $\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}} - \boldsymbol{\alpha}\right)$. Next, the FOC of (36) implies

$$-2\boldsymbol{G}'_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right)\widehat{\Psi}^{-1}\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right) = 0, \tag{39}$$

where $\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}\right)$ denotes the sample moment condition computed with prices sampled on the equally spaced clock with sampling interval given by $\delta = 1/n$ and $\boldsymbol{G}_{n_t}\left(\widehat{\boldsymbol{\alpha}}\right) = \nabla_{\widehat{\boldsymbol{\alpha}}}\,\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}\right)$ stands for the gradient of $\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}\right)$. Expanding $\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right)$ around $\boldsymbol{\alpha}$ and rearranging terms

$$\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}} - \boldsymbol{\alpha}\right) = \Big[\boldsymbol{G}'_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right)\widehat{\Psi}^{-1}\boldsymbol{G}'_{n_t}\left(\widehat{\boldsymbol{\alpha}}^{*}\right)\Big]^{-1}\boldsymbol{G}'_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right)\widehat{\Psi}^{-1}\sqrt{n}\boldsymbol{g}_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right), \tag{40}$$

where $\widehat{\boldsymbol{\alpha}}^{*}$ is a point in the line segment joining $\widehat{\boldsymbol{\alpha}}_{\text{IV}}$ and $\boldsymbol{\alpha}$. First, note that $\widehat{\Psi} \xrightarrow{p} \Psi$ by HAC large sample theory, see for instance, Theorem 6.9 in White (2000). To show $\boldsymbol{G}_{n_t}\left(\widehat{\boldsymbol{\alpha}}_{\text{IV}}\right) \xrightarrow{p} \boldsymbol{G}\left(\boldsymbol{\alpha}\right)$ with

$\boldsymbol{G}(\boldsymbol{\alpha}) = \mathbb{E}\left[\nabla_{\boldsymbol{\alpha}} \boldsymbol{g}_{n_{t_i}}(\boldsymbol{\alpha})\right] = \boldsymbol{G}$ and $\boldsymbol{g}_{n_{t_i}}(\alpha) = \left(\Delta p_{\tau_i} - \alpha_\delta(\boldsymbol{\alpha})\beta' p_{\tau_{i-1}}\right) \otimes Z_{\bar{k},\tau_{i-\bar{q}-k}}$, take a mean value expansion of sample and population moment conditions around $\boldsymbol{\alpha}$. Next, take the supremum norm and apply the triangle inequality, such that

$$\sup_{\widehat{\boldsymbol{\alpha}}^* \in \mathbf{A}^*} \left|\boldsymbol{G}_{n_t}(\widehat{\boldsymbol{\alpha}}^*) - \boldsymbol{G}(\widehat{\boldsymbol{\alpha}}^*)\right| \quad \sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}} \left|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right| < \sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}} \left|\boldsymbol{g}_{n_t}(\widehat{\boldsymbol{\alpha}}) - \boldsymbol{g}(\widehat{\boldsymbol{\alpha}})\right| + \left|\boldsymbol{g}_{n_t}(\boldsymbol{\alpha}) - \boldsymbol{g}(\boldsymbol{\alpha})\right|, \tag{41}$$

where $\widehat{\boldsymbol{\alpha}}^*$ is a point in the line segment joining an arbitrary point $\widehat{\boldsymbol{\alpha}}$ in $\mathbf{A}$ and $\boldsymbol{\alpha}$, and $\mathbf{A}^*$ denotes an open convex set containing $\mathbf{A}$. Assumptions GMM (c) and RROU deliver that the two terms on the right-hand side of (41) and $\sup_{\widehat{\boldsymbol{\alpha}} \in \mathbf{A}} |\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}|$ are $o_p(1)$ and $O(1)$, respectively. It then follows that $\sup_{\widehat{\boldsymbol{\alpha}}^* \in \mathbf{A}^*} |\boldsymbol{G}_{n_t}(\widehat{\boldsymbol{\alpha}}^*) - \boldsymbol{G}(\widehat{\boldsymbol{\alpha}}^*)|$ is $o_p(1)$, meaning that $\boldsymbol{G}_{n_t}(\widehat{\boldsymbol{\alpha}}_{\mathrm{IV}})$ converges uniformly in probability to $\boldsymbol{G}(\boldsymbol{\alpha})$.

Next, recall that the moment conditions are serially correlated. From Lemma 4, $\left\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ is an ergodic uniform integrable L^1- mixingale sequences, meaning that we can apply the Cramér-Wold device to obtain the limiting distribution of $\sqrt{n}\boldsymbol{g}_{n_t}(\boldsymbol{\alpha})$: $\sqrt{n}\boldsymbol{g}_{n_t}(\boldsymbol{\alpha}) \xrightarrow{d} N(0,\Psi)$, where $\Psi = \sum_{j=-\bar{q}-1}^{\bar{q}+1} \Psi(j)$ and $\Psi(j) = \mathbb{E}\left[\mathrm{v}_{t_i}\mathrm{v}'_{t_{i-j}} \otimes Z_{\bar{k},t_{i-\bar{q}-k}} Z'_{\bar{k},t_{i-\bar{q}-k-j}}\right]$ (see, for instance, White, 2000; Davidson, 1994). As the weighting matrix is chosen optimally to be Ψ, the limiting distribution of (40) reads

$$\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}_{\mathrm{IV}} - \boldsymbol{\alpha}\right) \xrightarrow{d} N\left(0, \left(\boldsymbol{G}'\Psi^{-1}\boldsymbol{G}\right)^{-1}\right). \tag{42}$$

It then follows that

$$\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}_{\perp,\mathrm{IV}} - \boldsymbol{\alpha}_\perp\right) \xrightarrow{d} N\left(0, \mathcal{S}\left(\psi' \otimes \Xi\right)\left(\boldsymbol{G}'\Psi^{-1}\boldsymbol{G}\right)^{-1}\left(\psi' \otimes \Xi\right)' \mathcal{S}'\right), \tag{43}$$

which proves item ii.

To prove (iii), it suffices to show that $RK \xrightarrow{p} \Sigma$ as item (i) in this theorem holds. To this end, we must show that our assumptions satisfy the conditions in Barndorff-Nielsen et al.'s (2011) Lemma 1. This is indeed the case. The continuous-time process in (1) meets the conditions in their assumption SH. Assumptions RT and K are the same as their assumptions D and K, whereas their assumption U holds under Assumption MMN(TS). It follows by Slutsky's theorem that $IS_{\mathrm{IV}} - IS = o_p(1)$. ■

A.1 Proofs of Lemma 1, 2 and 3

In this appendix, we show that $Z_{\bar{k},t_{i-\bar{q}-k}}$ and $\widetilde{\beta}' s_{t_{i-1}}$ are valid sets of instruments respectively under Assumptions MMN(TS) and MMN(CS) in that they correlate with $\beta' p_{t_{i-1}}$ even if orthogonal to v_{t_i}. For ease of exposition, we restrict attention to the case with $M = J = 2$ markets and focus without any loss of generality on the estimation of $\alpha_{\delta,1}$ in (12) using either $Z_{\bar{k},t_{i-\bar{q}-k}} = p_{1,t_{i-\bar{q}-k}} - p_{2,t_{i-\bar{q}-k}}$ or $\beta_2' s_{t_{i-1}} = s_{1,t_{i-1}} - s_{2,t_{i-1}}$ as a single instrument.

Proof of Lemma 1 Lemma 1 is a collection of established results. Specifically, rank $(\alpha\beta')$ = rank $(\alpha_\delta\beta')$ and the equivalence between the continuous- and discrete-time log-likelihood functions follow from Theorem 1 in Kessler and Rahbek (2004), whereas the remaining results are in Johansen (1995).

Proof of Lemma 2 We begin with

$$\begin{aligned}
\mathbb{E}(Z_{\bar{k},t_{i-\bar{q}-k}}\,\mathrm{v}_{1,t_i}) &= \mathbb{E}\left\{(p_{1,t_{i-\bar{q}-k}} - p_{2,t_{i-\bar{q}-k}})\left[\alpha_{\delta,1}(u_{1,t_{i-1}} - u_{2,t_{i-1}}) + \Delta u_{1,t_i} + \varepsilon_{t_i}\right]\right\} \\
&= \alpha_{\delta,1}\mathbb{E}\left[(p_{1,t_{i-\bar{q}-k}} - p_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(p_{1,t_{i-\bar{q}-k}} - p_{2,t_{i-\bar{q}-k}})\Delta u_{1,t_i}\right] + \mathbb{E}\left[(p_{1,t_{i-\bar{q}-k}} - p_{2,t_{i-\bar{q}-k}})\varepsilon_{1,t_i}\right] \\
&= \alpha_{\delta,1}\mathbb{E}\left[(P_{1,t_{i-\bar{q}-k}} - P_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(P_{1,t_{i-\bar{q}-k}} - P_{2,t_{i-\bar{q}-k}})\Delta u_{1,t_i}\right] + \mathbb{E}\left[(P_{1,t_{i-\bar{q}-k}} - P_{2,t_{i-\bar{q}-k}})\varepsilon_{1,t_i}\right] \\
&\quad + \alpha_{\delta,1}\mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})\Delta u_{1,t_i}\right] + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})\varepsilon_{1,t_i}\right],
\end{aligned}$$

given that $p_{j,t_i} = P_{j,t_i} + u_{j,t_i}$. It then follows from $P_{1,t_i} - P_{2,t_i} = \sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h}}$ that

$$\begin{aligned}
\mathbb{E}(Z_{\bar{k},t_{i-\bar{q}-k}}\,\mathrm{v}_{1,t_i}) &= \alpha_{\delta,1}\mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}\Delta u_{1,t_i}\right] + \mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}\varepsilon_{1,t_i}\right] \\
&\quad + \alpha_{\delta,1}\mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})\Delta u_{1,t_i}\right] + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}} - u_{2,t_{i-\bar{q}-k}})\varepsilon_{1,t_i}\right] = 0
\end{aligned}$$

under Assumption MMN(TS). We next show that the relevance of the instruments depend on the persistence in the price differentials:

$$
\begin{aligned}
\mathbb{E}\left[Z_{\bar{k},t_{i-\bar{q}-k}}\left(p_{1,t_{i-1}}-p_{2,t_{i-1}}\right)\right] &= \mathbb{E}\left[(p_{1,t_{i-\bar{q}-k}}-p_{2,t_{i-\bar{q}-k}})(p_{1,t_{i-1}}-p_{2,t_{i-1}})\right] \\
&= \mathbb{E}\left[(P_{1,t_{i-\bar{q}-k}}-P_{2,t_{i-\bar{q}-k}})(P_{1,t_{i-1}}-P_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(P_{1,t_{i-\bar{q}-k}}-P_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}}-u_{2,t_{i-\bar{q}-k}})(P_{1,t_{i-1}}-P_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}}-u_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&= \mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}\sum_{h_1=0}^{\infty}\bar{\Upsilon}_{\delta,h_1}\varepsilon_{t_{i-1-h_1}}\right] \\
&\quad + \mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}}-u_{2,t_{i-\bar{q}-k}})\sum_{h_1=0}^{\infty}\bar{\Upsilon}_{\delta,h_1}\varepsilon_{t_{i-1-h_1}}\right] \\
&\quad + \mathbb{E}\left[(u_{1,t_{i-\bar{q}-k}}-u_{2,t_{i-\bar{q}-k}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right].
\end{aligned} \tag{44}
$$

Note that as $\sum_{h=0}^{\infty}\left|\bar{\Upsilon}_{\delta,h}\right| < \infty$ and using the properties of the trace operator, the first term in (44) equals

$$
\begin{aligned}
\mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-\bar{q}-k-h}}\sum_{h_1=0}^{\infty}\bar{\Upsilon}_{\delta,h_1}\varepsilon_{t_{i-1-h_1}}\right] &= \sum_{h=0}^{\infty}\sum_{h_1=0}^{\infty}\bar{\Upsilon}_{\delta,h}\mathbb{E}\left[\varepsilon'_{t_{i-\bar{q}-k-h}}\otimes\varepsilon_{t_{i-1-h_1}}\right]\bar{\Upsilon}'_{\delta,h_1} \\
&= \sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_{\delta}\bar{\Upsilon}'_{\delta,\bar{q}+k-1+h} < \infty.
\end{aligned} \tag{45}
$$

Finally, under Assumption MMN(TS), the last three terms in (44) are zero. It then follows that (45) holds for any $2 \leq k \leq \bar{k}$, which proves Lemma 2.

Proof of Lemma 3 We first establish the exogeneity of the instruments:

$$\begin{aligned}
\mathbb{E}\left[\left(s_{1,t_{i-1}}-s_{2,t_{i-1}}\right)\mathrm{v}_{1,t_i}\right] &= \mathbb{E}\left\{(s_{1,t_{i-1}}-s_{2,t_{i-1}})\left[\alpha_{\delta,1}(u_{1,t_{i-1}}-u_{2,t_{i-1}})+\Delta u_{1,t_i}+\varepsilon_{t_i}\right]\right\} \\
&= \alpha_{\delta,1}\mathbb{E}\left[(s_{1,t_{i-1}}-s_{2,t_{i-q-11}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(s_{1,t_{i-1}}-s_{2,t_{i-1}})\Delta u_{1,t_i}\right]+\mathbb{E}\left[(s_{1,t_{i-1}}-s_{2,t_{i-1}})\varepsilon_{1,t_i}\right] \\
&= \alpha_{\delta,1}\mathbb{E}\left[(S_{1,t_{i-1}}-S_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(S_{1,t_{i-1}}-S_{2,t_{i-1}})\Delta u_{1,t_i}\right]+\mathbb{E}\left[(S_{1,t_{i-1}}-S_{2,t_{i-1}})\varepsilon_{1,t_i}\right] \\
&\quad + \alpha_{\delta,1}\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})\Delta u_{1,t_i}\right]+\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})\varepsilon_{1,t_i}\right],
\end{aligned}$$

given that $s_{j,t_i} = S_{j,t_i} + \upsilon_{j,t_i}$ for $j \in \{1,2\}$. Hansen's (2005) representation then yields

$$S_{1,t_i} - S_{2,t_i} = \widetilde{\beta}'\widetilde{\Xi}_\delta \sum_{h=1}^{i} \epsilon_{t_h} + \sum_{h=0}^{\infty} \widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-h}} + \widetilde{\beta}'\widetilde{\Xi}_\delta S_0 = \sum_{h=0}^{\infty} \widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-h}}.$$

This implies that

$$\begin{aligned}
\mathbb{E}\left[\left(s_{1,t_{i-1}}-s_{2,t_{i-1}}\right)\mathrm{v}_{1,t_i}\right] &= \alpha_{\delta,1}\mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-q-1-h}}(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-q-1-h}}\Delta u_{1,t_i}\right]+\mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-q-1-h}}\epsilon_{1,t_i}\right] \\
&\quad + \alpha_{\delta,1}\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad + \mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})\Delta u_{1,t_i}\right]+\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})\varepsilon_{1,t_i}\right].
\end{aligned}$$

Assumption MMN(CS) ensures that all of the above terms are equal to zero. It remains to show the relevance of the instrument. To this end, we compute its correlation with the price differential between markets:

$$
\begin{aligned}
\mathbb{E}\left[\left(s_{1,t_{i-1}}-s_{2,t_{i-1}}\right)\left(p_{1,t_{i-1}}-p_{2,t_{i-1}}\right)\right] &= \mathbb{E}\left[(S_{1,t_{i-1}}-S_{2,t_{i-1}})(P_{1,t_{i-1}}-P_{2,t_{i-1}})\right] \\
&\quad +\mathbb{E}\left[(S_{1,t_{i-1}}-S_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad +\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})(P_{1,t_{i-1}}-P_{2,t_{i-1}})\right] \\
&\quad +\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&= \mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}}\sum_{h_1=0}^{\infty}\bar{\Upsilon}_{\delta,h_1}\varepsilon_{t_{i-1-h_1}}\right] \\
&\quad +\mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}}(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right] \\
&\quad +\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-1-h}}\right] \\
&\quad +\mathbb{E}\left[(\upsilon_{1,t_{i-1}}-\upsilon_{2,t_{i-1}})(u_{1,t_{i-1}}-u_{2,t_{i-1}})\right].
\end{aligned}
$$

The last three terms are all equal to zero under Assumption MMN(CS), whereas the first term is nonzero as long as the covariance between ε_{t_i} and ϵ_{t_i} differs from zero. Specifically, define $\mathbf{\Sigma}_\delta(0)=\mathbb{E}\left[\epsilon'_{t_{i-h}}\otimes\varepsilon_{t_{i-h}}\right]<\infty$ and $\mathbf{\Sigma}_\delta(h)=\mathbb{E}\left[\epsilon'_{t_i}\otimes\varepsilon_{t_{i-h}}\right]=0$ for $h>0$. As $\sum_{h=0}^{\infty}\left\|\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\right\|<\infty$ and $\sum_{h'=0}^{\infty}\left\|\bar{\Upsilon}_{\delta,h'}\right\|<\infty$, and using the properties of the trace operator,

$$
\mathbb{E}\left[\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}}\sum_{h'=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-1-h'}}\right]=\sum_{h=0}^{\infty}\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h}\mathbf{\Sigma}_\delta(0)\,\bar{\Upsilon}'_{\delta,h}<\infty, \tag{46}
$$

which proves Lemma 3.

Proof of Lemma 4 We start by showing that $\left\{\beta' p_{t_{i-\bar{q}-\underline{k}}}\right\}$ is a stationary ergodic sequence. First, rewrite $\beta' p_{t_{i-\bar{q}-\underline{k}}}$ as

$$
\begin{aligned}
\beta' p_{t_{i-\bar{q}-\underline{k}}} &= \sum_{h=0}^{\infty}\beta'\Upsilon_{\delta,h}\varepsilon_{t_{i-\bar{q}-\underline{k}-h}}+\sum_{h_1}^{\bar{q}}\beta'\varpi_{h_1}\vartheta_{t_{i-\bar{q}-\underline{k}-h_1}}, \\
&= \sum_{h=0}^{\bar{q}}\Theta_h d_{t_{i-\bar{q}-\underline{k}-h}}+\sum_{h=\bar{q}+1}^{\bar{c}-1}\bar{\Theta}_h d_{t_{i-\bar{q}-\underline{k}-h}}+\sum_{h=\bar{c}}^{\infty}\bar{\Theta}_h d_{t_{i-\bar{q}-\underline{k}-h}},
\end{aligned} \tag{47}
$$

where $\Theta_h=\left(\bar{\Upsilon}_h,\varpi_h\right)$ and $\bar{\Theta}_h=\left(\bar{\Upsilon}_h,0_M\right)$ are a $R\times 2M$ matrix with $\bar{\Upsilon}_h=\beta'\Upsilon_{\delta,h}$ and $d_{t_{i-\bar{q}-\underline{k}-h}}=\operatorname{vec}\left(\varepsilon_{t_{i-\bar{q}-\underline{k}-h}},\vartheta_{t_{i-\bar{q}-\underline{k}-h_1}}\right)$ is Gaussian white noise process with dimension $2M\times 1$. Because (47) is a Gaussian VMA(∞) process, it is stationary. Next, $\beta' p_{t_{i-\bar{q}-\underline{k}}}$ is ergodic if $\bar{\Theta}_h$, $h=\bar{q}+1,\ldots$, is absolutely summable, i.e., $\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\left\|\bar{\Theta}_h\right\|=0$. It then follows that

$$\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\left\|\bar{\Theta}_h\right\|\leq\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\sum_{h_1=1}^{R}\sum_{h_2=1}^{2M}\left|\bar{\Theta}_{h_1,h_2,h}\right|\leq\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\sum_{h_1=1}^{R}\sum_{h_2=1}^{M}\left|\bar{\Upsilon}_{h_1,h_2,h}\right|, \tag{48}$$

where $\left|\bar{\Theta}_{h_1,h_2,h}\right|$ and $\left|\bar{\Upsilon}_{h_1,h_2,h}\right|$ denote the absolute values of the h_1h_2-th elements of $\bar{\Theta}$ and $\bar{\Upsilon}$, respectively. It then follows from Assumption RROU that $\bar{\Upsilon}_h$, $h=\bar{c}+1,\ldots$, is absolutely summable, which implies that $\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}+1}^{\infty}\sum_{h_1}^{R}\sum_{h_2}^{M}\left|\bar{\Upsilon}_{h_1,h_2,h}\right|=0$. As for v_{m,t_i}, it follows from $\mathrm{v}_{t_i}=\varepsilon_{t_i}+[I_M-(\alpha_\delta\beta'+I_M)L]\,u_{t_i}$ that v_{m,t_i} is a Guassian MA($\bar{q}$+1) process, meaning that it is stationary and ergodic process. Under Assumption MMN(CS), first note that $\widetilde{\beta}'s_{t_{i-1}}=\widetilde{\beta}'S_{t_{i-1}}+\widetilde{\beta}'\upsilon_{t_{i-1}}$. Next, by applying the GRT, it follows that $\widetilde{\beta}'s_{t_{i-1}}=\sum_{h=0}^{\bar{c}-1}\widetilde{\Theta}_h\widetilde{d}_{t_{i-1-h}}+\sum_{h_1=\bar{c}}^{\infty}\widetilde{\Theta}_{h_1}\widetilde{d}_{t_{i-1-h_1}}$ is a stationary Gaussian VMA(∞) process, where $\widetilde{\Theta}_h=\left(\widetilde{\beta}'\widetilde{\Upsilon}_{\delta,h},\,\varrho_h\right)$ is a $(J-1)\times 2J$ matrix and $\widetilde{d}_{t_{i-1-h_1}}=\mathrm{vec}\left(\epsilon_{t_{i-1-h}},\upsilon_{t_{i-1-h}}\right)$ is a $2J\times 1$ vector. Furthermore, $\widetilde{\beta}'s_{t_{i-1}}$ is ergodic if $\lim_{\bar{c}\to\infty}\sum_{h_1=\bar{c}}^{\infty}\left\|\widetilde{\Theta}_{h_1}\right\|=0$:

$$\lim_{\bar{c}\to\infty}\sum_{h_1=\bar{c}}^{\infty}\left\|\widetilde{\Theta}_{h_1}\right\|\leq\lim_{\bar{c}\to\infty}\sum_{h_1=\bar{c}}^{\infty}\sum_{h_2=1}^{J-1}\left\{\sum_{h_3=1}^{J}\left|\left[\widetilde{\beta}'\widetilde{\Upsilon}_{h_1}\right]_{h_2,h_3}\right|+\sum_{h_3=1}^{J}\left|\varrho_{h_2,h_3,h_1}\right|\right\}. \tag{49}$$

