Supplementary material to: S. Clinet and Y. Potiron, "Disentangling sources of high frequency market microstructure noise"

We give the assumptions related to Proposition 1 and a detailed proof of the consistency of the BIC. First, defining

$$\chi(\theta) := \mathbb{E}\left[(\Delta\phi(Q_{t_1},\theta) - \Delta\phi(Q_{t_1},\theta_0))^2\right],$$

we assume

[A] For any $m \in \mathcal{M}$, χ admits a unique minimum $\tilde{\theta}^{(m)}$ on the interior of m.

Note that **[A]** is automatically satisfied for linear models such as (2.1) as soon as the variance-covariance matrix of the vector of returns of information ΔQ_{t_1} is positive definite.

We also define

$$W_i(\theta) := \phi(Q_i,\theta) - \phi(Q_i,\theta_0),$$

and for any $i,j,k,l \in \mathbb{N}$, and for any multi-indices $\boldsymbol{q} = (q_1,q_2)$, $\boldsymbol{r} = (r_1,r_2,r_3,r_4)$, where the subcomponents of $\boldsymbol{q}$ and $\boldsymbol{r}$ are d dimensional multi-indices, the following quantities conditioned on the price process

$$\begin{aligned}
\mathbb{E}\left[W_i(\theta)\middle| X\right] &= 0 \text{ a.s},\\
\rho_j^{\boldsymbol{q}}(\theta) := \mathbb{E}\left[\frac{\partial^{q_1}W_i(\theta)}{\partial\theta^{q_1}}\frac{\partial^{q_2}W_{i+j}(\theta)}{\partial\theta^{q_2}}\middle| X\right] &= \mathbb{E}\left[\frac{\partial^{q_1}W_i(\theta)}{\partial\theta^{q_1}}\frac{\partial^{q_2}W_{i+j}(\theta)}{\partial\theta^{q_2}}\right] \text{ a.s},\\
\kappa_{j,k,l}^{\boldsymbol{r}}(\theta) := \operatorname{cum}\left[\frac{\partial^{r_1}W_i(\theta)}{\partial\theta^{r_1}},\frac{\partial^{r_2}W_{i+j}(\theta)}{\partial\theta^{r_2}},\frac{\partial^{r_3}W_{i+k}(\theta)}{\partial\theta^{r_3}},\frac{\partial^{r_4}W_{i+l}(\theta)}{\partial\theta^{r_4}}\middle| X\right]&\\
&= \operatorname{cum}\left[\frac{\partial^{r_1}W_i(\theta)}{\partial\theta^{r_1}},\frac{\partial^{r_2}W_{i+j}(\theta)}{\partial\theta^{r_2}},\frac{\partial^{r_3}W_{i+k}(\theta)}{\partial\theta^{r_3}},\frac{\partial^{r_4}W_{i+l}(\theta)}{\partial\theta^{r_4}}\right] \text{ a.s},
\end{aligned}$$

where $\rho_j^{\boldsymbol{q}}(\theta)$ and $\kappa_{j,k,l}^{\boldsymbol{r}}(\theta)$ are assumed independent of n. The following assumption is directly taken from [Clinet and Potiron, 2019]:

[B] The impact function ϕ is supposed to be of class C^m in θ with $m > \bar{d}/2 + 2$. Moreover, for any $i = 0,\cdots,m$ and $0 \leq |\boldsymbol{q}|,|\boldsymbol{r}| \leq m$, we have

$$\begin{aligned}
\sup_{\theta\in\Theta}\sum_{j=0}^{+\infty}\left|\rho_j^{\boldsymbol{q}}(\theta)\right| &< \infty \text{ a.s},\\
\sup_{\theta\in\Theta}\sum_{j,k,l=0}^{+\infty}\left|\kappa_{j,k,l}^{\boldsymbol{r}}(\theta)\right| &< \infty \text{ a.s},\\
\mathbb{E}\left[\sup_{\theta\in\Theta}\left|\frac{\partial^j\Delta\phi(Q_{t_i},\theta)}{\partial\theta^j}\right|^p\middle| X\right] &< \infty \text{ a.s, for any } p\geq 1,\ 0\leq j\leq 2,\\
\frac{\partial\rho_0(\theta)}{\partial\theta} = 0 &\Leftrightarrow \theta = \theta_0.
\end{aligned}$$

We are now ready to prove Proposition 1.

