Web Appendix

Order Exposure and Liquidity Coordination: Does Hidden Liquidity Harm Price Efficiency?

Gökhan Cebiroğlu
University of Vienna

Nikolaus Hautsch
University of Vienna
Center for Financial Studies

Ulrich Horst
Humboldt Universität zu Berlin

September 16, 2014

Descriptive Statistics

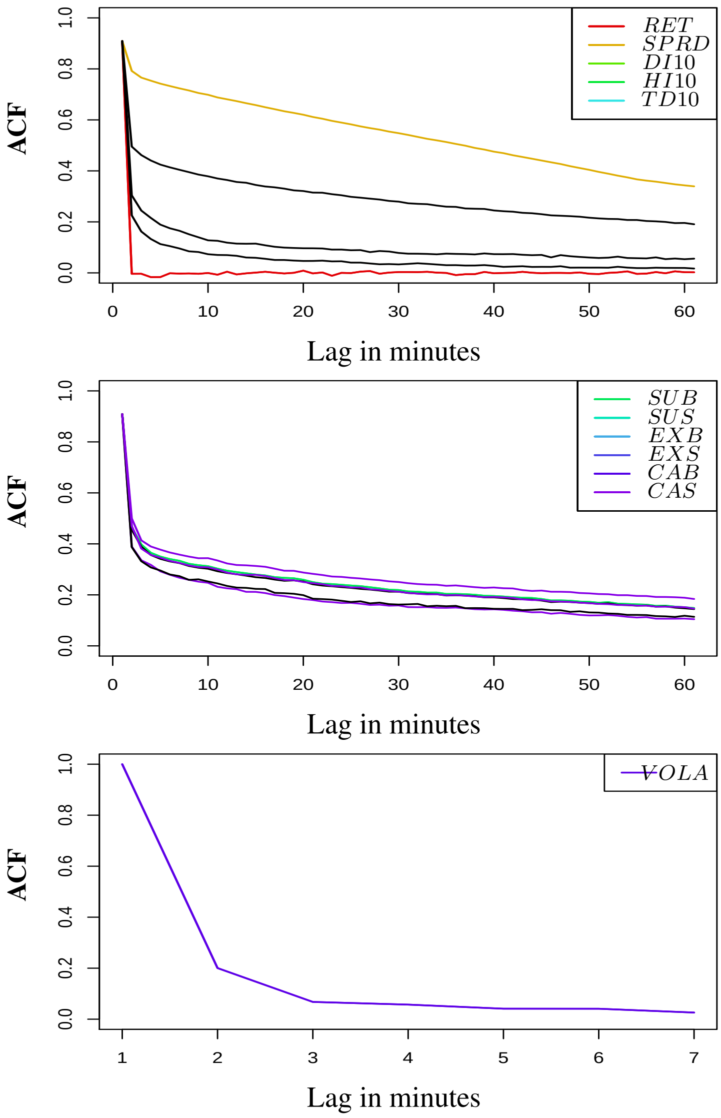
Table WA1: Time-series averages of mid-quotes, bid-ask spreads, visible and hidden depth, and order activities based on one-minute aggregates for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC. Reported variables: Averages of one-minute mid-quotes (MQ), the bid-ask spread (SPR), visible depth on the first 10 levels of the book ($D10$), and total hidden depth on the first 10 levels ($HD10$), and averages of minute-by-minute aggregated volume of limit buy order submissions (SUB), limit sell order submissions (SUS), limit buy order cancelations (CAB), limit sell order cancelations (CAS), buy limit order executions (EXB), and sell limit order executions (EXS). Data are based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data. The sample period is from November to December 2008, corresponding to $15,600$ one-minute intervals.