As $\widetilde{\beta}'\widetilde{\Upsilon}_{h_1}$ and ϱ_{h_1} are both absolutely summable (Assumptions OTP and MMN(CS), respectively), the limit in (49) equals zero. Finally, we show that v_{t_i} is a stationary ergodic sequence by rewriting it as a Gaussian VMA(∞) process with absolutely summable coefficients: $\mathrm{v}_{t_i}=\bar{\Theta}_0d_{t_i}+\sum_{h=1}^{\infty}\bar{\Theta}_hd_{t_{i-h}}$, where $\bar{\Theta}_0=(I_M,\,I_M)$ and $\bar{\Theta}_h=\left(\left(\varpi_{h+1}-(I_M+\alpha_\delta\beta')\,\varpi_h\right),\,0_M\right)$ are $M\times 2M$ matrices. Absolute summability of $\bar{\Theta}_h$ follows directly from ϖ_h, $h=1,\ldots$, being absolutely summable:

$$\lim_{\bar{c}\to\infty}\sum_{h_1=\bar{c}}^{\infty}\left\|\bar{\Theta}_h\right\|\leq\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\sum_{h=1}^{M}\sum_{h_2=1}^{M}\left|\varpi_{m,h_1,h_2,h}\right|+\lim_{\bar{c}\to\infty}\sum_{h=\bar{c}}^{\infty}\sum_{h=1}^{M}\sum_{h_2=1}^{M}\sum_{h_3=1}^{M}\left|D_{h_1,h_3}\right|\left|\varpi_{h_3,h_2,h}\right|=0, \tag{50}$$

where D_{h_1,h_3} is the h_1,h_3 element of $D=(I_M+\alpha_\delta\beta')$.

To show (ii), first note that from item (i), $\left\{\beta'_{\underline{m}}p_{t_{i-\bar{q}-\underline{k}}}\mathrm{v}_{m,t_i}\right\}$ and $\left\{\widetilde{\beta}'_{\underline{j}}s_{t_{i-1}}\mathrm{v}_{m,t_i}\right\}$ are stationary ergodic sequences (see Proposition 3.36 in White, 2000). Next, recall that a sequence is said to be a L^1-mixingale with respect to $\mathcal{F}_{t_i}$ if $\mathbb{E}\left\|\mathbb{E}\left(\beta'_{\underline{m}}p_{t_{i-\bar{q}-\underline{k}}}\mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\|\leq f_{t_i}\xi_c$, where $\{f_{t_i}\}_{i=1}^{n}$ is a deterministic sequence and $\lim_{c\to\infty}\xi_c=0$. First, it follows from a standard application of the GRT that $\beta'_{\underline{m}}p_{t_{i-\bar{q}-k}}=\sum_{h=0}^{\infty}\beta'_{\underline{m}}\Upsilon_{\delta,h}\varepsilon_{t_{i-\bar{q}-\underline{k}-h}}+\beta'_{\underline{m}}u_{t_{i-\bar{q}-\underline{k}-h}}$ and the sequence $\beta'_{\underline{m}}\Upsilon_{\delta,h}$, $h=1,2,\ldots$, is absolutely summable. Under Assumption MMN(TS), it follows that

$$\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right) = 0, \quad \text{for } c > \bar{q}+1; \tag{51}$$

$$\begin{aligned}\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right) = & \sum_{h=0}^{\infty} \beta'_{\underline{m}} \Upsilon_{\delta,h} \varepsilon_{t_{i-\bar{q}-\underline{k}-h}} \left(-\varpi_{m,\bar{q}} \vartheta_{m,t_{i-1-\bar{q}}} - \alpha_{\delta,m} \beta' \varpi_{\bar{q}} \vartheta_{t_{i-1-\bar{q}}}\right) \\ & + \sum_{h_1=0}^{\bar{q}} \beta'_{\underline{m}} \varpi_{h_1} \vartheta_{t_{i-1-\bar{q}-\underline{k}}} \left(-\varpi_{m,\bar{q}} \vartheta_{m,t_{i-1-\bar{q}}}\right. \\ & \left.-\alpha_{\delta,m} \beta' \varpi_{\bar{q}} \vartheta_{t_{i-1-\bar{q}}}\right), \quad \text{for } c = \bar{q}+1;\end{aligned} \tag{52}$$

$$\begin{aligned}\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right) = & \sum_{h=0}^{\infty} \beta'_{\underline{m}} \Upsilon_{\delta,h} \varepsilon_{t_{i-\bar{q}-\underline{k}-h}} \left(\sum_{h_2=c}^{\bar{q}} \varpi_{m,h_2} \vartheta_{m,t_{i-h_2}} - \sum_{h_2=c-1}^{\bar{q}} \varpi_{m,h_2} \vartheta_{m,t_{i-1-h_2}}\right. \\ & \left. - \sum_{h_2=c-1}^{\bar{q}} \varpi_{m,h_2} \alpha_{\delta,m} \beta' \varpi_{h_2} \vartheta_{m,t_{i-h_2}}\right) + \sum_{h_1=0}^{\bar{q}} \beta'_{\underline{m}} \varpi_{h_1} \vartheta_{t_{i-1-\bar{q}-\underline{k}}} \\ & \left(\sum_{h_2=c}^{\bar{q}} \varpi_{m,h_2} \vartheta_{m,t_{i-h_2}} - \sum_{h_2=c-1}^{\bar{q}} \varpi_{m,h_2} \vartheta_{m,t_{i-1-h_2}} - \right. \\ & \left. \sum_{h_2=c-1}^{\bar{q}} \varpi_{m,h_2} \alpha_{\delta,m} \beta' \varpi_{h_2} \vartheta_{m,t_{i-h_2}}\right), \quad \text{for } 1 \leq c \leq \bar{q},\end{aligned} \tag{53}$$

where $\alpha_{\delta,m}$ denotes the mth row of α_δ. From (51), it follows that $\mathbb{E}\left\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\right\| = 0$ for $c > \bar{q}+1$, which implies $\xi_c = 0$ and that it suffices to show that $\mathbb{E}\left\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\right\|$ is finite for $1 \leq c \leq \bar{q}+1$. As for $c = \bar{q}+1$, first recall that the sequence $\beta'_{\underline{m}} \Upsilon_{\delta,h}$, $h = 1, 2, \ldots$ is absolutely summable, that is $\sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| < \infty$, and thus we can interchange the order of the expectation operator. It then follows from (52) that

$$\begin{aligned}\mathbb{E}\left\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\right\| \leq & \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\varpi_{m,\bar{q}}\right\| \mathbb{E}\Big\{ \left\|\varepsilon_{t_{i-\bar{q}-\underline{k}-h}} \vartheta_{m,t_{i-1-\bar{q}}}\right\| \Big\} \\ & + \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\mathrm{vec}\left(\alpha_{\delta,m} \beta' \varpi_{\bar{q}}\right)\right\| \mathbb{E}\Big\{ \left\|\vartheta'_{t_{i-1-\bar{q}}} \otimes \varepsilon_{t_{i-\bar{q}-\underline{k}-h}}\right\| \Big\} \\ & + \sum_{h_1=0}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_1}\right\| \left\|\varpi_{m,\bar{q}}\right\| \mathbb{E}\Big\{ \left\|\vartheta_{t_{i-\bar{q}-\underline{k}-h_1}} \vartheta_{m,t_{i-1-\bar{q}}}\right\| \Big\} \\ & + \sum_{h_1=0}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_1}\right\| \left\|\mathrm{vec}\left(\alpha_{\delta,m} \beta' \varpi_{\bar{q}}\right)\right\| \mathbb{E}\Big\{ \left\|\vartheta'_{m,t_{i-1-\bar{q}}} \otimes \vartheta_{t_{i-\bar{q}-\underline{k}-h_1}}\right\| \Big\}.\end{aligned} \tag{54}$$

All terms on the right-hand side involve expectations taking from products of Gaussian variables and hence there exists finite constants $D_1, D_2 < \infty$ such that

$$\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| \leq \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| D_1 + \sum_{h'=0}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h'}\right\| D_2, \tag{55}$$

which implies $\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| < \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| D$ for some $D < \infty$. Next, we show that $\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| < \infty$ for $1 \leq c \leq \bar{q}$. From (53), it follows that

$$\begin{aligned}
\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| \leq & \sum_{h=0}^{\infty} \sum_{h_2=c}^{\bar{q}} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\varpi_{m,h_2}\right\| \mathbb{E}\Big\{\varepsilon_{t_{i-\bar{q}-\underline{k}-h}} \vartheta_{m,t_{i-h_2}}\Big\} \\
& + \sum_{h=0}^{\infty} \sum_{h_2=c-1}^{\bar{q}} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\varpi_{m,h_2}\right\| \mathbb{E}\Big\{\varepsilon_{t_{i-\bar{q}-\underline{k}-h}} \vartheta_{m,t_{i-1-h_2}}\Big\} \\
& + \sum_{h=0}^{\infty} \sum_{h_2=c-1}^{\bar{q}} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\mathrm{vec}\ \left(\alpha_{\delta,m} \beta' \varpi_{h_2}\right)\right\| \mathbb{E}\Big\{\vartheta'_{t_{i-h_2}} \otimes \varepsilon_{t_{i-\bar{q}-\underline{k}-h}}\Big\} \\
& + \sum_{h_1=0}^{\bar{q}} \sum_{h_2=c}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_2}\right\| \left\|\varpi_{m,h_2}\right\| \mathbb{E}\Big\{\vartheta_{t_{i-\bar{q}-\underline{k}-h_1}} \vartheta_{m,t_{i-h_2}}\Big\} \\
& + \sum_{h_1=0}^{\bar{q}} \sum_{h_2=c-1}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_2}\right\| \left\|\varpi_{m,h_2}\right\| \mathbb{E}\Big\{\vartheta_{t_{i-\bar{q}-\underline{k}-h_1}} \vartheta_{m,t_{i-1-h_2}}\Big\} \\
& + \sum_{h_1=0}^{\bar{q}} \sum_{h_2=c-1}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_2}\right\| \left\|\mathrm{vec}\ \left(\alpha_{\delta,m} \beta' \varpi_{h_2}\right)\right\| \mathbb{E}\Big\{\vartheta'_{t_{i-1-h_2}} \otimes \vartheta_{t_{i-\bar{q}-\underline{k}-h_1}}\Big\}.
\end{aligned}$$

There exists finite constants $D_1, D_2, D_3 < \infty$ such that

$$\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| \leq \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| D_1 + \sum_{h_1=0}^{\bar{q}} \left\|\beta'_{\underline{m}} \varpi_{h_2}\right\| D_2 \leq \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| D_3$$

Finally, as $\sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| < \infty$, it then follows that $\mathbb{E}\Big\|\mathbb{E}\left(\beta'_{\underline{m}} p_{t_{i-\bar{q}-\underline{k}}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\Big\| < \infty$, which proves item (ii).

Next, we show that $\left\{\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ is stationary ergodic adapted L^1-mixingale sequence.

To show that $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)\right\| \leq f_{t_i} \xi_c$ with $\lim_{c\to\infty} \xi_c = 0$, rewrite $\mathbb{E}\left(\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right)$ as

$$\mathbb{E}\left(\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right) = \mathbb{E}\left(\widetilde{\beta}'_{\underline{j}} \widetilde{\Upsilon}_{\delta,h} \epsilon_{t_{i-1-h}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right) + \mathbb{E}\left(\widetilde{\beta}'_{\underline{j}} \upsilon_{t_{i-1}} \mathrm{v}_{m,t_i} | \mathcal{F}_{t_{i-c}}\right), \tag{56}$$

where the first and second conditional expectations in (56) read:

$$\sum_{h=c-1}^{\infty} \sum_{h_1=c-1}^{\infty} \widetilde{\beta}'_{\underline{j}} \widetilde{\Upsilon}_{\delta,h} \epsilon_{t_{i-1-h}} \left(\varpi_{m,h_1+1} \vartheta_{ti+1-h_1} - \varpi_{m,h_1} \vartheta_{ti-h_1} - \alpha_{\delta,m} \beta' \varpi_{h_1} \vartheta_{ti-h_1}\right) \quad \text{and}$$

$$\sum_{h=c-1}^{\infty} \sum_{h_1=c-1}^{\infty} \widetilde{\beta}'_{\underline{j}} \varrho_h \varphi_{t_{i-1-h}} \left(\varpi_{m,h_1+1} \vartheta_{ti+1-h_1} - \varpi_{m,h_1} \vartheta_{ti-h_1} - \alpha_{\delta,m} \beta' \varpi_{h_1} \vartheta_{ti-h_1}\right),$$

respectively. Then, bound $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' s_{t_{i-1}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\|$ as $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' s_{t_{i-1}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| \leq \mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' \widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| + \left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' \upsilon_{t_{i-1}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\|$. The firts term reads

$$\begin{aligned}
\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' \widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| \leq & \sum_{h=c-1}^{\infty}\sum_{h_1=c-1}^{\infty}\left\|\widetilde{\beta}_{\underline{j}}'\widetilde{\Upsilon}_{\delta,h}\right\|\left\|\varpi_{m,h_1+1}\right\|\mathbb{E}\left\|\epsilon_{t_{i-1-h}}\vartheta_{ti+1-h_1}\right\| \\
& + \sum_{h=c-1}^{\infty}\sum_{h_1=c-1}^{\infty}\left\|\widetilde{\beta}_{\underline{j}}'\widetilde{\Upsilon}_{\delta,h}\right\|\left\|\varpi_{m,h_1}\right\|\mathbb{E}\left\|\epsilon_{t_{i-1-h}}\vartheta_{ti-h_1}\right\| \\
& + \sum_{h=c-1}^{\infty}\sum_{h_1=c-1}^{\infty}\left\|\widetilde{\beta}_{\underline{j}}'\widetilde{\Upsilon}_{\delta,h}\right\|\left\|\operatorname{vec}\left(\alpha_{\delta,m}\beta'\varpi_{h_1}\right)\right\|\mathbb{E}\left\|\epsilon_{t_{i-1-h}}\vartheta_{ti-h_1}\right\|.
\end{aligned}$$

Because $\widetilde{\Upsilon}_{\delta,h}$, $h = 1,\dots$ and ϖ_{h_1}, $h_1 = 1,\dots$ are absolutely summable, there exists an absolutely summable sequence $\bar{D}_{1,h}$ and finite constant D_1, such that $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' \widetilde{\Upsilon}_{\delta,h}\epsilon_{t_{i-1-h}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| \leq \sum_{h=c-1}^{\infty}\left\|\bar{D}_{1,h}\right\| D_1$. As for the second term in (56), as ϱ_h, $h = 1\dots$ is absolutely summable, we apply the same bounds such that there exists an absolutely summable sequence $\bar{D}_{2,h}$ and finite constant D_2, such that $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' \upsilon_{t_{i-1}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| \leq \sum_{h=c-1}^{\infty}\left\|\bar{D}_{2,h}\right\| D_2$. Finnaly, it follows that there exists an absolutely summable sequence $\bar{D}_h$ and finite constant D, such that $\mathbb{E}\left\|\mathbb{E}\left(\widetilde{\beta}_{\underline{j}}' s_{t_{i-1}} \mathrm{v}_{m,t_i}|\mathcal{F}_{t_{i-c}}\right)\right\| \leq \sum_{h=c-1}^{\infty}\left\|\bar{D}_h\right\| D$ with $f_{t_i} = D$ and $\xi_c = \sum_{h=c-1}^{\infty}\left\|\bar{D}_h\right\|$, which implies that $\lim_{c\to\infty}\xi_c = 0$.

To show that $\left\{\beta_{\underline{m}}' p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ is uniformly integrable, it suffices to prove that $\mathbb{E}\left\{\beta_{\underline{m}}' p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}' p_{t_{i-\bar{q}-k}}'\beta_{\underline{m}}\right\} < \infty$, (see, for instance, Proposition 7.7 in Hamilton, 1994). Rewriting it as

$$\mathbb{E}\left\{\left(\sum_{h=0}^{\infty}\beta_{\underline{m}}'\Upsilon_{\delta,h}\varepsilon_{t_{i-\bar{q}-\underline{k}-h}} + \beta_{\underline{m}}' u_{t_{i-\bar{k}-\underline{k}}}\right)\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\left(\sum_{h=0}^{\infty}\beta_{\underline{m}}'\Upsilon_{\delta,h}\varepsilon_{t_{i-\bar{q}-\underline{k}-h}} + \beta_{\underline{m}}' u_{t_{i-\bar{k}-\underline{k}}}\right)'\right\}, \tag{57}$$

it follows under Assumption MMN(TS) that

$$\begin{aligned}
\mathbb{E}\left\{\beta_{\underline{m}}' p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}' p_{t_{i-\bar{q}-k}}'\beta_{\underline{m}}\right\} = & \sum_{h=0}^{\infty}\sum_{h_1=0}^{\infty}\beta_{\underline{m}}'\Upsilon_{\delta,h}\mathbb{E}\left\{\varepsilon_{t_{i-\bar{q}-\underline{k}-h}}\mathbb{E}_{t_{i-\bar{q}-k}}\left\{\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\right\}\varepsilon_{t_{i-\bar{q}-\underline{k}-h_1}}'\right\}\Upsilon_{\delta,h_1}'\beta_{\underline{m}} \\
& + \sum_{h=0}^{\infty}\beta_{\underline{m}}'\Upsilon_{\delta,h}\mathbb{E}\left\{\varepsilon_{t_{i-\bar{q}-\underline{k}-h}}\mathbb{E}_{t_{i-\bar{q}-k}}\left\{\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\right\}u_{t_{i-\bar{q}-\underline{k}}}'\right\}\beta_{\underline{m}} \\
& + \sum_{h=0}^{\infty}\beta_{\underline{m}}'\mathbb{E}\left\{u_{t_{i-\bar{q}-\underline{k}}}\mathbb{E}_{t_{i-\bar{q}-k}}\left\{\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\right\}\varepsilon_{t_{i-\bar{q}-\underline{k}-h}}'\right\}\Upsilon_{\delta,h}'\beta_{\underline{m}} \\
& + \beta_{\underline{m}}'\mathbb{E}\left\{u_{t_{i-\bar{q}-\underline{k}}}\mathbb{E}_{t_{i-\bar{q}-k}}\left\{\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\right\}u_{t_{i-\bar{q}-\underline{k}}}'\right\}\beta_{\underline{m}}.
\end{aligned} \tag{58}$$

Next, v_{m,t_i} is a Gaussian MA$(\bar{q}+1)$ process, which implies that $\mathbb{E}_{t_{i-\bar{q}-k}}\left\{\mathrm{v}_{m,t_i}\mathrm{v}_{m,t_i}'\right\} = \sigma_{\mathrm{v},\bar{q}+1}^2 < \infty$

exists. It then follows from applying the unconditional expectations in (58),

$$\mathbb{E}\left\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i} \mathrm{v}'_{m,t_i} p'_{t_{i-\bar{q}-k}} \beta_{\underline{m}}\right\} = \sum_{h=0}^{\infty} \sigma^2_{\mathrm{v},\bar{q}+1} \beta'_{\underline{m}} \Upsilon_{\delta,h} \Sigma_\delta \Upsilon'_{\delta,h_1} \beta_{\underline{m}} + \sum_{h=0}^{\bar{q}} \sigma^2_{\mathrm{v},\bar{q}+1} \beta'_{\underline{m}} \Upsilon_{\delta,h} \Gamma_{\varepsilon,u}(h) \beta_{\underline{m}}$$
$$+ \sum_{h=0}^{\bar{q}} \sigma^2_{\mathrm{v},\bar{q}+1} \beta'_{\underline{m}} \Gamma_{\varepsilon,u}(h) \Upsilon'_{\delta,h} \beta_{\underline{m}} + \sigma^2_{\mathrm{v},\bar{q}+1} \beta'_{\underline{m}} \Omega_u \beta_{\underline{m}}, \tag{59}$$

where $\Gamma_{\varepsilon,u}(h) = \mathbb{E}\left\{\varepsilon_{t_{i-h}} u'_{t_i}\right\} < \infty$ for $0 \leq h \leq \bar{q}$, $\Gamma_{\varepsilon,u}(h) = 0_M$ for $h > \bar{q}$, and $\Omega_u = \mathbb{E}\left\{u_{t_i} u'_{t_i}\right\} < \infty$. The last three terms in (59) are finite sums,

$$\mathbb{E}\left\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i} \mathrm{v}'_{m,t_i} p'_{t_{i-\bar{q}-k}} \beta_{\underline{m}}\right\} \leq \sum_{h=0}^{\infty} \left\|\sigma^2_{\mathrm{v},\bar{q}+1}\right\| \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\Sigma_\delta\right\| \left|\left|\Upsilon'_{\delta,h_1} \beta_{\underline{m}}\right|\right| + D_1,$$
$$\leq \sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| \left\|\Upsilon'_{\delta,h_1} \beta_{\underline{m}}\right\| D \tag{60}$$

for some $D_1, D < \infty$. As $\sum_{h=0}^{\infty} \left\|\beta'_{\underline{m}} \Upsilon_{\delta,h}\right\| < \infty$, we conclude that $\left\{\beta'_{\underline{m}} p_{t_{i-\bar{q}-k}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ is uniformly integrable under Assumption MMN(TS). As for Assumption MMN(CS), $\left\{\widetilde{\beta}'_{\underline{j}} s_{t_{i-1}} \mathrm{v}_{m,t_i}, \mathcal{F}_{t_i}\right\}$ is uniformly integrable from a straight forward modification of the proof in Assumption MMN(TS).