Proof of Proposition 1. All we have to do is to show that for any $m \neq m_0$, $\text{BIC}(m) - \text{BIC}(m_0) \to^{\mathbb{P}} +\infty$.
Step 1. We prove our claim when m_0 is a submodel of m, and so $d > d_0$ where d is the number of parameters of m. By Theorem 3.1 from [Clinet and Potiron, 2019], and up to some reordering of the subcomponents of θ, the estimator $\widehat{\upsilon}^{(m)}$ is consistent and asymptotically normal, toward the limit $\upsilon_0 = (\overline{\sigma}_0^2, \theta_0^1, \cdots, \theta_0^{d_0}, 0, \cdots, 0)$ where $\overline{\sigma}_0^2 = \int_0^T \sigma_s^2 ds + \sum_{0<s\leq T} \Delta J_s^2$. We slightly reformulate the problem as follows: introducing $\widehat{w}^{(m)} = ((\widehat{\sigma}^2)^{(m)}, N^{1/2}(\widehat{\theta}^{(m)} - \theta_0))$, and $w_0 = (\overline{\sigma}_0^2, 0, \cdots, 0)$, we have by a Taylor expansion, for some $\overline{w} \in [\widehat{w}^{(m)}, w_0]$

$$\begin{aligned} 2(l_{exp}^{(m)}(\widehat{\upsilon}^{(m)}) - l_{exp}^{(m)}(\upsilon_0)) &= 2(\mathcal{L}_{exp}^{(m)}(\widehat{w}^{(m)}) - \mathcal{L}_{exp}^{(m)}(w_0)) \\ &= -N(\widehat{w}^{(m)} - w_0)^T H_{exp}^{(m)}(\overline{w})(\widehat{w}^{(m)} - w_0) \to^d \chi^2(d) \end{aligned}$$

by application of Theorem 3.1 and Lemma C.15 from [Clinet and Potiron, 2019], and with $\mathcal{L}_{exp}^{(d)}$ being the restriction of $\mathcal{L}_{exp}$ on m, where $\mathcal{L}_{exp}$ is defined in (C.88), p.323 of [Clinet and Potiron, 2019], and $H_{exp}^{(m)} = -N\partial^2\mathcal{L}_{exp}^{(m)}/\partial w^2$. In the previous equation, $\to^d \chi^2(d)$ stands for the convergence in law toward a chi-squared distribution with d degrees of freedom. We have a similar result for m_0, and thus $\mathcal{L}_{exp}^{(m)}(\widehat{\upsilon}^{(m)}) - \mathcal{L}_{exp}^{(m_0)}(\widehat{\upsilon}^{(m_0)}) = O_{\mathbb{P}}(1)$. This, in turn, implies that $\text{BIC}(m) - \text{BIC}(m_0) \sim (d-d_0)\log(N) \to^{\mathbb{P}} +\infty$.
Step 2. We prove our claim when m_0 is not a submodel of m. We recall that, by definition of the likelihood process, we have

$$\widehat{\theta}^{(m)} \in \text{argmin}_{\theta\in m} \Delta\widetilde{Z}(\theta)^T \Delta\widetilde{Z}(\theta),$$

and

$$(\widehat{\sigma}^2)^{(m)} = T^{-1}\Delta\widetilde{Z}(\widehat{\theta}^{(m)})^T \Delta\widetilde{Z}(\widehat{\theta}^{(m)}),$$

with $\widetilde{Z}_{t_i}(\theta) = Z_{t_i} - \phi(Q_{t_i}, \theta)$. By direct calculation similar to that of Section C.4 from [Clinet and Potiron, 2019], we have the uniform convergence for $\theta \in m$

$$N^{-1}\Delta\widetilde{Z}(\theta)^T \Delta\widetilde{Z}(\theta) \to^{\mathbb{P}} \chi(\theta) = \mathbb{E}\left[(\Delta\phi(Q_{t_1}, \theta) - \Delta\phi(Q_{t_1}, \theta_0))^2\right].$$