	MQ (in \$)	SPR (in \$)	$D10$ (10^3 sh.)	$HD10$ (10^3 sh.)	SUB (10^3 sh.)	SUS (10^3 sh.)	CAB (10^3 sh.)	CAS (10^3 sh.)	EXB (10^3 sh.)	EXS (10^3 sh.)
APC	36.42	0.07	1.41	0.89	22.20	18.99	21.47	21.29	1.63	1.68
AZO	117.14	0.53	0.24	0.51	7.36	4.80	7.18	6.93	0.28	0.31
CAH	33.87	0.05	2.24	0.40	8.08	7.58	8.43	8.41	0.60	0.62
EMR	33.08	0.04	3.41	0.45	18.81	19.50	18.23	19.10	1.82	1.54
GAS	37.84	0.24	1.11	0.26	3.94	4.01	4.05	4.27	0.14	0.13
GOOG	301.67	0.30	0.65	1.48	11.04	9.19	9.07	10.19	2.11	2.11
LEG	14.75	0.03	9.28	0.26	11.48	11.45	11.48	12.11	0.55	0.54
PAYX	26.31	0.02	9.88	0.28	31.51	26.39	29.50	30.02	2.09	2.10
STJ	31.52	0.08	2.23	0.27	13.40	12.36	13.44	13.43	0.87	0.85
TDC	13.87	0.03	6.06	0.45	4.96	4.80	5.23	4.97	0.25	0.26
Average	64.65	0.14	3.65	0.53	13.28	11.91	12.81	13.07	1.03	1.01

Table WA2: Summary statistics (mean, standard deviation and 10%, 25%, 75% and 90% quantiles) of mid-quotes, bid-ask spreads, visible and hidden depth and order activities on one-minute aggregates for the stock AZO. Reported variables: one-minute mid-quotes (MQ), minute-by-minute snapshots of bid-ask spreads (SPR), visible depth on the first 10 levels of the book ($D10$), total hidden depth on the first 10 levels ($HD10$), and of total depth and displayed depth imbalances, defined as standing buy volume in excess of sell volume ($DI10$ and $HI10$). Moreover, we report statistics of the minute-by-minute aggregated volume of limit buy order submissions (SUB) and limit sell order submissions (SUS), minute-by-minute aggregated volume of buy limit order cancellations (CAB) and sell limit order cancellations (CAS), and minute-by-minute aggregated volume of buy limit order executions (EXB) and sell limit order executions (EXS). Order flow minute-by-minute aggregation is based on NASDAQ ITCH data and one-minute snapshots are based on NASDAQ ModelView data. Sample period November to December 2008 corresponding to $15,600$ one-minute intervals.

Variable	Mean	St. Dev.	$q10$	$q25$	$q75$	$q90$
MQ (in $)	117.14	12.92	100.27	106.48	129.02	132.45
SPR (in $)	0.53	3.01	0.10	0.14	0.27	0.38
$DI10$ (in $)	−0.02	0.16	−0.23	−0.12	0.08	0.18
$HI10$ (in $)	0.01	0.60	−0.46	−0.16	0.19	0.49
$D10$ (in 1000 sh.)	0.24	0.16	0.04	0.11	0.33	0.46
$HD10$ (in 1000 sh.)	0.51	0.60	0.03	0.15	0.67	1.14
SUB (in 1000 sh.)	7.36	23.56	0.70	1.82	8.20	15.33
SUS (in 1000 sh.)	4.80	14.77	0.00	0.00	5.90	11.47
CAB (in 1000 sh.)	7.18	23.40	0.60	1.70	7.90	14.90
CAS (in 1000 sh.)	6.93	16.54	0.62	1.77	8.00	14.56
EXB (in 1000 sh.)	0.28	0.53	0.00	0.00	0.36	0.76
EXS (in 1000 sh.)	0.31	0.53	0.00	0.00	0.40	0.80

Figure WA3: Estimated average unconditional autocorrelation functions of buy- and sell-side order flow volume variables (EXB, EXS, CAB,CAS, SUB, SUS) in quantities that refer to the state of the order book, including the spread ($SPRD$), the hidden and displayed order imbalances ($HI10$, $DI10$), the total order depth ($TD10 = H10 + D10$), the midpoint return RET, and return volatility $VOLA$. $VOLA$ is computed as the sum of squared 1-min returns over a 10-min window. Accordingly, the ACF of $VOLA$ is computed based on 10-minute intervals. Autocorrelation estimates for the other variables are based on one-minute aggregates over snapshots. In particular, order flow volumes are aggregated on a minute-by-minute basis, while order-book quantities originate from one-minute snapshots of the order book. Based on one-minute aggregates of NASDAQ ITCH data and one-minute snapshots of NASDAQ ModelView data for the stocks APC, AZO, CAH, GAS, GOOG, EMR, LEG, PAYX, STJ, and TDC, November to December 2008.