Appendix B Bias of the LS estimator

In this appendix, we derive the bias of the LS estimator in the presence of microstructure noise. To simplify matters, we consider only $M = 2$ markets and focus on the estimation of $\alpha_{\delta,1}$ in

$$\Delta p_{1,t_i} = \alpha_{\delta,1}\left(p_{1,t_{i-1}} - p_{2,t_{i-1}}\right) + \mathrm{v}_{1,t_i},$$

where $\mathrm{v}_{1,t_i} = \Delta u_{1,t_i} + \varepsilon_{1,t_i} - \alpha_{\delta,1}\left(u_{1,t_{i-1}} - u_{2,t_{i-1}}\right)$. The LS estimator of $\alpha_{\delta,1}$ is given by

$$\begin{aligned}
\hat{\alpha}_{\delta,1} &= \left[\sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}})^2\right]^{-1} \sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}}) \Delta p_{1,t_i} \\
&= \left[\frac{1}{n} \sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}})^2\right]^{-1} \frac{1}{n} \sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}}) \left[\alpha_{\delta,1}(p_{1,t_{i-1}} - p_{2,t_{i-1}}) + \mathrm{v}_{1,t_i}\right] \\
&= \alpha_{\delta,1} + \left[\frac{1}{n} \sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}})^2\right]^{-1} \frac{1}{n} \sum_{i=1}^{n} (p_{1,t_{i-1}} - p_{2,t_{i-1}}) \mathrm{v}_{1,t_i}.
\end{aligned}$$

The average squared difference in prices converges in probability to some positive finite value Q, and hence $\operatorname*{plim}_{n\to\infty} \left[\frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})^2\right]^{-1} = Q^{-1} < \infty$. This means that

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} \hat{\alpha}_{\delta,1} &= \alpha_{\delta,1} + Q^{-1} \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})\mathrm{v}_{1,t_i} \\
&= \alpha_{\delta,1} - \alpha_{\delta,1}\, Q^{-1} \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})(u_{1,t_{i-1}} - u_{2,t_{i-1}}) \\
&\quad + Q^{-1} \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta u_{1,t_i} + Q^{-1} \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})\varepsilon_{1,t_i}.
\end{aligned}$$

We next consider each term individually. Let

$$\begin{aligned}
B_{1,n} &= \frac{1}{n}\sum_{i=1}^{n}(p_{1,t_{i-1}} - p_{2,t_{i-1}})(u_{1,t_{i-1}} - u_{2,t_{i-1}}) \\
&= \frac{1}{n}\sum_{i=1}^{n}(P_{1,t_{i-1}} - P_{2,t_{i-1}})(u_{1,t_{i-1}} - u_{2,t_{i-1}}) + \frac{1}{n}\sum_{i=1}^{n}(u_{1,t_{i-1}} - u_{2,t_{i-1}})^2
\end{aligned}$$

Corollary 2 in Hansen (2005) offers the following alternative representation to (6):

$$P_{t_i} = \Xi_\delta \sum_{h=1}^{i}\varepsilon_{t_h} + \sum_{h=0}^{\infty}\Upsilon^*_{\delta,h}\varepsilon_{t_{i-h}} + \Xi_\delta\, P_0$$

where $\Xi_\delta = \beta_\perp(\alpha'_{\delta\perp}\beta_\perp)^{-1}\alpha'_{\delta\perp}$ and $\Upsilon^*_{\delta,h} = (I_2 - \Xi_\delta)(I_2 + \alpha_\delta\beta')^h$. It then follows that

$$\begin{aligned}
P_{1,t_i} - P_{2,t_i} &= (\Xi_{\delta,1} - \Xi_{\delta,2})\sum_{h=1}^{i}\varepsilon_{t_h} + \sum_{h=0}^{\infty}(\Upsilon^*_{\delta,1,h} - \Upsilon^*_{\delta,2,h})\varepsilon_{t_{i-h}} + (\Xi_{\delta,1} - \Xi_{\delta,2})P_0 \\
&= \sum_{h=0}^{\infty}(\Upsilon^*_{\delta,1,h} - \Upsilon^*_{\delta,2,h})\varepsilon_{t_{i-h}} = \sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h}}
\end{aligned}$$

given that $\Xi_{\delta,1} = \Xi_{\delta,2}$. This means that

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} B_{1,n} &= \mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h-1}}(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] + \mathbb{E}(u_{1,t_{i-1}} - u_{2,t_{i-1}})^2 \\
&= \sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}[\rho_{\cdot,1}(h) - \rho_{\cdot,2}(h)] + \omega_1^2 + \omega_2^2 - 2\,\omega_{1,2},
\end{aligned}$$

where $\rho_{\cdot,j}(h) = \mathbb{E}(\varepsilon_{t_{i-h}}u_{j,t_i})$ for $j = 1,2$ and $\omega_{1,2} = \mathbb{E}(u_{1,t_i}u_{2,t_i})$. The first component reflects the dependence between efficient price and market microstructure noises, whereas the remaining terms correspond to the extra variation due to the presence of market microstructure noises. In turn,

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} B_{2,n} &= \mathbb{E}\left[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta u_{1,t_i}\right] \\
&= \mathbb{E}\left[(P_{1,t_{i-1}} - P_{2,t_{i-1}})\Delta u_{1,t_i}\right] + \mathbb{E}\left[(u_{1,t_{i-1}} - u_{2,t_{i-1}})\Delta u_{1,t_i}\right] \\
&= \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\mathbb{E}\left[\varepsilon_{t_{i-h-1}}\Delta u_{1,t_i}\right] + \mathbb{E}\left[(u_{1,t_{i-1}} - u_{2,t_{i-1}})\Delta u_{1,t_i}\right] \\
&= \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\left[\rho_{\cdot,1}(h+1) - \rho_{\cdot,1}(h)\right] + \omega_{1,2} - \omega_1^2 + [\gamma_{1,1}(1) - \gamma_{1,2}(1)].
\end{aligned}$$

The last term in brackets displays an additional source of bias due to cross-autocorrelation in the microstructure noises. Lastly,

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} B_{3,n} &= \mathbb{E}\left[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\varepsilon_{1,t_i}\right] \\
&= \mathbb{E}\left[(P_{1,t_{i-1}} - P_{2,t_{i-1}})\varepsilon_{1,t_i}\right] + \mathbb{E}\left[(u_{1,t_{i-1}} - u_{2,t_{i-1}})\varepsilon_{1,t_i}\right] \\
&= \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\mathbb{E}\left[\varepsilon_{t_{i-h-1}}\varepsilon_{1,t_i}\right] + \mathbb{E}\left[u_{1,t_{i-1}}\varepsilon_{1,t_i}\right] - \mathbb{E}\left[u_{2,t_{i-1}}\varepsilon_{1,t_i}\right] = 0
\end{aligned}$$

given that price innovations are white noises that may lead, but not lag, microstructure noises. The LS bias then amounts to

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty}(\hat{\alpha}_{\delta,1} - \alpha_{\delta,1}) = Q^{-1}&\left\{\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}[\rho_{\cdot,1}(h+1) - \rho_{\cdot,1}(h)] + \gamma_{1,1}(1) - \gamma_{1,2}(1) + \omega_{1,2} - \omega_1^2\right\} \\
&- \alpha_{\delta,1}Q^{-1}\left\{\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}[\rho_{\cdot,1}(h) - \rho_{\cdot,2}(h)] + \omega_1^2 + \omega_2^2 - 2\,\omega_{1,2}\right\}.
\end{aligned}$$

Altogether, there are four sources of bias: (1) dependence between efficient price and market microstructure noises; (2) variation in market microstructure noises; (3) correlation between microstructure noises across markets; and (4) persistence in the microstructure noises. Finally, we obtain a similar result to the overall variation in the observed price differential:

$$\begin{aligned}
Q &= \mathbb{E}(p_{1,t_i} - p_{2,t_i})^2 = \mathbb{E}\left[P_{1,t_i} - P_{2,t_i} + u_{1,t_i} - u_{2,t_i}\right]^2 = \mathbb{E}\left[\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h}} + u_{1,t_i} - u_{2,t_i}\right]^2 \\
&= \mathbb{E}\left[\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h}}\right]^2 + \mathbb{E}\left[u_{1,t_i} - u_{2,t_i}\right]^2 + 2\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\mathbb{E}\left[\varepsilon_{t_{i-h}}(u_{1,t_i} - u_{2,t_i})\right] \\
&= \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2 - 2\,\omega_{1,2} + 2\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}[\rho_{\cdot,1}(h) - \rho_{\cdot,2}(h)].
\end{aligned}$$

The first term corresponds to the genuine part of the variation in the price differential, that is to say, what we would get if you could observe prices without any market microstructure contamination.

The omega terms stem from the mere presence of (possibly correlated) market microstructure noises, whereas the last term reflects the additional contribution they entail if persistent over time.

B.1 Independent microstructure (white) noises

We now entertain a very simple setting in which microstructure noises are unrealistically well-behaved and independent across markets. Suppose that $u_{1,t_i}, \ldots, u_{M,t_i}$ are independent white noise processes with mean zero and covariance matrix diag $(\omega_1^2, \ldots, \omega_M^2)$. This means that $\mathrm{cov}(u_{m,t_i}, u_{m',t_i}) = 0$ and that $\gamma_{m,m'}(h) = 0$ for $1 \le m, m' \le M$ and any $h > 0$. In addition, assume there is no dependence with the efficient price. The LS bias reduces to

$$\operatorname*{plim}_{n\to\infty}(\hat{\alpha}_{\delta,1} - \alpha_{\delta,1}) = -Q^{-1}\left[(1+\alpha_{\delta,1})\omega_1^2 + \alpha_{\delta,1}\omega_2^2\right],$$

with

$$Q = \sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta \bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2.$$

This yields

$$\operatorname*{plim}_{n\to\infty}(\hat{\alpha}_{\delta,1} - \alpha_{\delta,1}) = -\frac{\omega_1^2}{\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta \bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2} - \alpha_{\delta,1}\frac{\omega_1^2 + \omega_2^2}{\sum_{h=0}^{\infty} \bar{\Upsilon}_{\delta,h}\Sigma_\delta \bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2}.$$

The bias of the LS estimator then depends on the magnitude of the market microstructure noise in that market as well as on the overall noise-to-signal ratio.

B.2 Bias of the LS estimator in VEC models with lags

We study what happens if one applies the usual fix of adding lags to the VEC model in order to deal with residual autocorrelation. For simplicity, we consider only the unrealistic case of independent microstructure (white) noises. Adding the past realization of the change in prices results in

$$\Delta p_{1,t_i} = \alpha_{\delta,1}\left(p_{1,t_{i-1}} - p_{2,t_{i-1}}\right) + \gamma\,\Delta p_{1,t_{i-1}} + \mathrm{v}_{1,t_i}.$$

The LS estimator of $\alpha_{\delta,1}$ then converges in probability to

$$\operatorname*{plim}_{n\to\infty}\hat{\alpha}_{\delta,1} = \frac{\mathbb{E}(\Delta p_{1,t_{i-1}}^2)\mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta p_{1,t_i}] - \mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta p_{1,t_{i-1}}]\mathbb{E}(\Delta p_{1,t_i}\Delta p_{1,t_{i-1}})}{\mathbb{E}(\Delta p_{1,t_{i-1}}^2)\mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})^2] - \{\mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta p_{1,t_{i-1}}]\}^2}.$$

In what follows, we treat each of the above expectations individually in order to characterize the bias. Recall that

$$\begin{aligned}\mathbb{E}[(p_{1,t_{i-1}}-p_{2,t_{i-1}})\Delta p_{1,t_i}] &= \mathbb{E}\left\{(p_{1,t_{i-1}}-p_{2,t_{i-1}})\left[\alpha_{\delta,1}(p_{1,t_{i-1}}-p_{2,t_{i-1}})+\mathrm{v}_{1,t_i}\right]\right\}\\ &= \alpha_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})^2\right]+\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})\mathrm{v}_{1,t_i}\right]\\ &= \alpha_{\delta,1}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h}+\omega_1^2+\omega_2^2\right]-(1+\alpha_{\delta,1})\omega_1^2-\alpha_{\delta,1}\omega_2^2\\ &= \alpha_{\delta,1}\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h}-\omega_1^2.\end{aligned}$$

whereas

$$\begin{aligned}\mathbb{E}[(p_{1,t_{i-1}}-p_{2,t_{i-1}})\Delta p_{1,t_{i-1}}] &= \mathbb{E}\left\{(p_{1,t_{i-1}}-p_{2,t_{i-1}})\left[\alpha_{\delta,1}(p_{1,t_{i-2}}-p_{2,t_{i-2}})+\mathrm{v}_{1,t_{i-1}}\right]\right\}\\ &= \alpha_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})(p_{1,t_{i-2}}-p_{2,t_{i-2}})\right]+\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})\mathrm{v}_{1,t_{i-1}}\right]\\ &= \alpha_{\delta,1}\,\mathbb{E}\left[(P_{1,t_{i-1}}-P_{2,t_{i-1}})(P_{1,t_{i-2}}-P_{2,t_{i-2}})\right]+\mathbb{E}\left[(P_{1,t_{i-1}}-P_{2,t_{i-1}})\mathrm{v}_{1,t_{i-1}}\right]\\ &\quad+\mathbb{E}\left[(u_{1,t_{i-1}}-u_{2,t_{i-1}})\mathrm{v}_{1,t_{i-1}}\right]\\ &= \alpha_{\delta,1}\,\mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h-1}}\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h-2}}\right]+\mathbb{E}\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\varepsilon_{t_{i-h-1}}\Delta u_{1,t_{i-1}}\right]\\ &\quad+\mathbb{E}\left[(u_{1,t_{i-1}}-u_{2,t_{i-1}})\Delta u_{1,t_{i-1}}\right]\\ &= \alpha_{\delta,1}\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h+1}+\bar{\Upsilon}_{\delta,0}\begin{pmatrix}\sigma^2_{\delta,1}\\ \sigma_{\delta,12}\end{pmatrix}+\omega_1^2\end{aligned}$$

with $\sigma^2_{\delta,1}=\mathbb{E}(\varepsilon^2_{1,t_i})$ and $\sigma_{\delta,12}=\mathbb{E}(\varepsilon_{1,t_i}\varepsilon_{2,t_i})$. In turn,

$$\begin{aligned}\mathbb{E}(\Delta p_{1,t_i}\Delta p_{1,t_{i-1}}) &= \mathbb{E}\left\{\left[\alpha_{\delta,1}(p_{1,t_{i-1}}-p_{2,t_{i-1}})+\mathrm{v}_{1,t_i}\right]\left[\alpha_{\delta,1}(p_{1,t_{i-2}}-p_{2,t_{i-2}})+\mathrm{v}_{1,t_{i-1}}\right]\right\}\\ &= \alpha^2_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})(p_{1,t_{i-2}}-p_{2,t_{i-2}})\right]+\alpha_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-1}}-p_{2,t_{i-1}})\mathrm{v}_{1,t_{i-1}}\right]\\ &\quad+\alpha_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-2}}-p_{2,t_{i-2}})\mathrm{v}_{1,t_i})\right]+\mathbb{E}\left(\mathrm{v}_{1,t_i}\mathrm{v}_{1,t_{i-1}}\right)\\ &= \alpha^2_{\delta,1}\,\mathbb{E}\left[(P_{1,t_{i-1}}-P_{2,t_{i-1}})(P_{1,t_{i-2}}-P_{2,t_{i-2}})\right]+\alpha_{\delta,1}\bar{\Upsilon}_{\delta,0}\begin{pmatrix}\sigma^2_{\delta,1}\\ \sigma_{\delta,12}\end{pmatrix}\\ &\quad+\mathbb{E}\left(\Delta u_{1,t_i}\Delta u_{1,t_{i-1}}\right)-\alpha_{\delta,1}\,\mathbb{E}\left[(u_{1,t_{i-1}}-u_{2,t_{i-1}})\Delta u_{1,t_{i-1}}\right]\\ &= \alpha^2_{\delta,1}\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h+1}+\alpha_{\delta,1}\bar{\Upsilon}_{\delta,0}\begin{pmatrix}\sigma^2_{\delta,1}\\ \sigma_{\delta,12}\end{pmatrix}-(1+\alpha_{\delta,1})\omega_1^2\end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}(\Delta p_{1,t_i}^2) &= \mathbb{E}\left\{\left[\alpha_{\delta,1}(p_{1,t_{i-1}} - p_{2,t_{i-1}}) + \mathrm{v}_{1,t_i}\right]^2\right\} \\
&= \alpha_{\delta,1}^2\,\mathbb{E}\left[(p_{1,t_{i-1}} - p_{2,t_{i-1}})^2\right] + 2\,\alpha_{\delta,1}\,\mathbb{E}\left[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\mathrm{v}_{1,t_i}\right] + \mathbb{E}(\mathrm{v}_{1,t_i}^2) \\
&= \alpha_{\delta,1}^2\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2\right] - 2\,\alpha_{\delta,1}\left[(1+\alpha_{\delta,1})\omega_1^2 + \alpha_{\delta,1}\omega_2^2\right] \\
&\quad + \mathbb{E}(\Delta u_{1,t_i}^2) + \mathbb{E}(\varepsilon_{1,t_i}^2) + \alpha_{\delta,1}^2\,\mathbb{E}(u_{1,t_{i-1}} - u_{2,t_{i-1}})^2 - 2\,\alpha_{\delta,1}\,\mathbb{E}\left[\Delta u_{1,t_i}(u_{1,t_{i-1}} - u_{2,t_{i-1}})\right] \\
&= \alpha_{\delta,1}^2\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2\right] - 2\,\alpha_{\delta,1}\left[(1+\alpha_{\delta,1})\omega_1^2 + \alpha_{\delta,1}\omega_2^2\right] \\
&\quad + 2\,\omega_1^2 + \sigma_1^2 + \alpha_{\delta,1}^2\,(\omega_1^2 + \omega_2^2) + 2\,\alpha_{\delta,1}\,\omega_1^2 \\
&= \alpha_{\delta,1}^2\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \sigma_{\delta,1}^2 + 2\,\omega_1^2.
\end{aligned}
$$

Altogether, this yields

$$
\operatorname*{plim}_{n\to\infty}\hat{\alpha}_{\delta,1} = \alpha_{\delta,1} - \frac{Q_B}{\mathbb{E}(\Delta p_{1,t_i}^2)\mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})^2] - \{\mathbb{E}[(p_{1,t_{i-1}} - p_{2,t_{i-1}})\Delta p_{1,t_{i-1}}]\}^2},
$$

with

$$
\begin{aligned}
Q_B &= \left[(1+\alpha_{\delta,1})\omega_1^2 + \omega_2^2\right]\left[\alpha_{\delta,1}^2\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \sigma_{\delta,1}^2 + 2\,\omega_1^2\right] \\
&\quad + (1+2\,\alpha_{\delta,1})\omega_1^2\left[\alpha_{\delta,1}\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h+1} + \bar{\Upsilon}_{\delta,0}\begin{pmatrix}\sigma_{\delta,1}^2 \\ \sigma_{\delta,12}\end{pmatrix} + \omega_1^2\right].
\end{aligned}
$$

The expression for the bias is now much more intricated because it has to account for the measurement error in the additional regressor. Similarly, adding the other market's past price change or more lags would further increase the magnitude of the bias.