As a direct consequence, we obtain that $\widehat{\theta}^{(m)} \to^{\mathbb{P}} \tilde{\theta}^{(m)}$ where we recall that $\tilde{\theta}^{(m)}$ is the unique minimum of χ on the interior of m by Assumption **[A]**. Similarly, we easily obtain that

$$(\widehat{\sigma}^2)^{(m)} = \Delta_N^{-1}\chi\left(\tilde{\theta}^{(m)}\right) + o_{\mathbb{P}}\left(\Delta_N^{-1}\right),$$

where $\Delta_N = T/N$, and where $\chi(\tilde{\theta}^{(m)}) > 0$ by the identifiability assumption (2.14) from [Clinet and Potiron, 2019] along with the fact that $\tilde{\theta}^{(m)} \neq \theta_0$. Moreover, we have

$$\frac{\partial^2 l_{exp}^{(m)}(\widehat{\upsilon}^{(m)})}{\partial \upsilon^2} = \begin{pmatrix} \frac{-T}{2(\widehat{\sigma}^4)^{(m)}\Delta_N} & 0 \\ 0 & \frac{-1}{2(\widehat{\sigma}^2)^{(m)}\Delta_N} \frac{\partial^2\left(\Delta\widetilde{Z}(\widehat{\theta}^{(m)})^T\Delta\widetilde{Z}(\widehat{\theta}^{(m)})\right)}{\partial\theta^2} \end{pmatrix},$$

and therefore by a Taylor expansion at $\widehat{\upsilon}^{(m)}$, we get for some $\overline{\upsilon} \in [\upsilon_0, \widehat{\upsilon}^{(m)}]$

$$\begin{aligned} 2(l_{exp}^{(m)}(\widehat{\upsilon}^{(m)}) - l_{exp}^{(m)}(\upsilon_0)) &= (\widehat{\upsilon}^{(d)} - \upsilon_0)^T \frac{\partial^2 l_{exp}^{(m)}(\overline{\upsilon})}{\partial \upsilon^2} (\widehat{\upsilon}^{(m)} - \upsilon_0) \\ &= \frac{-T\left((\widehat{\sigma}^2)^{(m)} - \overline{\sigma}_0^2\right)^2}{2(\widehat{\sigma}^4)^{(m)}\Delta_N} - \frac{\left(\widehat{\theta}^{(m)} - \theta_0\right)^T}{2(\widehat{\sigma}^2)^{(m)}\Delta_N} \frac{\partial^2\left(\Delta\widetilde{Z}(\widehat{\theta}^{(m)})^T \Delta\widetilde{Z}(\widehat{\theta}^{(m)})\right)}{\partial\theta^2}\left(\widehat{\theta}^{(m)} - \theta_0\right) \\ &= -\frac{T\Delta_N^{-1}}{2} - T\Delta_N^{-1}\frac{\left(\widehat{\theta}^{(m)} - \theta_0\right)^T}{2\chi(\tilde{\theta}^{(m)})}\frac{\partial^2\chi(\tilde{\theta}^{(m)})}{\partial\theta^2}\left(\widehat{\theta}^{(m)} - \theta_0\right) + o_{\mathbb{P}}\left(\Delta_N^{-1}\right). \end{aligned}$$

Now, since $\tilde{\theta}^{(m)}$ is the unique minimum of χ on the interior of m, we deduce that $\frac{\partial^2\chi(\tilde{\theta}^{(m)})}{\partial\theta^2}$ is a positive matrix and thus $-T\Delta_N^{-1}\frac{\left(\widehat{\theta}^{(m)} - \theta_0\right)^T}{2\chi(\tilde{\theta}^{(m)})}\frac{\partial^2\chi(\tilde{\theta}^{(m)})}{\partial\theta^2}\left(\widehat{\theta}^{(m)} - \theta_0\right) \leq 0$. Therefore, we have $2(l_{exp}^{(m)}(\widehat{\upsilon}^{(m)}) - l_{exp}^{(m)}(\upsilon_0)) \leq -\frac{T}{2}\Delta_N^{-1} + o_{\mathbb{P}}(\Delta_N^{-1})$. Thus,

$$\text{BIC}(m) - \text{BIC}(m_0) \geq \frac{N}{2} + \underbrace{(d - d_0)\log(N)}_{o_{\mathbb{P}}(N)} + o_{\mathbb{P}}(N),$$

which proves our claim.

□

References

[Clinet and Potiron, 2019] Clinet, S. and Potiron, Y. (2019). Testing if the market microstructure noise is fully explained by the informational content of some variables from the limit order book. *Journal of Econometrics*.