Asymptotic Properties of Generalized Impulse Response Functions

To derive the asymptotic properties of the cumulative impulse response functions Ξ, we follow Lütkepohl (2007). Therefore, consider the $K-$dimensional VAR(p) process

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + u_t, \tag{wa1}$$

with $y_t = (y_{1t}, \ldots, y_{Kt})'$ and the $(K \times K)$ coefficients matrices A_i and $K-$dimensional white noise with $E(u_t) = 0$ and

$$E(u_t u_s') = \begin{cases} \Sigma_u, & \text{if} \quad t = s, \\ 0, & \text{otherwise}, \end{cases} \tag{wa2}$$

with

$$\Sigma_u = \begin{bmatrix} \sigma_{11} & \ldots & \sigma_{1K} \\ \vdots & \ddots & \vdots \\ \sigma_{K1} & \ldots & \sigma_{KK} \end{bmatrix}. \tag{wa3}$$

Let vec denote the column stacking operator and $vech$ the corresponding operator that stacks only the elements on and below the diagonal. Then, the duplication operator D_k is such that for any $(K \times K)$-matrix T, $D_K vech(T) = vec(T)$ holds. Furthermore, in the spirit of Lütkepohl (2007) we define the following matrices

$$D_K^+ = (D_K' D_K)^{-1} D_K', \qquad \sigma = vech(\Sigma_u), \tag{wa4}$$

$$J = \begin{bmatrix} I_k & \ldots & 0 & 0 \end{bmatrix}, \qquad \alpha = vec(A_1, A_2, \ldots, A_p), \tag{wa5}$$

and

$$\Gamma = \left(\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ t_{t-p+1} \end{bmatrix} \begin{bmatrix} y_t', & \ldots & y_{t-p+1} \end{bmatrix} \right), \qquad \overline{\Sigma}_u = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \ldots & 0 \\ 0 & \sqrt{\sigma_{22}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \sqrt{\sigma_{KK}} \end{bmatrix}, \tag{wa6}$$

$$\mathbf{A} = \begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & I_K & 0 \end{bmatrix}. \tag{wa7}$$

Let $\hat{\alpha}$ and $\hat{\sigma}$ denote the least squares estimators with respect to (wa1). According to Lütkepohl (2007) these estimators have asymptotic covariances Σ_α and Σ_σ with

$$\Sigma_\alpha = \Gamma^{-1} \otimes \Sigma_u \qquad \Sigma_\sigma = 2D_K^{+}\Big(\Sigma_u \otimes \Sigma_u\Big)D_K^{+\prime}. \tag{wa8}$$

Furthermore, Lütkepohl (2007) derives the asymptotic properties of the orthogonalized impulse response functions Θ^o and Ξ^o. The orthogonalized impulse response arises from diagonalizing the residual covariance matrix Σ_u such that $\Sigma_u = PP'$ holds. With the definition $w_t = P^{-1}u_t$, w_t obeys $\Sigma_w = E[w_t w_t'] = I_K$. The corresponding MA representation of y_t can be written as $y_t = \sum_{i=0}^{\infty} \Theta^o(i) w_{t-i}$ with the *orthogonalized impulse response function* $\Theta^o(i)$, i.e.,

$$\Theta^o(i) = \Phi_i P \qquad \Xi^o(n) = \sum_{i=0}^{n} \Theta^o(i). \tag{wa9}$$

Comparing with the corresponding *generalized* impulse responses in (6.4) and (6.5), we have

$$\Theta^g(n) = \Phi_n Q \qquad \Xi^g(n) = \sum_{i=0}^{n} \Phi_i Q, \tag{wa10}$$

$$\Theta^o(n) = \Phi_n P \qquad \Xi^o(n) = \sum_{i=0}^{n} \Phi_i P, \tag{wa11}$$

with $Q = \Sigma_u \left(\overline{\Sigma}_u\right)^{-1}$. Observe that the only difference between orthogonalized and generalized impulse response lies in the right-multiplication of the matrices Q and P. Thus, we can use the analogy of the asymptotic properties of the orthogonalized impulse response as of Lütkepohl (2007) to derive the asymptotic properties for the generalized impulse response. Therefore, recall first the result with respect to orthogonalized impulse responses.