B.3 Bias in leading market

Suppose now that the first market leads the second market at all times with $\alpha_{\delta,1} = 0$. The bias becomes

$$
\operatorname*{plim}_{n\to\infty}\hat{\alpha}_{\delta,1} = Q^{-1}\left\{\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}[\rho_{\cdot,1}(h+1) - \rho_{\cdot,1}(h)] + \gamma_{1,1}(1) - \gamma_{1,2}(1) + \omega_{1,2} - \omega_1^2\right\},
$$

which reduces to $-\omega_1^2/Q$ if the market microstructure noises are uncorrelated white noises independent of the efficient prices. If we now add the past price changes to control the serial correlation in the error term, the LS estimator of $\alpha_{\delta,1}$ converges in probability to

$$
\operatorname*{plim}_{n\to\infty}\hat{\alpha}_{\delta,1} = -\frac{\left[\omega_1^2 + \omega_2^2\right]\left[\sigma_{\delta,1}^2 + 2\,\omega_1^2\right] + \omega_1^2\left[\bar{\Upsilon}_{\delta,0}(\sigma_{\delta,1}^2, \sigma_{\delta,12})' + \omega_1^2\right]}{(\sigma_{\delta,1}^2 + 2\,\omega_1^2)\left[\sum_{h=0}^{\infty}\bar{\Upsilon}_{\delta,h}\Sigma_\delta\bar{\Upsilon}'_{\delta,h} + \omega_1^2 + \omega_2^2\right] - \left[\bar{\Upsilon}_{\delta,0}(\sigma_{\delta,1}^2, \sigma_{\delta,12})' + \omega_1^2\right]^2}.
$$

As the least-square estimator overestimates the magnitude of $\alpha_{\delta,1}$, it will underestimate by how much the first market leads the price discovery.

References

Ait-Sahalia, Y., 2007, Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, Vol. 3 of Econometric Society Monographs, Cambridge University Press.

Ait-Sahalia, Y., Xiu, D., forthcoming, A hausman test for the presence of noise in high frequency data, Journal of Econometrics.

Amemiya, T., 1985, Advanced Econometrics, Harvard University Press.

Andrews, D. W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 817–858.

Baillie, R. T., Booth, G. G., Tse, Y., Zabotina, T., 2002, Price discovery and common factor models, Journal of Financial Markets 5, 309–321.

Bandi, F. M., Russell, J. R., 2005, Realized covariation, realized beta, and microstructure noise, Technical report, working paper, University of Chicago and John Hopkins University.

Bandi, F. M., Russell, J. R., 2008, Microstructure noise, realized variance, and optimal sampling, Review of Economic Studies 75, 339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008, Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, Econometrica 76, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2009, Realized kernels in practice: Trades and quotes, Econometrics Journal 12, C1–C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011, Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading, Journal of Econometrics 162, 149–169.

Booth, G. G., So, R. W., Tseh, Y., 1999, Price discovery in the German equity index derivatives markets, Journal of Futures Markets 19, 619–643.

Bound, J., Jaeger, D. A., Baker, R. M., 1995, Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak, Journal of the American Statistical Association 90, 443–450.

Chakravarty, S., Gulen, H., Mayhew, S., 2004, Informed trading in stock and option markets, Journal of Finance 59, 1235–1257.

Chu, Q. C., Hsieh, W. G., Tse, Y., 1999, Price discovery on the S&P 500 index markets: An analysis of spot index, index futures and SPDRs, International Review of Financial Analysis 8, 21–34.

Davidson, J., 1994, Stochastic Limit Theory. An Introduction for Econometricians, Oxford University Press.

Dias, G. F., Fernandes, M., Scherrer, C., 2019, Price discovery in a continuous-time setting, forthcoming at Journal of Financial Econometrics.

Eun, C. S., Sabherwal, S., 2003, Cross-border listings and price discovery: Evidence from US-listed Canadian stocks, Journal of Finance 58, 549–575.

Fernandes, M., Scherrer, C. M., 2018, Price discovery in dual-class shares across multiple markets, Journal of Futures Markets 38, 129–155.

Figuerola-Ferretti, I., Gonzalo, J., 2010, Modelling and measuring price discovery in commodity markets, Journal of Econometrics 158, 95–107.

Frijns, B., Gilbert, A., Tourani-Rad, A., 2015, The determinants of price discovery: Evidence from US-Canadian cross-listed shares, Journal of Banking and Finance pp. 457–468.

Gonzalo, J., Granger, C., 1995, Estimation of common long-memory components in cointegrated systems, Journal of Business and Statistics 13, 27–35.

Gonzalo, J., Ng, S., 2001, A systematic framework for analyzing the dynamic effects of permanent and transitory shocks, Journal of Economics Dynamics and Control 25, 1527–1546.

Grammig, J., Melvin, M., Schlag, C., 2005, Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects, Journal of Empirical Finance 12, 139–164.

Hamilton, J. D., 1994, Time Series Analysis, Princeton University Press.

Han, C., Phillips, P. C. B., 2006, Gmm with many moment conditions, Econometrica 74(1), 147–192.

Hansen, L. P., Heaton, J., Yaron, A., 1996, Finite-sample properties of some alternative gmm estimators, Journal of Business & Economic Statistics 14(3), 262–280.

Hansen, P. R., 2005, Granger's representation theorem: A closed-form expression for I(1) processes, Econometrics Journal 8, 23–38.

Hansen, P. R., Lunde, A., 2006, Realized variance and market microstructure noise, Journal of Business and Economic Statistics 24, 127–161.

Hansen, P. R., Lunde, A., 2014, Estimating the persistence and the autocorrelation function of a time series that is measured with error, Ecocnometric Theory 30, 6.

Harris, F., McInish, T. H., Wood, R. A., 2002, Security price adjustment across exchanges: An investigation of common factor components for Dow stocks, Journal of Financial Markets 5, 277–308.

Hasbrouck, J., 1995, One security, many markets: Determining the contributions to price discovery, Journal of Finance 50, 1175–1198.

Hasbrouck, J., 2003, Intraday price formation in U.S. equity index markets, Journal of Finance 58, 2375–2399.

Hasbrouck, J., 2018, Price discovery in high resolution, Halbert White Jr. Memorial JFEC invited lecture, forthcoming at Journal of Financial Econometrics.

Hausman, J., Lewis, R., Menzel, K., Newey, W., 2011, Properties of the cue estimator and a modification with moments, Journal of Econometrics 165(1), 45 – 57. Moment Restriction-Based Econometric Methods.

Jacod, J., n.d., Limit of random measures associated with the increments of a brownian semimartingale, Technical report.

Jacod, J., Li, Y., Mykland, P., Podolskij, M., Vetter, M., 2009, Microstructure noise in the continuous case: The pre-averaging approach, Stochastic processes and their applications 119, 2249–2276.

Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica 59, 1551–1580.

Johansen, S., 1995, Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press.

Kessler, M., Rahbek, A., 2001, Asymptotic likelihood based inference for co-integrated homogenous Gaussian diffusions, Scandinavian Journal of Statistics 28, 455–470.

Kessler, M., Rahbek, A., 2004, Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions, Statistical Inference for Stochastic Processes 7, 137–151.

Lien, D., Shrestha, K., 2009, A new information share measure, Journal of Futures Markets 4, 377–395.

Liu, L., Patton, A., Sheppard, K., 2015, Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, Journal of Econometrics 187, 293–311.

Menkveld, A. J., 2014, High-frequency traders and market structure, Financial Review 49, 333–344.

Menkveld, A. J., 2016, The economics of high-frequency trading: Taking stock, Annual Review of Financial Economics 8, 1–24.

Menkveld, A. J., Koopman, S. J., Lucas, A., 2007, Modeling around-the-clock price discovery for cross-listed stocks using state space methods, Journal of Business and Economic Statistics 25, 213–225.

Mizrach, B., Neely, C. J., 2008, Information shares in the US Treasury market, Journal of Banking and Finance 32, 1221–1233.

Mykland, P. A., Zhang, L., 2006, ANOVA for diffusions and Ito processes, Annals of Statistics 34, 1931–1963.

Newey, W. K., McFadden, D., 1994, Large sample estimation and hypothesis testing, in: R. F. Engle D. McFadden (eds), Handbook of Econometrics, Vol. 4 of Handbook of Econometrics, Elsevier, chapter 36, pp. 2111–2245.

Newey, W. K., Windmeijer, F., 2009, Generalized method of moments with many weak moment conditions, Econometrica 77(3), 687–719.

Nguenang, C., 2016, Evaluating markets contribution to price discovery in high-frequency for co-listed assets, working paper, Toulouse School of Economics.

O'Hara, M., 2015, High frequency market microstructure, Journal of Financial Economics 116, 257–270.

Ozturk, S., van der Wel, M., van Dijk, D., 2014, Intraday price discovery in fragmented markets, Technical report, working paper, Tinbergen Institute Discussion Paper.

Phillips, P. C. B., Yu, J., 2008, Information loss in volatility measurement with flat price trading, Technical report, working paper, Cowles Foundation for Research in Economics, Yale University.

Scherrer, C., 2013, Price discovery and instantaneous effects among cross listed stocks, Technical report, working paper, Aarhus University, Department of Economics and Business, and CREATES.

Staiger, D., Stock, J. H., 1997, Instrumental variables regression with weak instruments, Econometrica 65(3), 557–586.

Stock, J. H., Watson, M. W., 1988, Testing for common trends, Journal of the American Statistical Association 83, 1097–1107.

Stock, J. H., Wright, J. H., 2000, Gmm with weak identification, Econometrica 68(5), 1055–1096.

Stock, J., Yogo, M., 2002, A survey of weak instruments and weak identification in generalized method of moments, Journal of Business and Economic Statistics 20, 518–529.

Stock, J., Yogo, M., 2005a, Asymptotic distributions of instrumental variables statistics with many instruments, Identification and Inference for Econometric Models, Cambridge University Press, chapter 6, pp. 109–120.

Stock, J., Yogo, M., 2005b, Testing for weak instruments in linear iv regression, Identification and Inference for Econometric Models, Cambridge University Press, chapter 5.

White, H., 2000, Asymptotic Thoery for Econometricians: Revised Edition, Academic Press.

Yan, B., Zivot, E., 2010, A structural analysis of price discovery measures, Journal of Financial Markets 13, 1–19.

Zhang, L., 2011, Estimating covariation: Epps effect and microstructure noise, Journal of Econometrics 160, 33–47.

Table 1: Performance of the LS and 2SLS estimators of $\alpha_{\perp,1}$ under Assumption MMN(TS): iid market microstructure noise

The upper panel reports the median bias of the estimates of $\alpha_{\perp,1}$ and the lower panel displays the ratio of their root median squared errors (RRMSE) with respect to the estimates obtained using LS at the tick-by-tick frequency. We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are serially uncorrelated white noises with variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. We estimate $\alpha_\perp$ by LS and 2SLS. Specifically, VEC(0), VEC(0)$_1$, VEC(0)$_5$ denote estimates of $\alpha_\perp$ that are estimated by LS from a VEC(0) model fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; VEC(1), VEC(1)$_1$, VEC(1)$_5$ denote estimates of $\alpha_\perp$ that are estimated by LS from VEC(1) model fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; 2SLS and 2SLS$_1$ are the estimates of $\alpha_\perp$ which are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ for $2 \leq \kappa \leq 6$ and $3 \leq \kappa \leq 7$, respectively; and R^2_{XZ} ($R^2_{XZ,1}$) and F_{XZ} ($F_{XZ,1}$) are the coefficient of determination and F statistics of the first-stage equation of the 2SLS (2SLS$_1$) estimators.

Bias

$\alpha_\perp = (1.00, 0.00)'$

	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS	2SLS$_1$	F_{XZ}	$F_{XZ,1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.43	-0.35	-0.18	-0.16	-0.15	-0.19	0.00	0.00	2655	1118
$\omega^2 = (0.0005, 0.0005)'$	-0.48	-0.46	-0.36	-0.35	-0.37	-0.33	-0.01	-0.02	494	242
$\omega^2 = (0.0010, 0.0010)'$	-0.49	-0.48	-0.42	-0.40	-0.43	-0.45	-0.02	-0.04	195	102
$\omega^2 = (0.0010, 0.0005)'$	-0.65	-0.63	-0.53	-0.51	-0.52	-0.54	-0.03	-0.04	290	148

$\alpha_\perp = (0.67, 0.33)'$

	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS	2SLS$_1$	F_{XZ}	$F_{XZ,1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.14	-0.12	-0.07	-0.07	-0.06	-0.06	0.00	0.00	1499	596
$\omega^2 = (0.0005, 0.0005)'$	-0.16	-0.15	-0.13	-0.13	-0.14	-0.10	-0.01	-0.01	240	109
$\omega^2 = (0.0010, 0.0010)'$	-0.16	-0.16	-0.15	-0.14	-0.16	-0.18	-0.01	-0.02	87	42
$\omega^2 = (0.0010, 0.0005)'$	-0.33	-0.32	-0.28	-0.27	-0.27	-0.29	-0.02	-0.04	133	63

Relative root mean squared error (RRMSE)

$\alpha_\perp = (1.00, 0.00)'$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS	2SLS$_1$	R^2_{XZ}	$R^2_{XZ,1}$
$\omega^2 = (0.0001, 0.0001)'$	0.80	0.41	0.41	1.05	1.50	0.09	0.10	0.47	0.40
$\omega^2 = (0.0005, 0.0005)'$	0.94	0.74	0.71	0.83	1.01	0.14	0.14	0.14	0.13
$\omega^2 = (0.0010, 0.0010)'$	0.97	0.85	0.81	0.87	0.99	0.21	0.19	0.06	0.06
$\omega^2 = (0.0010, 0.0005)'$	0.97	0.81	0.78	0.81	0.90	0.15	0.14	0.09	0.08

$\alpha_\perp = (0.67, 0.33)'$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS	2SLS$_1$	R^2_{XZ}	$R^2_{XZ,1}$
$\omega^2 = (0.0001, 0.0001)'$	0.80	0.71	0.92	3.19	4.44	0.23	0.23	0.33	0.26
$\omega^2 = (0.0005, 0.0005)'$	0.95	0.82	0.82	1.68	2.34	0.38	0.35	0.07	0.06
$\omega^2 = (0.0010, 0.0010)'$	0.97	0.92	0.88	1.31	1.94	0.54	0.53	0.03	0.02
$\omega^2 = (0.0010, 0.0005)'$	0.97	0.86	0.83	0.98	1.19	0.25	0.24	0.04	0.04

Table 2: Performance of the LS and 2SLS estimators of $\alpha_{\perp,1}$ under Assumption MMN(TS): endogenous MA(1) market microstructure noise

The upper panel reports the median bias of the estimates of $\alpha_{\perp,1}$ and the lower panel displays the ratio of their root median squared errors (RRMSE) with respect to the estimates obtained using LS at the tick-by-tick frequency. We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are endogenous MA(1) processes with $\mathrm{Corr}\,(u_{m,t_i}, \varepsilon_{m',t_i}) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. We estimate $\alpha_\perp$ by LS and 2SLS. Specifically, VEC(0), VEC(0)$_1$, VEC(0)$_5$ denote estimates of $\alpha_\perp$ that are estimated by LS from a VEC(0) model fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; VEC(1), VEC(1)$_1$, VEC(1)$_5$ denote estimates of $\alpha_\perp$ that are estimated by LS from VEC(1) model fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; 2SLS$_1$ and 2SLS$_2$ are the estimates of $\alpha_\perp$ which are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ for $3 \leq \kappa \leq 7$ and $4 \leq \kappa \leq 8$, respectively; and $R^2_{XZ,1}$ ($R^2_{XZ,2}$) and $F_{XZ,1}$ ($F_{XZ,2}$) are the coefficient of determination and F statistics of the first-stage equation of the 2SLS$_1$ (2SLS$_2$) estimators.

Bias

$\alpha_\perp = (1.00, 0.00)'$

	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_1$	2SLS$_2$	$F_{XZ,1}$	$F_{XZ,2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.36	-0.28	-0.11	-0.10	-0.10	-0.18	0.00	0.00	3770	1410
$\omega^2 = (0.0005, 0.0005)'$	-0.46	-0.43	-0.28	-0.24	-0.29	-0.29	0.00	0.00	945	412
$\omega^2 = (0.0010, 0.0010)'$	-0.48	-0.46	-0.36	-0.32	-0.36	-0.34	-0.01	-0.01	402	191
$\omega^2 = (0.0010, 0.0005)'$	-0.76	-0.74	-0.54	-0.51	-0.52	-0.54	0.00	-0.02	548	251

$\alpha_\perp = (0.67, 0.33)'$

	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_1$	2SLS$_2$	$F_{XZ,1}$	$F_{XZ,2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.12	-0.09	-0.05	-0.04	-0.04	-0.09	0.00	0.00	2089	740
$\omega^2 = (0.0005, 0.0005)'$	-0.16	-0.15	-0.11	-0.10	-0.12	-0.06	0.00	0.00	452	184
$\omega^2 = (0.0010, 0.0010)'$	-0.16	-0.16	-0.13	-0.12	-0.14	-0.12	0.00	0.00	176	77
$\omega^2 = (0.0010, 0.0005)'$	-0.44	-0.44	-0.35	-0.35	-0.35	-0.36	0.00	-0.02	247	105

Relative root mean squared error (RRMSE)

$\alpha_\perp = (1.00, 0.00)'$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_1$	2SLS$_2$	$R^2_{XZ,1}$	$R^2_{XZ,2}$
$\omega^2 = (0.0001, 0.0001)'$	0.77	0.35	0.42	1.27	1.91	0.10	0.12	0.50	0.43
$\omega^2 = (0.0005, 0.0005)'$	0.93	0.61	0.53	0.85	1.12	0.14	0.14	0.20	0.18
$\omega^2 = (0.0010, 0.0010)'$	0.96	0.75	0.66	0.81	0.97	0.22	0.20	0.10	0.09
$\omega^2 = (0.0010, 0.0005)'$	0.97	0.71	0.67	0.71	0.81	0.13	0.11	0.13	0.12

$\alpha_\perp = (0.67, 0.33)'$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_1$	2SLS$_2$	$R^2_{XZ,1}$	$R^2_{XZ,2}$
$\omega^2 = (0.0001, 0.0001)'$	0.78	0.80	1.10	3.76	5.66	0.26	0.31	0.36	0.28
$\omega^2 = (0.0005, 0.0005)'$	0.93	0.76	0.79	2.08	3.06	0.42	0.43	0.11	0.09
$\omega^2 = (0.0010, 0.0010)'$	0.96	0.83	0.80	1.65	2.43	0.65	0.62	0.04	0.04
$\omega^2 = (0.0010, 0.0005)'$	0.99	0.79	0.80	0.88	1.09	0.20	0.18	0.06	0.05

Table 3: Performance of the LS-RV, 2SLS-RK, and 2SLS-HAC estimators of the continuous-time Information Share measure under Assumption MMN(TS): iid market microstructure noise

The upper panel reports the median bias of the estimates of the continuous-time Information Share (IS) measure of the first market, IS_1 and the lower panel displays the ratio of their root median squared errors (RRMSE) with respect to the estimates obtained using LS at the tick-by-tick frequency. We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters and sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are serially uncorrelated white noises with variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. VEC(0), VEC$(0)_1$, VEC$(0)_5$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by LS from VEC(0) models fitted at tick-data, 1-minute, and 5-minute sampling intervals, respectively, and integrated variance estimated with the realized variance (RV) estimator at tick-data, 1-minute, and 5-minute sampling intervals, respectively; VEC(1), VEC$(1)_1$, VEC$(1)_5$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by LS from VEC(1) models fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively, and continuous-time covariance matrix estimated with the realized variance (RV) estimator at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; 2SLS_{RK} and $2\text{SLS}_{\text{RK},1}$ denote estimates of IS_1 that are constructed with estimates of $\alpha_\perp$ computed by 2SLS from VEC(0) models fitted at the tick-by-tick frequency and integrated variance estimated with the realized kernel (RK) estimator at the tick-by-tick. The $\alpha_\perp$ estimates in 2SLS_{RK} and $2\text{SLS}_{\text{RK},1}$ use instruments selected as $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ for $2 \leq \kappa \leq 6$ and $3 \leq \kappa \leq 7$, respectively; and 2SLS_{HAC} and $2\text{SLS}_{\text{HAC},1}$ denote estimates of the IS measures based on the 2SLS estimator with instruments defined in the same fashion as in the 2SLS_{RK} and $2\text{SLS}_{\text{RK},1}$ measures, and the integrated variance based on the exact discretization of the discrete-time unconditional variance computed with the HAC estimator.