Theorem 1 (Lütkepohl (2007))**.** *Suppose*

$$\sqrt{T}\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\sigma} - \sigma \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \Sigma_\alpha & 0 \\ 0 & \Sigma_\sigma \end{bmatrix}\right). \tag{wa12}$$

Then

$$\sqrt{T} vec\Big(\hat{\Xi}^o(n) - \Xi^o(n)\Big) \xrightarrow{d} N\Big(0, B_n \Sigma_\alpha B_n' + \overline{B}_n \Sigma_\sigma \overline{B}_n'\Big), \qquad n = 1, 2 \ldots, \tag{wa13}$$

where the matrices B_n and $\overline{B}_n$ obey

$$B_n = \Big(P' \otimes I_K\Big)F_n, \qquad \overline{B}_n = (I_K \otimes \Psi_n)H, \tag{wa14}$$

(wa15)

with the F_n matrices obeying $F_n = \sum_{i=1}^{n} G_i$ and $G_i = \sum_{m=0}^{i-1} J(A')^{i-1-m} \otimes \Phi_m$ and the H matrix being defined as $H = \partial vec(P)/\partial \sigma'$. Moreover, $\Psi_j = \sum_{i=1}^{j} \Phi_i$.

With (wa10) and (wa11) in mind, it is easy to check that the asymptotic property of the corresponding cumulative generalized impulse response Ξ^g are derived similarly by replacing the matrix P with Q. Together with (wa8) and Theorem 1 we finally obtain:

Corollary WA1 (Asymptotic Distribution of Generalized Impulse Response)**.**

$$\sqrt{T}vec\Big(\hat{\Xi}^g(n) - \Xi^g(n)\Big) \xrightarrow{d} N\Big(0, B_n^g \Sigma_\alpha B_n^{g\prime} + \overline{B}_n^g \Sigma_\sigma \overline{B}_n^{g\prime}\Big), \qquad n = 1, 2 \dots \tag{wa16}$$

with F_n, G_n Ψ_n as in Theorem 1 and

$$B_n^g = \Big(Q' \otimes I_K\Big)F_n, \qquad \overline{B}_n^g = (I_K \otimes \Psi_n)H^g, \qquad H^g = \partial vec(Q)/\partial\sigma', \tag{wa17}$$

and

$$\Sigma_\alpha = \Gamma^{-1} \otimes \Sigma_u, \qquad \Sigma_\sigma = 2D_K^+\Big(\Sigma_u \otimes \Sigma_u\Big)D_K^{+\prime}. \tag{wa18}$$

Remark WA1 (Impulse Response for a Single Variable)**.** *From Corollary WA1, it is straightforward that the cumulative impulse response $\Xi_j^g(n)$ of the j-th endogenous variable at time n after the shock is obtained by right-multiplying Ξ_n^g with the column-vector e_i, which consists of zeros except at the jth entry. Thus, we have*

$$\sqrt{T}vec\Big(\hat{\Xi}_j^g(n) - \Xi_i^g(n)\Big) \xrightarrow{d} N\Big(0, \Lambda_{jn}\Big), \tag{wa19}$$

and

$$\Lambda_{jn} = e_j'\Big(B_n^g \Sigma_\alpha B_n^{g\prime} + \overline{B}_n^g \Sigma_\sigma \overline{B}_n^{g\prime}\Big)e_j. \tag{wa20}$$

References

LÜTKEPOHL, H. (2007): *New Introduction to Multiple Time Series Analysis*, Springer.