Bias

$IS_1 = 0.88$

	VEC(0)	VEC(1)	VEC$(0)_1$	VEC$(1)_1$	VEC$(0)_5$	VEC$(1)_5$	2SLS_{RK}	$2\text{SLS}_{\text{RK},1}$	2SLS_{HAC}	$2\text{SLS}_{\text{HAC},1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.26	-0.12	-0.35	-0.25	-0.37	-0.35	-0.17	-0.17	0.01	0.01
$\omega^2 = (0.0005, 0.0005)'$	-0.35	-0.29	-0.36	-0.27	-0.37	-0.34	-0.22	-0.23	0.03	0.03
$\omega^2 = (0.0010, 0.0010)'$	-0.36	-0.33	-0.37	-0.28	-0.37	-0.36	-0.24	-0.24	0.05	0.05
$\omega^2 = (0.0010, 0.0005)'$	-0.52	-0.47	-0.48	-0.36	-0.41	-0.38	-0.23	-0.23	0.04	0.04

$IS_1 = 0.66$

	VEC(0)	VEC(1)	VEC$(0)_1$	VEC$(1)_1$	VEC$(0)_5$	VEC$(1)_5$	2SLS_{RK}	$2\text{SLS}_{\text{RK},1}$	2SLS_{HAC}	$2\text{SLS}_{\text{HAC},1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.12	-0.07	-0.15	-0.13	-0.16	-0.15	-0.10	-0.10	0.01	0.00
$\omega^2 = (0.0005, 0.0005)'$	-0.15	-0.13	-0.16	-0.13	-0.16	-0.15	-0.12	-0.12	0.03	0.02
$\omega^2 = (0.0010, 0.0010)'$	-0.16	-0.15	-0.16	-0.14	-0.16	-0.17	-0.12	-0.12	0.04	0.04
$\omega^2 = (0.0010, 0.0005)'$	-0.32	-0.30	-0.27	-0.22	-0.20	-0.19	-0.12	-0.13	0.04	0.03

Relative root mean squared error (RRMSE)

$IS_1 = 0.88$

	VEC(1)	VEC$(0)_1$	VEC$(1)_1$	VEC$(0)_5$	VEC$(1)_5$	2SLS_{RK}	$2\text{SLS}_{\text{RK},1}$	2SLS_{HAC}	$2\text{SLS}_{\text{HAC},1}$
$\omega^2 = (0.0001, 0.0001)'$	0.47	1.35	0.98	1.43	1.36	0.67	0.67	0.06	0.06
$\omega^2 = (0.0005, 0.0005)'$	0.85	1.05	0.77	1.08	1.00	0.64	0.65	0.10	0.10
$\omega^2 = (0.0010, 0.0010)'$	0.92	1.02	0.78	1.04	1.01	0.66	0.67	0.16	0.15
$\omega^2 = (0.0010, 0.0005)'$	0.91	0.92	0.69	0.79	0.74	0.45	0.45	0.09	0.08

$IS_1 = 0.66$

	VEC(1)	VEC$(0)_1$	VEC$(1)_1$	VEC$(0)_5$	VEC$(1)_5$	2SLS_{RK}	$2\text{SLS}_{\text{RK},1}$	2SLS_{HAC}	$2\text{SLS}_{\text{HAC},1}$
$\omega^2 = (0.0001, 0.0001)'$	0.60	1.25	1.03	1.30	1.25	0.83	0.83	0.22	0.25
$\omega^2 = (0.0005, 0.0005)'$	0.89	1.04	0.88	1.06	0.98	0.79	0.80	0.36	0.37
$\omega^2 = (0.0010, 0.0010)'$	0.95	1.01	0.88	1.03	1.07	0.78	0.80	0.64	0.61
$\omega^2 = (0.0010, 0.0005)'$	0.94	0.86	0.71	0.63	0.60	0.38	0.40	0.24	0.23

Table 4: Performance of the LS-RV, 2SLS-RK, and 2SLS-HAC estimators of the continuous-time Information Share measure under Assumption MMN(TS): endogenous MA(1) market microstructure noise

The upper panel reports the median bias of the estimates of the continuous-time Information Share (IS) measure of the first market, IS_1 and the lower panel displays the ratio of their root median squared errors (RRMSE) with respect to the estimates obtained using LS at the tick-by-tick frequency. We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are endogenous MA(1) processes with Corr $(u_{m,t_i}, \varepsilon_{m',t_i}) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. VEC(0), VEC(0)$_1$, VEC(0)$_5$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by LS from VEC(0) models fitted at tick-data, 1-minute, and 5-minute sampling intervals, respectively, and integrated variance estimated with the realized variance (RV) estimator at tick-data, 1-minute, and 5-minute sampling intervals, respectively; VEC(1), VEC(1)$_1$, VEC(1)$_5$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by LS from VEC(1) models fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively, and continuous-time covariance matrix estimated with the realized variance (RV) estimator at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively; 2SLS$_{\text{RK}}$ and 2SLS$_{\text{RK},1}$ denote estimates of IS_1 that are constructed with estimates of $\alpha_\perp$ computed by 2SLS from VEC(0) models fitted at the tick-by-tick frequency and integrated variance estimated with the realized kernel (RK) estimator at the tick-by-tick. The $\alpha_\perp$ estimates in 2SLS$_{\text{RK}}$ and 2SLS$_{\text{RK},1}$ use instruments selected as $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ for $2 \leq \kappa \leq 6$ and $3 \leq \kappa \leq 7$, respectively; and 2SLS$_{\text{HAC}}$ and 2SLS$_{\text{HAC},1}$ denote estimates of the IS measures based on the 2SLS estimator with instruments defined in the same fashion as in the 2SLS$_{\text{RK}}$ and 2SLS$_{\text{RK},1}$ measures, and the integrated variance based on the exact discretization of the discrete-time unconditional variance computed with the HAC estimator.

Bias

$IS_1 = 0.88$

	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_{\text{RK},1}$	2SLS$_{\text{RK},2}$	2SLS$_{\text{HAC},1}$	2SLS$_{\text{HAC},2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.21	-0.13	-0.33	-0.25	-0.37	-0.35	-0.16	-0.16	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	-0.33	-0.30	-0.36	-0.26	-0.37	-0.35	-0.20	-0.20	0.01	0.01
$\omega^2 = (0.0010, 0.0010)'$	-0.35	-0.33	-0.36	-0.27	-0.37	-0.35	-0.22	-0.22	0.02	0.02
$\omega^2 = (0.0010, 0.0005)'$	-0.58	-0.56	-0.50	-0.37	-0.41	-0.38	-0.21	-0.21	0.01	0.01

$IS_1 = 0.66$

.	VEC(0)	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_{\text{RK},1}$	2SLS$_{\text{RK},2}$	2SLS$_{\text{HAC},1}$	2SLS$_{\text{HAC},2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.11	-0.08	-0.15	-0.12	-0.16	-0.16	-0.10	-0.10	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	-0.15	-0.14	-0.16	-0.13	-0.16	-0.15	-0.11	-0.11	0.01	0.01
$\omega^2 = (0.0010, 0.0010)'$	-0.15	-0.15	-0.16	-0.13	-0.16	-0.15	-0.11	-0.11	0.02	0.02
$\omega^2 = (0.0010, 0.0005)'$	-0.39	-0.38	-0.30	-0.25	-0.20	-0.19	-0.11	-0.12	0.02	0.01

Relative root mean squared error (RRMSE)

$IS_1 = 0.88$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_{\text{RK},1}$	2SLS$_{\text{RK},2}$	2SLS$_{\text{HAC},1}$	2SLS$_{\text{HAC},2}$
$\omega^2 = (0.0001, 0.0001)'$	0.62	1.54	1.18	1.71	1.65	0.74	0.74	0.07	0.07
$\omega^2 = (0.0005, 0.0005)'$	0.89	1.07	0.78	1.11	1.04	0.61	0.61	0.07	0.07
$\omega^2 = (0.0010, 0.0010)'$	0.94	1.03	0.76	1.05	0.98	0.62	0.62	0.10	0.10
$\omega^2 = (0.0010, 0.0005)'$	0.96	0.86	0.63	0.71	0.65	0.36	0.37	0.05	0.04

$IS_1 = 0.66$

	VEC(1)	VEC(0)$_1$	VEC(1)$_1$	VEC(0)$_5$	VEC(1)$_5$	2SLS$_{\text{RK},1}$	2SLS$_{\text{RK},2}$	2SLS$_{\text{HAC},1}$	2SLS$_{\text{HAC},2}$
$\omega^2 = (0.0001, 0.0001)'$	0.71	1.39	1.17	1.48	1.46	0.89	0.89	0.25	0.30
$\omega^2 = (0.0005, 0.0005)'$	0.92	1.05	0.87	1.08	1.01	0.76	0.76	0.39	0.39
$\omega^2 = (0.0010, 0.0010)'$	0.96	1.02	0.87	1.04	0.99	0.74	0.75	0.61	0.56
$\omega^2 = (0.0010, 0.0005)'$	0.99	0.76	0.64	0.52	0.49	0.29	0.30	0.20	0.17

Table 5: Performance (RRMSE and bias) of the LS and 2SLS-HAC estimators of the discrete-time Information Share measure under Assumption MMN(TS): iid and endogenous MA(1) market microstructure noise

The first and second panels report the median bias and relative root median squared error (RRMSE) of the estimates of the discrete-time Information Share (IS) measures (tick-by-tick frequency) for the case of iid noise, whereas the last two panels cover the case of endogenous MA(1) market microstructure noise. We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. The market microstructure noises are serially uncorrelated white noises and endogenous MA(1) processes with $\text{Corr}\,(u_{m,t_i}, \varepsilon_{m',t_i}) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. VEC(0) and VEC(1) denote estimates of the discrete-time IS measures that are constructed with estimates of $\alpha_\perp$ computed by least squares from VEC(0) and VEC(1) models fitted at the tick-by-tick, respectively, and estimates of the unconditional variance that are not robust to heteroskedaticity and serial correlation; $2SLS_{HAC}$, $2SLS_{HAC,1}$, and $2SLS_{HAC,2}$ denote estimates of the discrete-time IS measures that are constructed with estimates of $\alpha_\perp$ computed by 2SLS from VEC(0) models fitted at the tick-by-tick frequency and estimates of the unconditional variance computed with the HAC estimator. Specifically, $2SLS_{HAC}$, $2SLS_{HAC,1}$, and $2SLS_{HAC,2}$ use instruments selected as $(p_{1,t_{i-\kappa}} - p_{2,t_{i-\kappa}})'$ for $2 \leq \kappa \leq 6$, $3 \leq \kappa \leq 7$, and $4 \leq \kappa \leq 8$ respectively. The RRMSE measure is the ratio of the root median squared errors (RMSE) of a given estimator relative to the RMSE of the LS estimator.

iid noise, $IS_{\delta,1} = 0.88$

	Bias				RRMSE		
	VEC(0)	VEC(1)	$2SLS_{HAC}$	$2SLS_{HAC,1}$	VEC(1)	$2SLS_{HAC}$	$2SLS_{HAC,1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.25	0.24	0.00	0.00	0.50	0.06	0.06
$\omega^2 = (0.0005, 0.0005)'$	-0.33	0.08	0.03	0.03	0.85	0.10	0.10
$\omega^2 = (0.0010, 0.0010)'$	-0.35	0.04	0.05	0.05	0.92	0.17	0.16
$\omega^2 = (0.0010, 0.0005)'$	-0.50	-0.10	0.04	0.04	0.93	0.09	0.09

iid noise, $IS_{\delta,1} = 0.66$

	Bias				RRMSE		
	VEC(0)	VEC(1)	$2SLS_{HAC}$	$2SLS_{HAC,1}$	VEC(1)	$2SLS_{HAC}$	$2SLS_{HAC,1}$
$\omega^2 = (0.0001, 0.0001)'$	-0.11	0.08	0.01	0.00	0.61	0.23	0.25
$\omega^2 = (0.0005, 0.0005)'$	-0.14	0.02	0.03	0.02	0.89	0.36	0.37
$\omega^2 = (0.0010, 0.0010)'$	-0.15	0.01	0.04	0.04	0.94	0.65	0.62
$\omega^2 = (0.0010, 0.0005)'$	-0.29	-0.14	0.04	0.03	0.97	0.25	0.24

endogenous MA(1) noise, $IS = 0.88$

	Bias				RRMSE		
	VEC(0)	VEC(1)	$2SLS_{HAC,1}$	$2SLS_{HAC,2}$	VEC(1)	$2SLS_{HAC,1}$	$2SLS_{HAC,2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.21	0.23	0.00	0.00	0.63	0.07	0.07
$\omega^2 = (0.0005, 0.0005)'$	-0.33	0.07	0.01	0.01	0.89	0.08	0.08
$\omega^2 = (0.0010, 0.0010)'$	-0.34	0.04	0.01	0.02	0.95	0.10	0.10
$\omega^2 = (0.0010, 0.0005)'$	-0.54	-0.17	0.01	0.01	0.98	0.05	0.05

endogenous MA(1) noise, $IS = 0.66$

	Bias				RRMSE		
	VEC(0)	VEC(1)	$2SLS_{HAC,1}$	$2SLS_{HAC,2}$	VEC(1)	$2SLS_{HAC,1}$	$2SLS_{HAC,2}$
$\omega^2 = (0.0001, 0.0001)'$	-0.10	0.08	0.00	0.00	0.72	0.25	0.30
$\omega^2 = (0.0005, 0.0005)'$	-0.14	0.02	0.01	0.01	0.92	0.38	0.39
$\omega^2 = (0.0010, 0.0010)'$	-0.15	0.01	0.02	0.02	0.96	0.62	0.56
$\omega^2 = (0.0010, 0.0005)'$	-0.34	-0.20	0.02	0.01	1.01	0.21	0.18

Table 6: Performance of the LS, 2SLS, and CU-GMM estimators of $\alpha_{\perp,1}$ under Assumption MMN(CS): endogenous MA(∞) market microstructure noise

We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1,-1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1,-1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are endogenous MA(∞) processes with Corr $\left(u_{m,t_i}, \varepsilon_{m',t_i}\right) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50^q for $q = 0, 1, 2, \ldots$, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. We estimate $\alpha_\perp$ by LS, 2SLS, and CU-GMM estimators using other assets traded at different exchanges as instruments. Specifically, the cross-sectional dimension of instruments, V, ranges from $V \in (1, 5, 10, 20)$ for the 2SLS estimator and $V \in (10, 20)$ for the CU-GMM estimator, so that $\text{2SLS}_{V=1}$, $\text{2SLS}_{V=5}$, $\text{2SLS}_{V=10}$, and $\text{2SLS}_{V=20}$ denote the estimates obtained when using 1, 5, 10 and 20 assets as instruments, respectively, and $\text{CU-GMM}_{V=10}$ and $\text{CU-GMM}_{V=20}$ denote the CU-GMM estimator obtained with 10 and 20 assets as instruments. For each asset, we consider five lags as instruments. VEC(0), VEC(0)_1 and VEC(0)_5 denote estimates of $\alpha_\perp$ that are estimated by LS from VEC(0) models fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively. We report the median bias and the ratio of their root median square errors (RRMSE) with respect to the estimates obtained using LS at the tick-by-tick frequency, as well as the coefficient of determination R^2_{XZ} of the first-stage equation as a measure of the relevance of the instruments.

$\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\alpha_\perp = (1.00, 0.00)'$

Bias	VEC(0)	VEC(0)_1	VEC(0)_5	$\text{2SLS}_{V=1}$	$\text{2SLS}_{V=5}$	$\text{2SLS}_{V=10}$	$\text{2SLS}_{V=20}$	$\text{CU-GMM}_{V=10}$	$\text{CU-GMM}_{V=20}$	F-stat	$\text{F-stat}_{V=5}$	$\text{F-stat}_{V=10}$	$\text{F-stat}_{V=20}$
$\omega^2 = (0.0001, 0.0001)'$	-0.40	-0.16	-0.12	0.03	0.01	0.00	-0.01	0.01	0.00	56.34	153.65	98.55	81.39
$\omega^2 = (0.0005, 0.0005)'$	-0.48	-0.38	-0.35	-0.13	-0.07	-0.11	-0.16	0.01	0.01	12.45	32.50	20.04	14.51
$\omega^2 = (0.0010, 0.0010)'$	-0.49	-0.44	-0.43	-0.31	-0.17	-0.24	-0.29	0.03	0.02	4.79	13.46	8.39	6.06
$\omega^2 = (0.0010, 0.0005)'$	-0.66	-0.56	-0.54	-0.27	-0.14	-0.22	-0.28	0.02	0.02	9.14	23.57	14.62	10.58

$\alpha_{\delta=1/23400} = (-0.025, 0.050)'$, $\alpha_\perp = (0.67, 0.33)'$

Bias	VEC(0)	VEC(0)_1	VEC(0)_5	$\text{2SLS}_{V=1}$	$\text{2SLS}_{V=5}$	$\text{2SLS}_{V=10}$	$\text{2SLS}_{V=20}$	$\text{CU-GMM}_{V=10}$	$\text{CU-GMM}_{V=20}$	F-stat	$\text{F-stat}_{V=5}$	$\text{F-stat}_{V=10}$	$\text{F-stat}_{V=20}$
$\omega^2 = (0.0001, 0.0001)'$	-0.13	-0.08	-0.05	0.01	0.01	0.00	0.00	0.00	0.00	47.23	114.83	71.84	58.19
$\omega^2 = (0.0005, 0.0005)'$	-0.16	-0.14	-0.12	-0.05	-0.02	-0.04	-0.06	0.01	0.00	8.67	19.70	12.28	9.04
$\omega^2 = (0.0010, 0.0010)'$	-0.16	-0.15	-0.15	-0.10	-0.06	-0.09	-0.11	0.01	0.00	3.32	7.79	5.11	3.86
$\omega^2 = (0.0010, 0.0005)'$	-0.33	-0.30	-0.29	-0.14	-0.09	-0.13	-0.16	0.01	0.01	6.25	13.99	8.84	6.59

$\alpha_{\delta=1/23400} = (-0.005, 0.010)'$, $\alpha_\perp = (0.67, 0.33)'$

RRMSE	VEC(0)_1	VEC(0)_5	$\text{2SLS}_{V=1}$	$\text{2SLS}_{V=5}$	$\text{2SLS}_{V=10}$	$\text{2SLS}_{V=20}$	$\text{CU-GMM}_{V=10}$	$\text{CU-GMM}_{V=20}$	R^2	$\text{R}^2_{V=5}$	$\text{R}^2_{V=10}$	$\text{R}^2_{V=20}$
$\omega^2 = (0.0001, 0.0001)'$	0.41	0.97	0.53	0.16	0.14	0.11	0.15	0.11	0.02	0.21	0.25	0.35
$\omega^2 = (0.0005, 0.0005)'$	0.78	0.79	0.72	0.23	0.25	0.33	0.21	0.19	0.00	0.05	0.06	0.09
$\omega^2 = (0.0010, 0.0010)'$	0.89	0.87	0.89	0.37	0.49	0.60	0.30	0.28	0.00	0.02	0.03	0.04
$\omega^2 = (0.0010, 0.0005)'$	0.85	0.83	0.65	0.25	0.34	0.43	0.18	0.16	0.00	0.04	0.05	0.07

$\alpha_{\delta=1/23400} = (-0.025, 0.050)'$, $\alpha_\perp = (0.67, 0.33)'$

RRMSE	VEC(0)_1	VEC(0)_5	$\text{2SLS}_{V=1}$	$\text{2SLS}_{V=5}$	$\text{2SLS}_{V=10}$	$\text{2SLS}_{V=20}$	$\text{CU-GMM}_{V=10}$	$\text{CU-GMM}_{V=20}$	R^2	$\text{R}^2_{V=5}$	$\text{R}^2_{V=10}$	$\text{R}^2_{V=20}$
$\omega^2 = (0.0001, 0.0001)'$	0.74	3.08	1.17	0.35	0.32	0.24	0.33	0.25	0.02	0.16	0.19	0.28
$\omega^2 = (0.0005, 0.0005)'$	0.86	1.59	1.49	0.49	0.43	0.41	0.47	0.42	0.00	0.03	0.04	0.06
$\omega^2 = (0.0010, 0.0010)'$	0.94	1.27	1.75	0.63	0.62	0.66	0.68	0.68	0.00	0.01	0.02	0.02
$\omega^2 = (0.0010, 0.0005)'$	0.91	0.93	0.86	0.32	0.40	0.49	0.25	0.23	0.00	0.02	0.03	0.04

Table 7: Performance of the LS, 2SLS, and CU-GMM estimators of the continuous-time Information Share measure under Assumption MMN(CS): endogenous MA(∞) market microstructure noise

We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are endogenous MA(∞) processes with $\mathrm{Corr}\left(u_{m,t_i}, \varepsilon_{m',t_i}\right) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50^q for $q = 0, 1, 2, \ldots$, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. We estimate continuous-time Information Share (IS) measures based on the LS, 2SLS, and CU-GMM estimators using other assets traded at different exchanges as instruments. Specifically, the cross-sectional dimension of instruments, V, ranges from $V \in (1, 5, 10, 20)$ for the 2SLS estimator and $V \in (10, 20)$ for the CU-GMM estimator. $2\mathrm{SLS}_{V=1,\mathrm{RK}}$, $2\mathrm{SLS}_{V=5,\mathrm{RK}}$, $2\mathrm{SLS}_{V=10,\mathrm{RK}}$, and $2\mathrm{SLS}_{V=20,\mathrm{RK}}$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by 2SLS with $V \in (1, 5, 10, 20)$ from VEC(0) models fitted at the tick-by-tick frequency and integrated variance estimated with the realized kernel (RK) estimator at the tick-by-tick, while $\text{CU-GMM}_{V=10,\mathrm{RK}}$ and $\text{CU-GMM}_{V=20,\mathrm{RK}}$ denote estimates of the IS measures computed from tick-by-tick by the CU-GMM estimator with $V = 10$ and $V = 20$, respectively, and the RK; the $2\mathrm{SLS}_{V=1,\mathrm{HAC}}$, $2\mathrm{SLS}_{V=5,\mathrm{HAC}}$, $2\mathrm{SLS}_{V=10,\mathrm{HAC}}$, and $2\mathrm{SLS}_{V=20,\mathrm{HAC}}$ denote the of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by 2SLS with $V \in (1, 5, 10, 20)$ from VEC(0) models fitted at the tick-by-tick frequency and estimates of the integrated variance obtained from the exact discretization applied to the discrete-time unconditional variance computed with the HAC estimator. VEC(0), $\mathrm{VEC}(0)_1$ and $\mathrm{VEC}(0)_5$ denote estimates of IS measures that combines estimates of $\alpha_\perp$ and the integrated variance computed by LS and the RV estimator, respectively, from VEC(0) models fitted at the tick-by-tick, 1-minute, and 5-minute sampling intervals, respectively. We report the median bias of the estimates of the continuous-time Information Share (IS) measures and the ratio of their root median square errors (RRMSE) with respect to the estimates obtained using LS and RV at the tick-by-tick frequency, as well as the coefficient of determination R^2_{XZ} of the first-stage equation as a measure of the relevance of the instruments.

$\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\alpha_\perp = (1.00, 0.00)'$

Bias	VEC(0)	$\mathrm{VEC}(0)_1$	$\mathrm{VEC}(0)_5$	$2\mathrm{SLS}_{V=1,\mathrm{RK}}$	$2\mathrm{SLS}_{V=5,\mathrm{RK}}$	$2\mathrm{SLS}_{V=10,\mathrm{RK}}$	$2\mathrm{SLS}_{V=20,\mathrm{RK}}$	$\text{CU-GMM}_{V=10,\mathrm{RK}}$	$\text{CU-GMM}_{V=20,\mathrm{RK}}$	$2\mathrm{SLS}_{V=1,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=5,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=10,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=20,\mathrm{HAC}}$	$\text{CU-GMM}_{V=10,\mathrm{HAC}}$	$\text{CU-GMM}_{V=20,\mathrm{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$	-0.21	-0.25	-0.35	-0.13	-0.14	-0.14	-0.15	-0.14	-0.14	0.00	0.00	0.00	0.00	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	-0.34	-0.29	-0.35	-0.21	-0.19	-0.20	-0.22	-0.16	-0.16	-0.03	0.03	0.02	0.01	0.04	0.04
$\omega^2 = (0.0010, 0.0010)'$	-0.36	-0.31	-0.35	-0.28	-0.22	-0.25	-0.27	-0.15	-0.15	-0.14	0.02	-0.01	-0.04	0.06	0.06
$\omega^2 = (0.0010, 0.0005)'$	-0.52	-0.40	-0.38	-0.25	-0.19	-0.23	-0.25	-0.13	-0.14	-0.07	0.02	-0.01	-0.04	0.05	0.05

$\alpha_{\delta=1/23400} = (-0.025, 0.050)'$, $\alpha_\perp = (0.67, 0.33)'$

Bias	VEC(0)	$\mathrm{VEC}(0)_1$	$\mathrm{VEC}(0)_5$	$2\mathrm{SLS}_{V=1,\mathrm{RK}}$	$2\mathrm{SLS}_{V=5,\mathrm{RK}}$	$2\mathrm{SLS}_{V=10,\mathrm{RK}}$	$2\mathrm{SLS}_{V=20,\mathrm{RK}}$	$\text{CU-GMM}_{V=10,\mathrm{RK}}$	$\text{CU-GMM}_{V=20,\mathrm{RK}}$	$2\mathrm{SLS}_{V=1,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=5,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=10,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=20,\mathrm{HAC}}$	$\text{CU-GMM}_{V=10,\mathrm{HAC}}$	$\text{CU-GMM}_{V=20,\mathrm{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$	-0.11	-0.13	-0.15	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08	0.00	0.00	0.00	0.00	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	-0.15	-0.14	-0.15	-0.11	-0.10	-0.11	-0.11	-0.09	-0.09	-0.01	0.02	0.01	0.00	0.04	0.03
$\omega^2 = (0.0010, 0.0010)'$	-0.16	-0.15	-0.15	-0.13	-0.11	-0.12	-0.13	-0.08	-0.08	-0.08	0.00	-0.03	-0.05	0.07	0.06
$\omega^2 = (0.0010, 0.0005)'$	-0.32	-0.25	-0.19	-0.14	-0.11	-0.13	-0.15	-0.06	-0.07	-0.07	0.00	-0.05	-0.09	0.08	0.07

$\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\alpha_\perp = (1.00, 0.00)'$

RRMSE		$\mathrm{VEC}(0)_1$	$\mathrm{VEC}(0)_5$	$2\mathrm{SLS}_{V=1,\mathrm{RK}}$	$2\mathrm{SLS}_{V=5,\mathrm{RK}}$	$2\mathrm{SLS}_{V=10,\mathrm{RK}}$	$2\mathrm{SLS}_{V=20,\mathrm{RK}}$	$\text{CU-GMM}_{V=10,\mathrm{RK}}$	$\text{CU-GMM}_{V=20,\mathrm{RK}}$	$2\mathrm{SLS}_{V=1,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=5,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=10,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=20,\mathrm{HAC}}$	$\text{CU-GMM}_{V=10,\mathrm{HAC}}$	$\text{CU-GMM}_{V=20,\mathrm{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$		1.19	1.62	0.63	0.67	0.67	0.69	0.66	0.67	0.19	0.09	0.08	0.07	0.08	0.07
$\omega^2 = (0.0005, 0.0005)'$		0.85	1.03	0.63	0.55	0.60	0.65	0.47	0.47	0.24	0.13	0.11	0.09	0.13	0.13
$\omega^2 = (0.0010, 0.0010)'$		0.87	0.99	0.78	0.61	0.69	0.76	0.41	0.42	0.38	0.18	0.16	0.16	0.19	0.19
$\omega^2 = (0.0010, 0.0005)'$		0.76	0.74	0.48	0.37	0.43	0.49	0.26	0.26	0.20	0.09	0.08	0.10	0.11	0.11

$\alpha_{\delta=1/23400} = (-0.025, 0.050)'$, $\alpha_\perp = (0.67, 0.33)'$

RRMSE		$\mathrm{VEC}(0)_1$	$\mathrm{VEC}(0)_5$	$2\mathrm{SLS}_{V=1,\mathrm{RK}}$	$2\mathrm{SLS}_{V=5,\mathrm{RK}}$	$2\mathrm{SLS}_{V=10,\mathrm{RK}}$	$2\mathrm{SLS}_{V=20,\mathrm{RK}}$	$\text{CU-GMM}_{V=10,\mathrm{RK}}$	$\text{CU-GMM}_{V=20,\mathrm{RK}}$	$2\mathrm{SLS}_{V=1,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=5,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=10,\mathrm{HAC}}$	$2\mathrm{SLS}_{V=20,\mathrm{HAC}}$	$\text{CU-GMM}_{V=10,\mathrm{HAC}}$	$\text{CU-GMM}_{V=20,\mathrm{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$		1.20	1.40	0.85	0.76	0.77	0.78	0.76	0.77	1.06	0.33	0.31	0.25	0.31	0.26
$\omega^2 = (0.0005, 0.0005)'$		0.93	1.01	0.83	0.66	0.71	0.76	0.58	0.59	1.36	0.53	0.44	0.36	0.49	0.44
$\omega^2 = (0.0010, 0.0010)'$		0.94	0.98	0.95	0.71	0.79	0.85	0.52	0.53	1.42	0.72	0.57	0.47	0.83	0.79
$\omega^2 = (0.0010, 0.0005)'$		0.77	0.59	0.48	0.35	0.41	0.46	0.21	0.22	0.72	0.29	0.27	0.29	0.30	0.28

Table 8: Performance of the LS, 2SLS, and CU-GMM estimators of the discrete-time Information Share measure under Assumption MMN(CS): endogenous MA(∞) market microstructure noise

We simulate from the exact discretization of (1) with alternative speed-of-adjustment parameters, $\beta = (1,-1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1,-1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample at alternative frequencies, varying from tick-by-tick to 5 minutes. The market microstructure noises are endogenous MA(∞) processes with $\mathrm{Corr}\left(u_{m,t_i}, \varepsilon_{m',t_i}\right) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50^q for $q = 0, 1, 2, \ldots$, and variances ω_j^2 with $j = 1, 2$ ranging from 0.0001 to 0.001. We estimate discrete-time Information Share (IS) measures based on the LS, 2SLS, and CU-GMM estimators using other assets traded at different exchanges as instruments. Specifically, the cross-sectional dimension of instruments, V, ranges from $V \in (1, 5, 10, 20)$ for the 2SLS estimator and $V \in (10, 20)$ for the CU-GMM estimator. $\text{2SLS}_{V=1,\text{HAC}}$, $\text{2SLS}_{V=5,\text{HAC}}$, $\text{2SLS}_{V=10,\text{HAC}}$, and $\text{2SLS}_{V=20,\text{HAC}}$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by 2SLS from VEC(0) models fitted at the tick-by-tick frequency and estimates of the unconditional variance computed with the HAC estimator; $\text{CU-GMM}_{V=10,\text{HAC}}$ and $\text{CU-GMM}_{V=20,\text{HAC}}$ denote estimates of the IS measures that are constructed with estimates of $\alpha_\perp$ computed by CU-GMM from VEC(0) models fitted at the tick-by-tick frequency and estimates of the unconditional variance computed with the HAC estimator; and VEC(0) denotes estimates of the discrete-time IS measures that are constructed with estimates of $\alpha_\perp$ computed by least squares from VEC(0)models fitted at the tick-by-tick and estimates of the unconditional variance that are not robust to heteroskedaticity and serial correlation. We report the median bias of the estimates of the discrete-time Information Share (IS) measures and the ratio of their root median square errors (RRMSE) with respect to the estimates obtained using LS estimator.

$\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\alpha_\perp = (1.00, 0.00)'$

	RRMSE						Bias						
	$\text{2SLS}_{V=1,\text{HAC}}$	$\text{2SLS}_{V=5,\text{HAC}}$	$\text{2SLS}_{V=10,\text{HAC}}$	$\text{2SLS}_{V=20,\text{HAC}}$	$\text{CU-GMM}_{V=10,\text{HAC}}$	$\text{CU-GMM}_{V=20,\text{HAC}}$	VEC(0)	$\text{2SLS}_{V=1,\text{HAC}}$	$\text{2SLS}_{V=5,\text{HAC}}$	$\text{2SLS}_{V=10,\text{HAC}}$	$\text{2SLS}_{V=20,\text{HAC}}$	$\text{CU-GMM}_{V=10,\text{HAC}}$	$\text{CU-GMM}_{V=20,\text{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$	0.21	0.09	0.09	0.08	0.09	0.08	-0.20	0.00	0.00	0.00	0.00	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	0.26	0.13	0.11	0.10	0.14	0.14	-0.33	-0.02	0.03	0.02	0.01	0.04	0.04
$\omega^2 = (0.0010, 0.0010)'$	0.37	0.19	0.16	0.16	0.22	0.21	-0.35	-0.12	0.02	-0.01	-0.04	0.07	0.07
$\omega^2 = (0.0010, 0.0005)'$	0.21	0.10	0.09	0.10	0.12	0.11	-0.50	-0.07	0.02	-0.01	-0.05	0.06	0.06

$\alpha_{\delta=1/23400} = (-0.025, 0.050)'$, $\alpha_\perp = (0.67, 0.33)'$

	RRMSE						Bias						
	$\text{2SLS}_{V=1,\text{HAC}}$	$\text{2SLS}_{V=5,\text{HAC}}$	$\text{2SLS}_{V=10,\text{HAC}}$	$\text{2SLS}_{V=20,\text{HAC}}$	$\text{CU-GMM}_{V=10,\text{HAC}}$	$\text{CU-GMM}_{V=20,\text{HAC}}$	VEC(0)	$\text{2SLS}_{V=1,\text{HAC}}$	$\text{2SLS}_{V=5,\text{HAC}}$	$\text{2SLS}_{V=10,\text{HAC}}$	$\text{2SLS}_{V=20,\text{HAC}}$	$\text{CU-GMM}_{V=10,\text{HAC}}$	$\text{CU-GMM}_{V=20,\text{HAC}}$
$\omega^2 = (0.0001, 0.0001)'$	1.09	0.34	0.32	0.26	0.32	0.26	-0.10	0.00	0.00	0.00	0.00	0.00	0.00
$\omega^2 = (0.0005, 0.0005)'$	1.44	0.54	0.46	0.36	0.51	0.46	-0.14	0.00	0.02	0.01	0.00	0.04	0.03
$\omega^2 = (0.0010, 0.0010)'$	1.59	0.74	0.59	0.47	0.88	0.82	-0.15	-0.07	0.00	-0.03	-0.04	0.07	0.06
$\omega^2 = (0.0010, 0.0005)'$	0.75	0.29	0.27	0.29	0.31	0.29	-0.31	-0.06	0.00	-0.05	-0.09	0.08	0.07

Figure 1: The upper panel presents the empirical densities of the t-statistics based on Theorem 1(ii), while the second panel displays the empirical densities of the t-statistics based on Theorem 2(ii). The columns refer to alternative market microstructure levels, namely the first columns sets $\omega^2 = (0.0001, 0.0001)'$, the second columns presents the results for $\omega^2 = (0.0005, 0.0005)'$, and the third and fourth columns display the results for $\omega^2 = (0.001, 0.0005)'$ and $\omega^2 = (0.001, 0.001)'$, respectively. Each plot is computed from 1000 replications. The solid lines are the empirical densities and the dashed lines are the theoretical normal distribution with mean zero and unit variance.

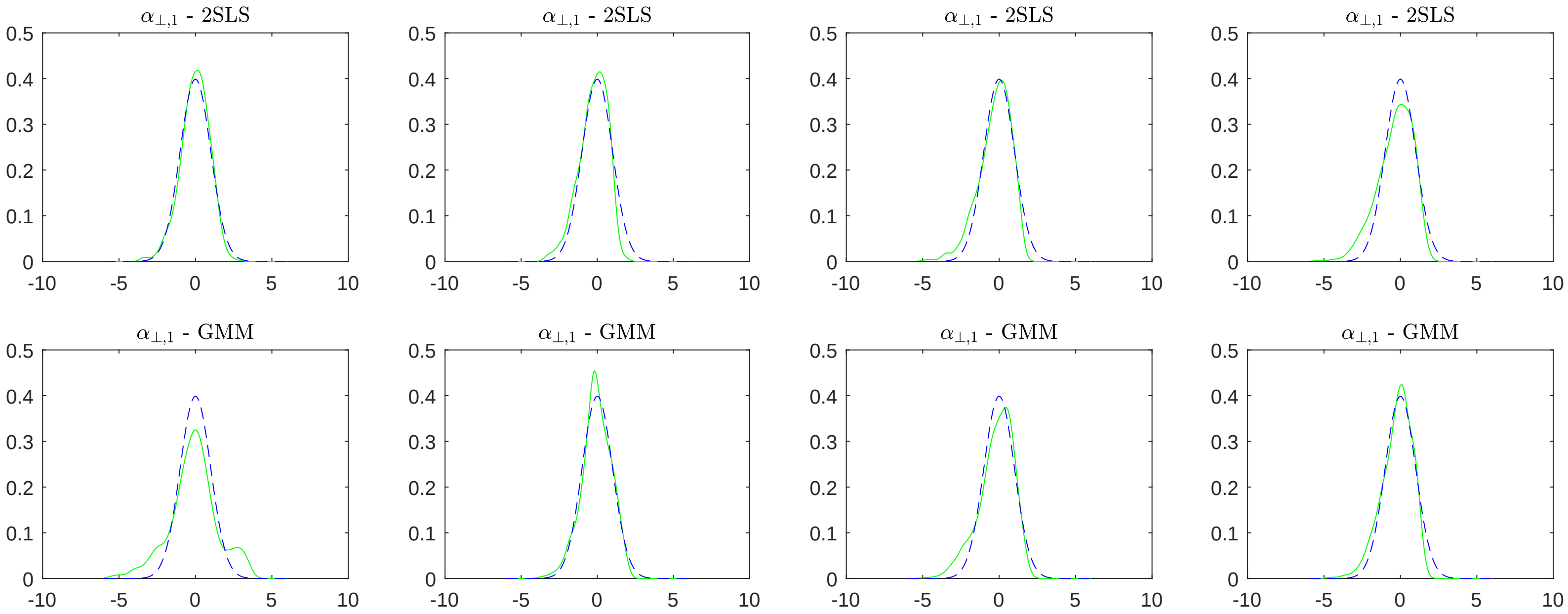

Figure 2: $\alpha_{\perp,1}$ estimates from prices with (without) market microstructure noise

We report the box plot of the estimates of $\alpha_{\perp,1}$. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample using refresh time at the tick-by-tick. The first panel reports $\alpha_{\perp,1}$ estimates computed using the LS and 2SLS from VEC(0) models fitted at tick data without market microstructure noise contamination. The 2SLS estimator uses valid and relevant instruments which are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$. The second panel displays the boxplot of $\widehat{\alpha}_{\perp,1}$ estimated by LS and 2SLS from VEC(0) models fitted at tick data in which prices are contaminated with market microstructure noise. The market microstructure noises are iid processes with variances ω_j^2 with $j = 1, 2$ given by $\omega^2 = (0.0001, 0.00075)'$. The third panel displays the boxplot of $\widehat{\alpha}_{\perp,1}$ estimated by LS_1 and LS_5 from prices that are contaminated with market microstructure noise. Specifically, LS_1 and LS_5 denote the LS estimators from VEC(0) models fitted at 1- and 5-minute sampling intervals; and the market microstructure noises are iid processes as discussed in the second panel. The black stars are the true values, the edges of boxes in the box plots refer to the 25% and 75% percentiles, and the red line represents the median.

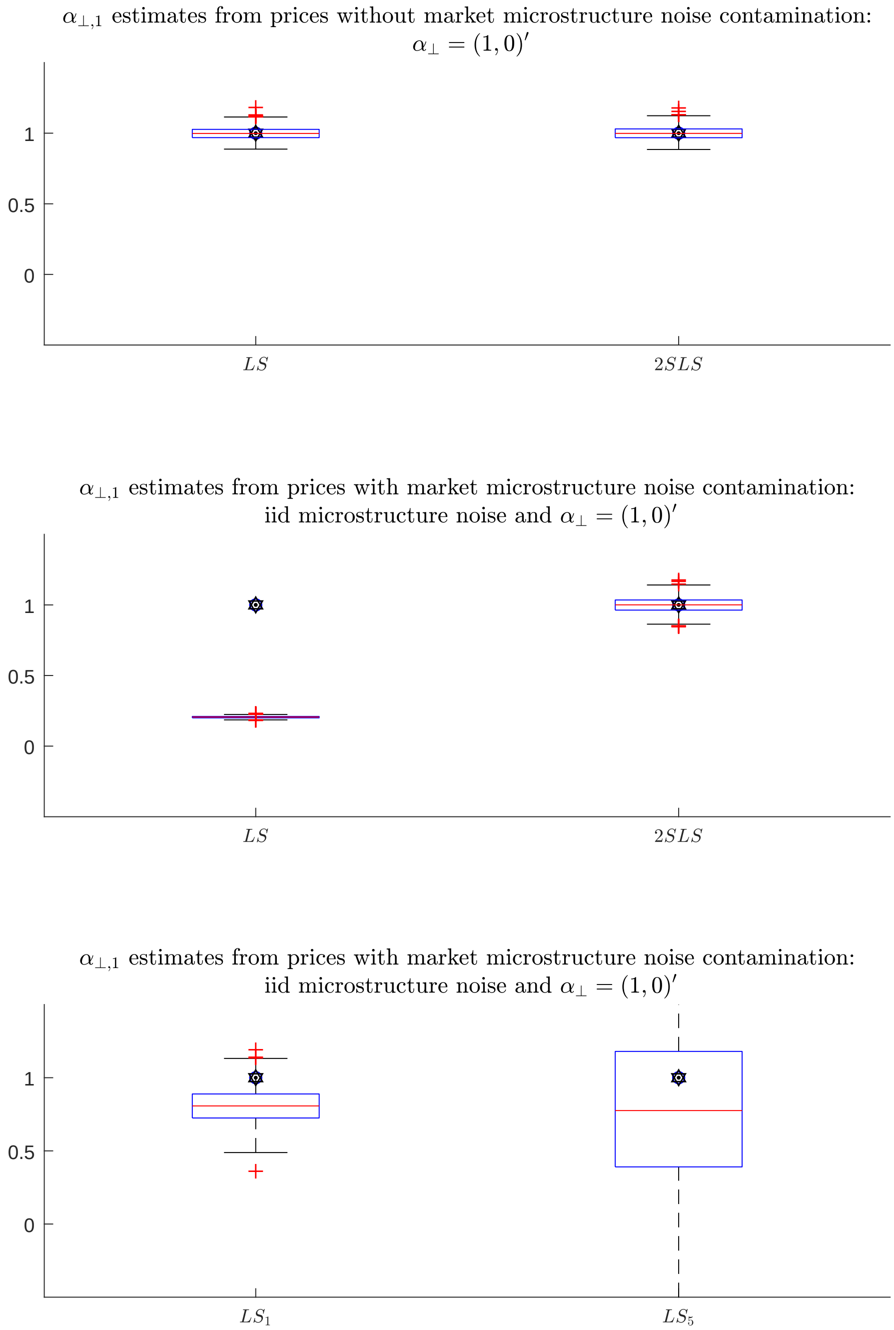

Figure 3: Impulse response function (IRF) under Assumption MMN(TS): iid market microstructure noise

We report the estimated IRFs for p_{2,t_i} sampled at the tick-by-tick frequency under different estimators and market microstructure noise processes. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample using refresh time at the tick-by-tick. The market microstructure noises are serially uncorrelated white noises with variances ω_j^2 with $j = 1, 2$ given by $\omega^2 = (0.0001, 0.0001)'$, $\omega^2 = (0.0005, 0.0005)'$, $\omega^2 = (0.001, 0.001)'$, and $\omega^2 = (0.001, 0.0005)'$ for the first, second, third, and fourth columns, respectively. We estimate $\alpha_{\delta,1}$ by LS and by 2SLS using $(p_{1,t_{i-1-\bar{q}-k}} - p_{2,t_{i-1-\bar{q}-k}})'$ as instruments, where k and $\bar{k}$ are integers such that $2 \leq k \leq \bar{k} \leq \infty$. Specifically, VEC(1) denotes estimates of $\alpha_{\delta,1}$ that are estimated by LS from VEC(1) model fitted at the tick-by-tick; LS denotes estimates of $\alpha_\perp$ that are estimated by LS from VEC(0) models fitted at the tick-by-tick; and 2SLS and 2SLS$_R$ are the estimates of $\alpha_\perp$ that are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $6 \leq \kappa \leq 10$, respectively. The solid lines correspond to the median IRFs, the shaded area is bounded by the 5% and 95% quantiles, and the black dotted lines are the true IRFs.

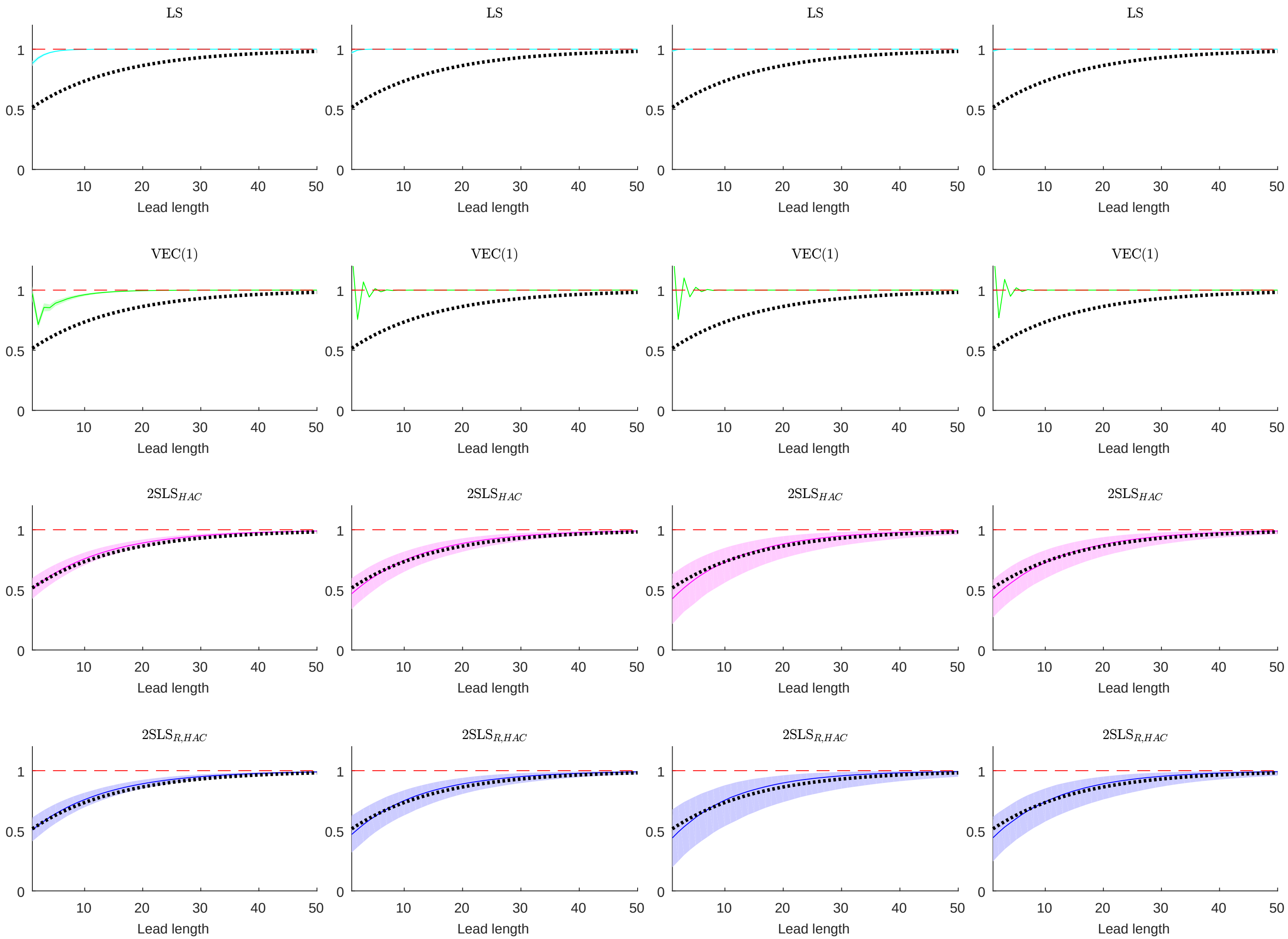

Figure 4: Impulse response function (IRF) under Assumption MMN(TS): endogenous MA(1) market microstructure noise

We report the estimated IRFs for p_{2,t_i} sampled at the tick-by-tick frequency under different estimators and market microstructure noise processes. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample using refresh time at the tick-by-tick. The market microstructure noises are endogenous MA(1) processes with $\mathrm{Corr}\,(u_{m,t_i}, \varepsilon_{m',t_i}) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50, and variances ω_j^2 with $j = 1, 2$ given by $\omega^2 = (0.0001, 0.0001)'$, $\omega^2 = (0.0005, 0.0005)'$, $\omega^2 = (0.001, 0.001)'$, and $\omega^2 = (0.001, 0.0005)'$ for the first, second, third, and fourth columns, respectively. We estimate $\alpha_{\delta,1}$ by LS and by 2SLS using $(p_{1,t_{i-1-\bar{q}-k}} - p_{2,t_{i-1-\bar{q}-k}})'$ as instruments, where k and $\bar{k}$ are integers such that $2 \leq k \leq \bar{k} \leq \infty$. Specifically, VEC(1) denotes estimates of $\alpha_{\delta,1}$ that are estimated by LS from VEC(1) model fitted at the tick-by-tick; LS denotes estimates of $\alpha_\perp$ that are estimated by LS from VEC(0) models fitted at the tick-by-tick; and 2SLS and 2SLS$_R$ are the estimates of $\alpha_\perp$ that are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $6 \leq \kappa \leq 10$, respectively. The solid lines correspond to the median IRFs, the shaded area is bounded by the 5% and 95% quantiles, and the black dotted lines are the true IRFs.

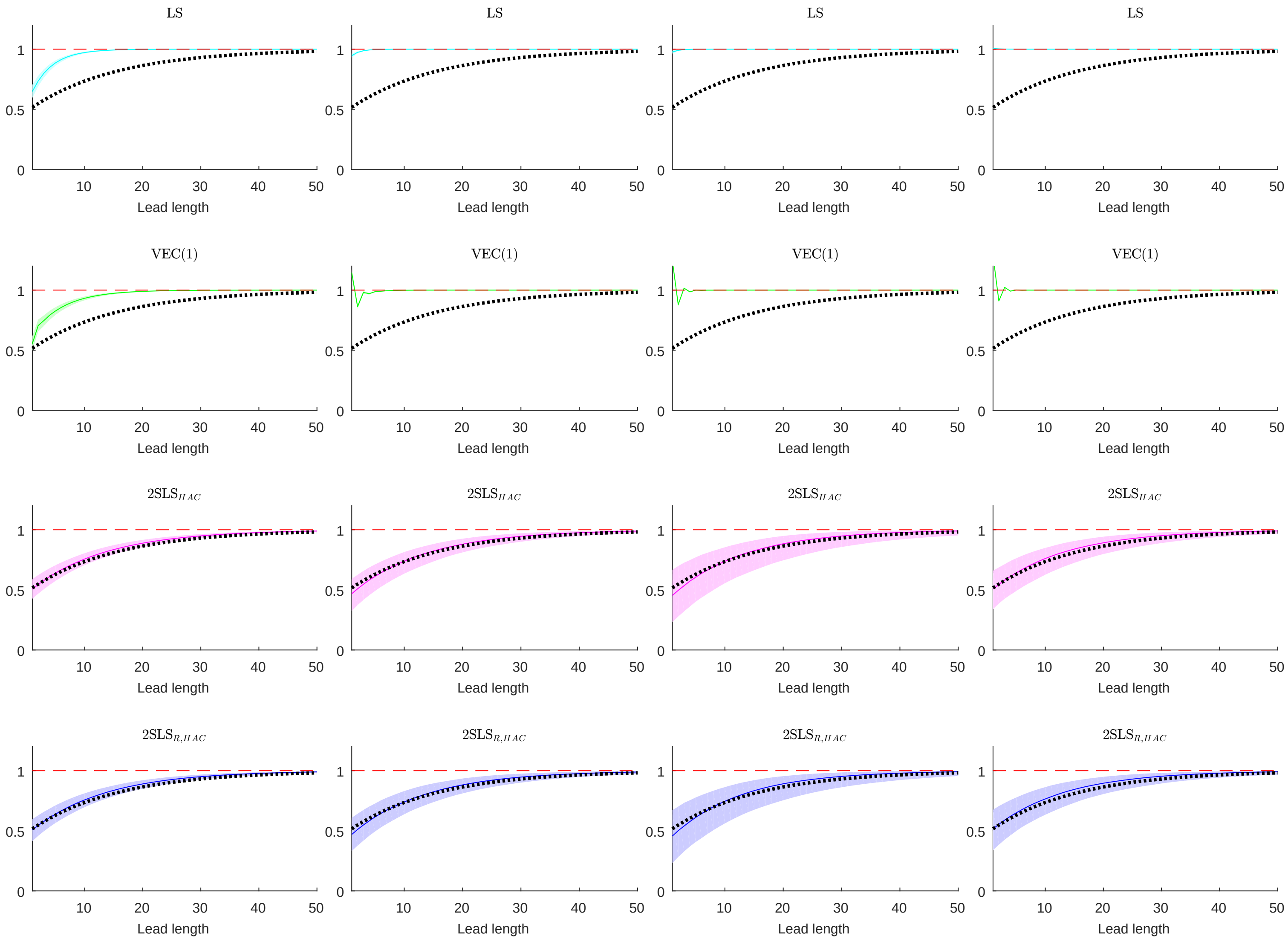

Figure 5: Impulse response function (IRF) under Assumption MMN(CS): endogenous MA(∞) market microstructure noise

We report the estimated IRFs for p_{2,t_i} sampled at the tick-by-tick frequency under different estimators and market microstructure noise processes. We simulate from the exact discretization of (1) $\alpha_{\delta=1/23400} = (0.000, 0.050)'$, $\beta = (1, -1)'$, market-specific integrated variances equal to one, and $\rho = 0.5$. We sample using refresh time at the tick-by-tick. The market microstructure noises are endogenous MA(∞) processes with $\text{Corr}\left(u_{m,t_i}, \varepsilon_{m',t_i}\right) = -0.2$ for $m, m' = 1, 2$, moving average parameters equal to -0.50^q for $q = 0, 1, 2, ...,$ and variances ω_j^2 with $j = 1, 2$ given by $\omega^2 = (0.0001, 0.0001)'$, $\omega^2 = (0.0005, 0.0005)'$, $\omega^2 = (0.001, 0.001)'$, and $\omega^2 = (0.001, 0.0005)'$ for the first, second, third, and fourth columns, respectively. We estimate α_δ the by 2SLS and CU-GMM estimators using other assets traded at different exchanges as instruments. Specifically, the cross-sectional dimension of instruments, V, ranges from $V \in (1, 5, 10, 20)$ for the 2SLS estimator and $V \in (10, 20)$ for the CU-GMM estimator. $\text{2SLS}_{V=1,\text{HAC}}$, $\text{2SLS}_{V=5,\text{HAC}}$, $\text{2SLS}_{V=10,\text{HAC}}$, and $\text{2SLS}_{V=20,\text{HAC}}$ denote estimates of the IRFs that are constructed with estimates of α_δ computed by 2SLS from VEC(0) models fitted at the tick-by-tick frequency and estimates of the unconditional variance computed with the HAC estimator; $\text{CU-GMM}_{V=10,\text{HAC}}$ and $\text{CU-GMM}_{V=20,\text{HAC}}$ denote estimates of the IS measures that are constructed with estimates of α_δ computed by CU-GMM from VEC(0) models fitted at the tick-by-tick frequency and estimates of the unconditional variance computed with the HAC estimator. The solid lines correspond to the median IRFs, the shaded area is bounded by the 5% and 95% quantiles, and the black dotted lines are the true IRFs.

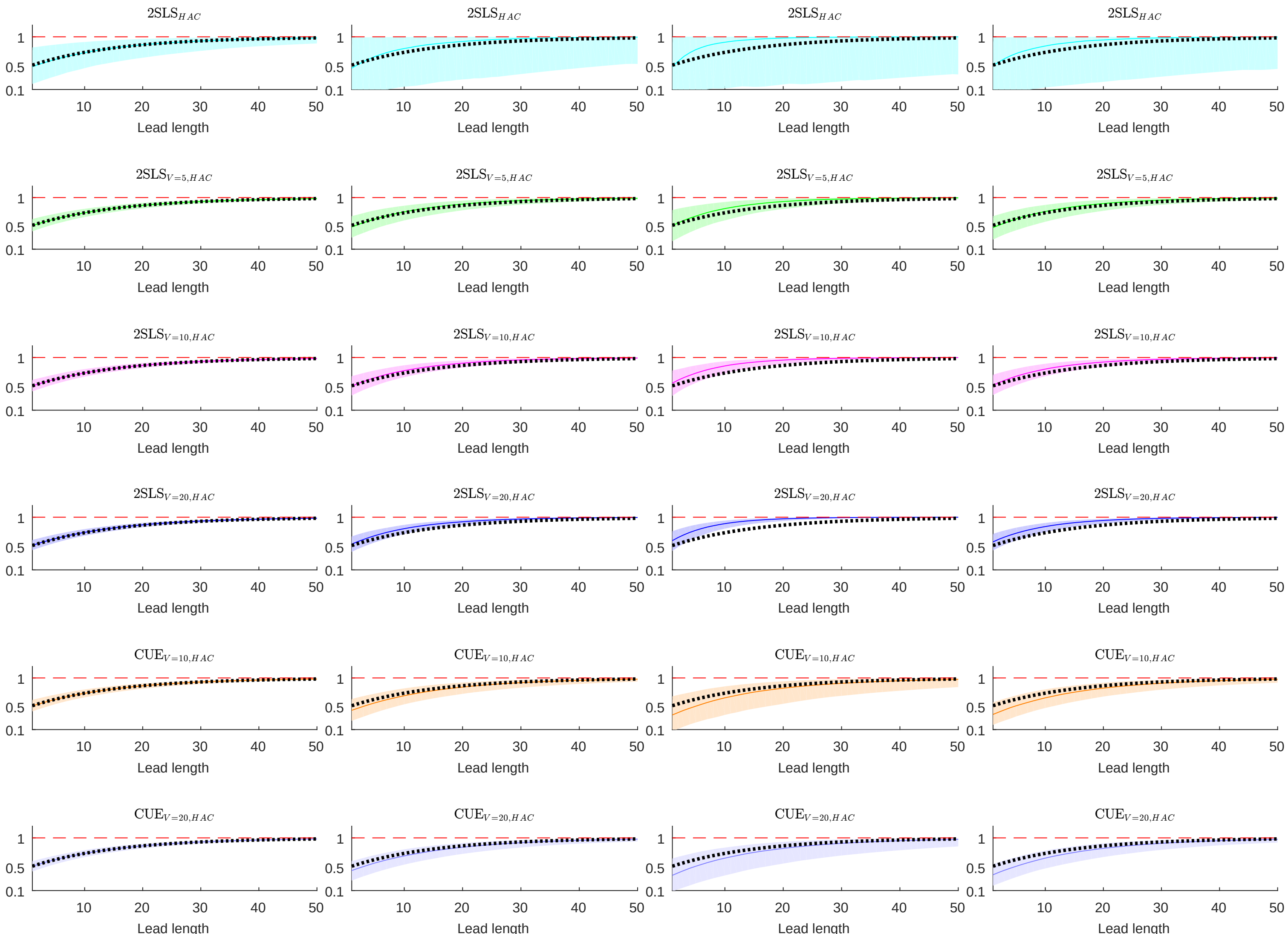

Table 9: Data description

We report summary statistics for raw and cleaned data of AA and 25 assets (potential instruments) covering NYSE (N), Nasdaq (T), NASDAQ OMX BX (B), Arca (P), and BATS (Z). The first column presents the ticker symbols for all assets; the first panel (Raw quotes) presents the number of quotes (in millions) for each stock on the five trading venues before any cleaning filter; the second panel (Mkt share: raw quotes) displays the market share (in %) for each trading venue as well as the total market share of all five exchanges; the third panel (Cleaned obs) presents the number of quotes (in millions) for each stock on the five trading venues after the implementation of the cleaning procedure; and the fourth panel (Average duration) shows the average duration (in seconds) between quotes for each trading venue. “-” denotes that the asset does not trade at this trading venue.

	Raw quotes ('000,000)					Mkt share: raw quotes (%)						Cleaned obs ('000,000)					Average duration (seconds)				
	N	T	B	P	Z	N	T	B	P	Z	Total	N	T	B	P	Z	N	T	B	P	Z
aa	35.5	9.8	9.7	11.2	9.8	27.0%	7.4%	7.4%	8.5%	7.4%	57.8%	3.8	3.0	2.1	2.8	2.6	1.6	2.0	2.7	2.1	2.3
aapl	-	17.7	32.9	25.4	17.7	-	10.4%	19.4%	15.0%	10.4%	55.3%	-	3.5	3.9	3.7	3.3	-	1.7	1.5	1.6	1.8
bac	69.5	23.8	29.6	27.2	23.8	21.6%	7.4%	9.2%	8.5%	7.4%	54.1%	4.4	3.7	3.3	3.8	3.4	1.3	1.6	1.8	1.5	1.7
brkb	30.6	16.9	3.1	6.3	16.9	33.4%	18.4%	3.4%	6.8%	18.4%	80.3%	3.7	2.3	1.0	1.8	2.4	1.6	2.6	5.6	3.3	2.4
csco	-	23.1	16.6	21.1	23.1	-	11.8%	8.5%	10.8%	11.8%	42.8%	-	1.5	1.1	1.3	1.5	-	3.9	5.2	4.5	3.8
dal	27.7	12.0	5.9	10.8	12.0	23.1%	10.0%	4.9%	9.0%	10.0%	57.0%	2.8	2.3	1.3	2.1	1.9	2.1	2.6	4.5	2.8	3.1
gm	36.5	14.2	4.9	10.7	14.2	26.4%	10.3%	3.5%	7.7%	10.3%	58.2%	3.2	2.3	1.1	2.0	2.0	1.8	2.6	5.2	2.9	3.0
goog	-	5.9	12.1	13.1	5.9	-	8.5%	17.3%	18.8%	8.5%	53.1%	-	2.0	2.0	2.1	1.7	-	2.9	2.9	2.8	3.5
hpq	50.0	21.0	14.5	20.6	21.0	21.0%	8.8%	6.1%	8.7%	8.8%	53.5%	3.7	3.2	2.1	2.9	2.6	1.6	1.9	2.7	2.0	2.2
ibm	23.2	8.8	9.8	14.9	8.8	24.2%	9.2%	10.2%	15.5%	9.2%	68.2%	3.6	2.3	2.0	2.7	2.2	1.6	2.6	2.9	2.2	2.6
jcp	26.0	16.2	5.5	14.9	16.2	19.5%	12.1%	4.2%	11.2%	12.1%	59.1%	2.9	2.5	1.3	2.3	2.1	2.0	2.3	4.6	2.5	2.7
jnj	39.4	23.0	6.2	15.6	23.0	22.0%	12.8%	3.4%	8.7%	12.8%	59.8%	3.9	3.1	1.6	2.8	3.0	1.5	1.9	3.7	2.1	2.0
jpm	120.6	58.3	21.0	42.6	58.3	25.6%	12.4%	4.5%	9.1%	12.4%	63.9%	4.8	4.2	3.0	4.1	4.2	1.2	1.4	2.0	1.4	1.4
ko	37.9	15.3	6.1	11.6	15.3	25.1%	10.2%	4.0%	7.7%	10.2%	57.1%	3.9	2.9	1.3	2.4	2.3	1.5	2.1	4.4	2.5	2.5
mo	32.1	11.7	9.5	10.5	11.7	24.4%	8.9%	7.2%	7.9%	8.9%	57.2%	3.4	2.5	1.9	2.2	2.0	1.7	2.3	3.1	2.7	2.9
mrk	42.9	19.8	9.8	14.2	19.8	23.6%	10.9%	5.4%	7.8%	10.9%	58.5%	3.9	3.0	2.0	2.6	2.6	1.5	1.9	3.0	2.3	2.2
mrvl	-	10.8	6.7	9.5	10.8	-	12.5%	7.7%	11.0%	12.5%	43.8%	-	2.5	1.3	2.1	2.1	-	2.3	4.4	2.8	2.7
msft	-	31.9	22.7	27.4	31.9	-	12.1%	8.6%	10.3%	12.1%	43.1%	-	4.1	2.9	3.7	3.6	-	1.4	2.0	1.6	1.6
nok	24.9	8.6	7.9	10.5	8.6	24.3%	8.4%	7.8%	10.2%	8.4%	59.1%	2.7	1.9	1.5	2.1	1.7	2.2	3.1	3.9	2.8	3.5
orcl	-	32.8	16.7	25.1	32.8	-	14.1%	7.2%	10.8%	14.1%	46.1%	-	3.7	2.5	3.5	3.7	-	1.6	2.3	1.7	1.6
pfe	49.9	18.5	14.5	17.7	18.5	22.9%	8.5%	6.7%	8.1%	8.5%	54.8%	4.0	3.1	2.3	3.0	2.7	1.5	1.9	2.6	2.0	2.2
pg	31.1	20.9	7.6	18.8	20.9	18.0%	12.1%	4.4%	10.9%	12.1%	57.5%	3.6	3.1	1.7	2.8	2.7	1.6	1.9	3.4	2.1	2.2
vz	41.0	17.0	9.2	15.2	17.0	23.6%	9.8%	5.3%	8.7%	9.8%	57.2%	3.9	2.9	1.9	2.7	2.5	1.5	2.0	3.1	2.2	2.3
wfc	59.6	35.7	22.2	29.6	35.7	18.9%	11.4%	7.1%	9.4%	11.4%	58.1%	4.3	3.6	2.8	3.7	3.7	1.4	1.6	2.1	1.6	1.6
xom	62.1	48.8	5.2	33.6	48.8	22.0%	17.2%	1.8%	11.9%	17.2%	70.1%	4.8	3.9	1.4	4.0	4.2	1.2	1.5	4.3	1.5	1.4
yhoo	-	15.1	10.0	12.8	15.1	-	10.6%	7.0%	9.0%	10.6%	37.2%	-	3.0	1.8	2.5	2.4	-	2.0	3.2	2.3	2.4
Total	892.6	561.4	329.1	482.6	551.6	17.4%	11.2%	6.7%	9.8%	11.2%	56.2%	67.6	72.6	49.1	71.6	69.5	1.6	2.0	2.9	2.1	2.2

Table 10: Continuous-time CS estimates for NYSE, $\alpha_{N,\perp}$, under assumption MMN(TS)

We estimate monthly orthogonal complements of the speed-of-adjustment parameters for the NYSE, $\alpha_{N,\perp}$. Specifically, VEC(0), VEC(0)$_1$ and VEC(0)$_5$ denote estimates of $\alpha_{N,\perp}$ that are estimated by LS from VEC(0) models fitted at tick data, 1-minute, and 5-minute sampling intervals, respectively; and 2SLS and 2SLS$_1$ are the estimates of $\alpha_{N,\perp}$ that are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $3 \leq \kappa \leq 7$, respectively. Finally, we report the p-values of the Endogeneity test using the VEC(0), 2SLS, and 2SLS$_1$ estimates of α_δ, F and F$_1$ denote the F-statistics ($\times 10^{-3}$ from the auxiliary regressions using the two choices of instruments, and R^2 and R$^2_1$ are the R^2 measures from the auxiliary regressions using the two choices of instruments.

	VEC(0)	VEC(0)$_1$	VEC(0)$_5$	2SLS	2SLS$_1$	H	H$_1$	F ($\times 10^{-3}$)	R^2	F$_1$ ($\times 10^{-3}$)	R^2_R
Jun	0.47	1.11	0.76	0.87	0.95	0.00	0.00	23.11	0.31	12.25	0.20
	(0.01)	(0.146)	(0.863)	(0.029)	(0.04)						
Jul	0.46	1.09	1.27	0.79	0.91	0.00	0.00	42.82	0.48	24.51	0.35
	(0.012)	(0.166)	(0.478)	(0.027)	(0.034)						
Aug	0.40	0.83	1.07	0.88	1.02	0.00	0.35	20.42	0.29	11.20	0.18
	(0.012)	(0.152)	(0.767)	(0.043)	(0.063)						
Sep	0.29	0.66	0.38	0.90	1.05	0.00	0.10	21.11	0.32	12.63	0.22
	(0.012)	(0.156)	(0.57)	(0.034)	(0.052)						
Oct	0.29	0.66	0.41	0.76	0.83	0.00	0.70	28.85	0.37	17.54	0.26
	(0.009)	(0.10)	(0.323)	(0.029)	(0.04)						
Nov	0.44	0.68	0.09	0.85	0.90	0.00	0.00	26.89	0.39	15.71	0.28
	(0.01)	(0.069)	(0.251)	(0.027)	(0.039)						
Dec	0.36	0.67	0.79	0.82	0.98	0.00	0.91	30.22	0.41	17.30	0.29
	(0.009)	(0.069)	(0.326)	(0.024)	(0.033)						
Jan	0.39	0.39	-1.71	0.75	0.83	0.00	0.00	33.97	0.44	20.81	0.33
	(0.015)	(0.16)	(2.588)	(0.052)	(0.064)						
Feb	0.39	0.82	0.59	0.86	0.99	0.00	0.00	40.77	0.51	25.49	0.39
	(0.011)	(0.075)	(0.257)	(0.028)	(0.038)						
Mar	0.42	0.64	0.57	0.91	1.12	0.00	0.00	30.10	0.44	18.57	0.33
	(0.013)	(0.085)	(0.38)	(0.035)	(0.053)						
Apr	0.47	0.80	1.25	0.92	1.07	0.00	0.00	32.71	0.40	19.33	0.28
	(0.009)	(0.106)	(0.514)	(0.028)	(0.04)						
May	0.41	0.79	1.09	0.78	0.88	0.00	0.00	32.65	0.42	20.61	0.31
	(0.011)	(0.102)	(0.552)	(0.031)	(0.043)						
Year	0.40	0.75	0.53	0.84	0.95	0.00	0.00	361.08	0.40	214.84	0.28
	(0.003)	(0.033)	(0.112)	(0.009)	(0.013)						

Table 11: Continuous- and discrete-time IS estimates for NYSE under assumption MMN(TS)

We report monthly and annual averages of the continuous- and discrete-time IS measures for the NYSE market. Specifically, the first panel presents the estimates for the continuous-time IS measure, IS, and the second panel displays the estimates for the discrete-time IS measure, IS_δ. As for the estimates of IS, VEC(0) and VEC(0)$_1$ denote the estimates computed by LS from VEC(0) models fitted at tick-data and 1-minute sampling intervals, respectively; 2SLS$_{\text{RK}}$ and 2SLS$_{\text{RK},1}$ are the estimates that combine the realized kernel estimates of the integrated variance and the 2SLS estimates of α_δ from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $3 \leq \kappa \leq 7$, respectively; and 2SLS$_{\text{HAC}}$ and 2SLS$_{\text{HAC},1}$ denote the estimates of IS measures constructed using the 2SLS estimates of α_δ and the HAC estimates of the unconditional variance from the residuals of a VEC(0) model fitted at tick data.

	IS						IS_δ			
	VEC(0)	VEC(0)$_1$	2SLS$_{\text{RK}}$	2SLS$_{\text{RK},1}$	2SLS$_{\text{HAC}}$	2SLS$_{\text{HAC},1}$	VEC(0)	2SLS$_{\text{HAC}}$	2SLS$_{\text{HAC},1}$	Duration (sec)
Jun	0.49	0.50	0.50	0.51	0.61	0.57	0.50	0.56	0.56	1.96
	(0.034)	(0.003)	(0.003)	(0.004)	(0.063)	(0.044)	(0.028)	(0.027)	(0.037)	(0.168)
Jul	0.48	0.50	0.50	0.51	0.60	0.57	0.48	0.55	0.56	2.22
	(0.059)	(0.006)	(0.005)	(0.008)	(0.054)	(0.045)	(0.051)	(0.035)	(0.038)	(0.426)
Aug	0.43	0.50	0.50	0.50	0.59	0.56	0.44	0.55	0.55	2.17
	(0.051)	(0.004)	(0.002)	(0.004)	(0.079)	(0.064)	(0.045)	(0.05)	(0.047)	(0.247)
Sep	0.37	0.49	0.50	0.51	0.59	0.55	0.39	0.55	0.55	1.95
	(0.041)	(0.004)	(0.002)	(0.004)	(0.075)	(0.047)	(0.032)	(0.03)	(0.04)	(0.14)
Oct	0.36	0.49	0.50	0.51	0.58	0.55	0.38	0.54	0.55	2.03
	(0.041)	(0.004)	(0.003)	(0.006)	(0.066)	(0.043)	(0.036)	(0.028)	(0.038)	(0.232)
Nov	0.46	0.49	0.51	0.51	0.59	0.56	0.47	0.57	0.55	2.51
	(0.043)	(0.007)	(0.004)	(0.006)	(0.058)	(0.035)	(0.037)	(0.027)	(0.032)	(0.744)
Dec	0.41	0.49	0.51	0.51	0.64	0.60	0.42	0.57	0.59	2.29
	(0.056)	(0.009)	(0.004)	(0.008)	(0.078)	(0.049)	(0.05)	(0.042)	(0.045)	(0.833)
Jan	0.46	0.49	0.50	0.51	0.60	0.58	0.46	0.56	0.57	2.31
	(0.049)	(0.006)	(0.004)	(0.006)	(0.087)	(0.063)	(0.044)	(0.05)	(0.054)	(0.311)
Feb	0.42	0.48	0.51	0.51	0.63	0.57	0.43	0.57	0.57	2.30
	(0.108)	(0.029)	(0.007)	(0.014)	(0.098)	(0.06)	(0.101)	(0.065)	(0.056)	(0.35)
Mar	0.46	0.49	0.51	0.51	0.65	0.59	0.46	0.57	0.58	2.49
	(0.068)	(0.007)	(0.005)	(0.008)	(0.078)	(0.05)	(0.058)	(0.035)	(0.044)	(0.377)
Apr	0.50	0.50	0.51	0.51	0.64	0.58	0.50	0.58	0.57	2.15
	(0.067)	(0.006)	(0.004)	(0.006)	(0.066)	(0.047)	(0.057)	(0.038)	(0.039)	(0.266)
May	0.46	0.50	0.50	0.51	0.57	0.55	0.47	0.54	0.54	2.30
	(0.078)	(0.009)	(0.004)	(0.006)	(0.084)	(0.058)	(0.067)	(0.039)	(0.05)	(0.317)
Year	0.44	0.49	0.50	0.51	0.61	0.57	0.45	0.56	0.56	2.22
	(0.074)	(0.011)	(0.004)	(0.007)	(0.077)	(0.052)	(0.064)	(0.041)	(0.045)	(0.444)

Table 12: Estimates under assumption MMN(CS)

We estimate monthly orthogonal complements of the speed-of-adjustment parameters for the NYSE, $\alpha_{N,\perp}$, using the CU-GMM estimator. Specifically, we select instruments using two rules: the first panel reports the results for instruments selected from stocks which posts at least 7,500 quotes per day within an entire month; the second panel selects instruments from stocks which post at least 5,000 quotes per day within an entire month. $\alpha_{\perp,N}$ denotes the CU-GMM estimate of the CS measure for NYSE; $\alpha_{\perp,N}=0$ and $\alpha_{\perp,T}=0$, denote p-values from the null hypothesis of $\alpha_{\perp,N}=0$ and $\alpha_{\perp,T}=0$, R^2_{XZ} and F_{XZ} are the R^2 and the F-statistics from the first stage of the 2SLSLS, V is the number of instruments selected, and Duration (sec.) is the average duration after aggregating the data using the refresh-time scheme.

	min. 7,500 quotes per day							min. 5,000 quotes per day						
	$\alpha_{\perp,N}$	$\alpha_{\perp,N}=0$	$\alpha_{\perp,T}=0$	R^2_{XZ}	F_{XZ}	V	Duration (sec.)	$\alpha_{\perp,N}$	$\alpha_{\perp,N}=0$	$\alpha_{\perp,T}=0$	R^2_{XZ}	F_{XZ}	V	Duration (sec.)
Jun-12	1.00	0.00	1.00	0.00	9.22	19	8.27	1.00	0.52	1.00	0.00	3.22	34	19.86
Jul-12	1.00	0.10	1.00	0.00	10.46	3	3.89	1.00	0.04	1.00	0.00	5.98	16	10.88
Aug-12	1.00	0.02	1.00	0.00	10.89	7	5.59	1.00	0.50	1.00	0.01	4.39	24	17.04
Sep-12	1.00	0.01	1.00	0.01	3.37	13	7.20	1.00	0.09	1.00	0.01	4.82	30	19.83
Oct-12	0.94	0.01	0.87	0.00	2.97	15	6.72	1.00	0.04	1.00	0.00	4.33	31	18.38
Nov-12	1.00	0.03	1.00	0.00	12.31	3	4.37	1.00	0.03	1.00	0.00	2.36	15	8.92
Dec-12						0	2.11	1.00	0.02	1.00	0.00	5.59	9	9.31
Jan-13	1.00	0.17	1.00	0.00	2.84	15	7.90	0.75	0.25	0.72	0.00	2.23	36	24.25
Feb-13	1.00	0.64	1.00	0.00	2.82	15	9.30	1.00	0.04	1.00	0.01	3.58	32	19.98
Mar-13	1.00	0.07	1.00	0.00	12.67	17	9.64	1.00	0.62	1.00	0.01	5.11	37	25.38
Apr-13	1.00	0.05	1.00	0.00	9.79	23	11.13	1.00	0.00	1.00	0.01	5.34	40	20.55
May-13	1.00	0.00	1.00	0.01	3.27	20	11.03	1.00	0.01	1.00	0.00	4.36	36	18.42

Figure 6: Daily estimates of $\alpha_\perp$: AA

We report the daily estimates of $\alpha_\perp$ for AA computed from different estimators. We estimate $\alpha_{\perp,1}$ at tick-by-tick frequency by LS and by 2SLS using $(p_{1,t_{i-1-\bar{q}-k}} - p_{2,t_{i-1-\bar{q}-k}})'$ as instruments, where k and $\bar{k}$ are integers such that $2 \leq k \leq \bar{k} \leq \infty$. Specifically, VEC(1) denotes estimates of $\alpha_{\delta,1}$ that are estimated by LS from VEC(1) model fitted at tick-by-tick; LS denotes estimates of $\alpha_\perp$ that are estimated by LS from VEC(0) models fitted at tick-by-tick; and 2SLS and 2SLS$_R$ are the estimates of $\alpha_\perp$ that are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $3 \leq \kappa \leq 7$, respectively. The solid lines correspond to the estimates of $\alpha_\perp$ and the shaded area is bounded by the 95% confidence interval.

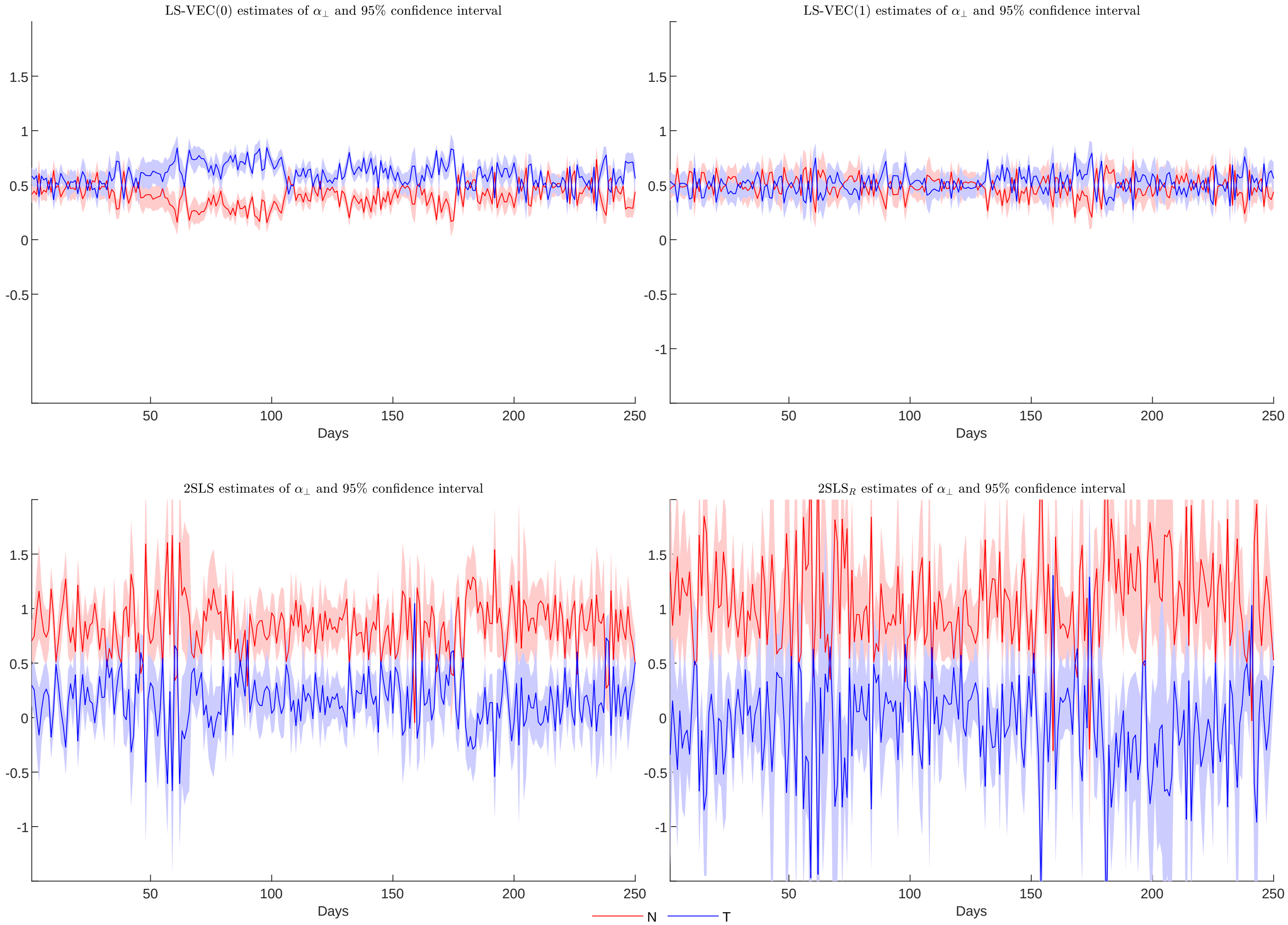

Figure 7: Impulse response function (IRF): AA

We report the estimated IRFs for AA sampled at tick-by-tick frequency under different estimators. We estimate $\alpha_{\delta,1}$ by LS and by 2SLS using $(p_{1,t_{i-1-\bar{q}-k}} - p_{2,t_{i-1-\bar{q}-k}})'$ as instruments, where k and $\bar{k}$ are integers such that $2 \leq k \leq \bar{k} \leq \infty$. Specifically, VEC(1) denotes estimates of $\alpha_{\delta,1}$ that are estimated by LS from VEC(1) model fitted at tick-by-tick; LS denotes estimates of $\alpha_{\perp}$ that are estimated by LS from VEC(0) models fitted at tick-by-tick; and 2SLS and 2SLS$_R$ are the estimates of $\alpha_{\perp}$ that are estimated by 2SLS from VEC(0) models fitted at tick data, where the instruments are selected as $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $2 \leq \kappa \leq 6$ and $(p_{1,t_{i-k}} - p_{2,t_{i-k}})'$ for $3 \leq \kappa \leq 7$, respectively. The solid lines correspond to the median IRFs and the shaded area is bounded by the 5% and 95% quantiles.