Practical Guidelines for the Estimation and Inference of a Dynamic Logistic Model with Fixed-Effects
Supplementary Material

Romain Aeberhardt* Laurent Davezies§

December 7, 2010

Theorem 2.1 (detailed assumptions and extended proof).
(A1) $\{(y_{it})_{t=1..T}, (x_{mit})_{m=1..M,t=1..T}\}_{i=1}^{n}$ *is a random sample of* n *observations from a distribution satisfying Equation (1) in our main paper*
(A2) $\theta_0 \in int(\Theta)$, *with* Θ *a compact subset of* $\mathbb{R}^{(k+1)(M-1)}$
(A3) $\forall 2 \leq t < s \leq T-1$, $x_{it+1} - x_{is+1}$ *is absolutely continuously distributed with a density* $f_{t,s}$, *bounded above on its support, and strictly positive at zero, twice differentiable on its support, with bounded derivatives.* $\forall 2 \leq t < s \leq T-1$, $\forall 2 \leq t' < s' \leq T-1$ *such that* $(t,s) \neq (t',s')$, $(x_{it+1} - x_{is+1}, x_{it'+1} - x_{is'+1})$ *is absolutely continuously distributed with a density* $f_{t,s,t',s'}$, *bounded above on its support, and strictly positive at zero, twice differentiable on its support, with bounded derivatives.*
(A4) $\forall 2 \leq t < s \leq T-1$, $E\left[||x_{it} - x_{is}||^6 | x_{it+1} - x_{is+1}\right]$ *is bounded on its support.*
(A5) Let $h_{itsml}(\theta) = h^{(0)}_{itsml}(\theta) = \mathbb{1}\{\{y_{it}, y_{is}\} = \{m,l\}\} \ln\left(\frac{\exp(\mathbb{1}\{y_{it}=m\} Z_{i,t,s,m,l}\theta)}{1+\exp(Z_{i,t,s,m,l}\theta)}\right)$, $h^{(1)}_{itsml}(\theta) = \frac{\partial h_{itsml}(\theta)}{\partial \theta}$ *and* $h^{(2)}_{itsml}(\theta) = \frac{\partial^2 h_{itsml}(\theta)}{\partial \theta \partial \theta'}$.

- $\forall 2 \leq t < s \leq T-1$, $\forall m, l \in [1..M]$,
  - $E\left[h^{(j)}_{itsml}(\theta) | x_{it+1} - x_{is+1}\right]$ *is twice differentiable in a neighborhood of zero for all* $\theta \in \Theta$ *for* $j = 0,1,2$.
  - $E\left[h^{(j)}_{itsml}(\theta) h^{(j)}_{itsml}(\theta)' | x_{it+1} - x_{is+1}\right]$ *is twice differentiable in a neighborhood of zero for all* $\theta \in \Theta$ *for* $j = 0,1$.
- $\forall 2 \leq t < s \leq T-1$, $\forall 2 \leq t' < s' \leq T-1$ *such that* $(t,s) \neq (t',s')$, $\forall m, l, m', l' \in [1..M]$, $E\left[h^{(j)}_{itsml}(\theta) h^{(j)}_{it's'm'l'}(\theta)' | x_{it+1} - x_{is+1}, x_{it'+1} - x_{is'+1}\right]$ *is twice differentiable in a neighborhood of* $(0,0)$ *for all* $\theta \in \Theta$ *for* $j = 0,1$.

(A6) $\forall 2 \leq t < s \leq T-1$, $\forall m, l \in [1..M]$, $E\left[(x_{it} - x_{is})'(x_{it} - x_{is}) | x_{it+1} - x_{is+1}\right]$ *has full rank* k *in a neighborhood of zero.*
(A7) $K : \mathbb{R}^k \to \mathbb{R}$ *is a bounded and symmetric kernel such that* $\int K(u)du = 1$.

*CREST (INSEE), romain.aeberhardt@travail.gouv.fr
§CREST, laurent.davezies@ensae.fr, (corresponding author)

(A8) $\sqrt{n}\sigma_n^{2+k/2} \to \sigma \in \mathbb{R}^+$.

If (A1)-(A8) hold, then

$$\sqrt{n}\sigma_n^{k/2}\left(\widehat{\theta}_n - \theta_0\right) \to \mathcal{N}\left(B, J^{-1}VJ^{-1}\right)$$

With $B = O\left(\sqrt{n}\sigma_n^{2+k/2}\right)$, *J and V being consistently estimated by*

$$\widehat{J}_n = -\frac{1}{n\sigma_n^k}\sum_{i=1}^n \sum_{\substack{2\le t<s\le T-1\\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h^{(2)}_{itsml}(\widehat{\theta}_n) \text{ and}$$

$$\widehat{V}_n = \frac{1}{n\sigma_n^k}\sum_{i=1}^n \left[\sum_{\substack{2\le t<s\le T-1\\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h^{(1)}_{itsml}(\widehat{\theta}_n)\right]\left[\sum_{\substack{2\le t<s\le T-1\\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h^{(1)}_{itsml}(\widehat{\theta}_n)'\right]$$

or

$$\widetilde{V}_n = \frac{1}{n\sigma_n^k}\sum_{i=1}^n \left[\sum_{\substack{2\le t<s\le T-1\\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right)^2 h^{(1)}_{itsml}(\widehat{\theta}_n) h^{(1)}_{itsml}(\widehat{\theta}_n)'\right]$$

The proof is very close to the proofs of theorems 1 to 3, of Honoré and Kyriazidou (Honoré and Kyriazidou, 2000) :

We will use some useful results on the kernel estimators and on the uniform convergence in probability:

- (Kernel estimators) If Z is a random variable iid across individuals,
  - If $E\left(Z|x_{it+1}-x_{is+1}=x\right)$ exists and is twice differentiable in a neighborhood of $x=0$, then:

    $$\forall \nu \ge 0,\quad E\left(\frac{1}{\sigma_n^k}K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right)^{\nu} Z_i\right) = f_{t+1,s+1}(0)E\left(Z|x_{it+1}=x_{is+1}\right)\int K(u)^{\nu}du + O\left(\sigma_n^2\right)$$

  - If $E\left(Z|x_{it+1}-x_{is+1}=x; x_{it'+1}-x_{is'+1}=x'\right)$ exists and is twice differentiable in a neighborhood of $(x,x')=(0,0)$ for $(t,s)\neq(t',s')$, then:

    $$E\left(\frac{1}{\sigma_n^{2k}}K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right)K\left(\frac{x_{it'+1}-x_{is'+1}}{\sigma_n}\right)Z_i\right) =$$

    $$f_{t+1,s+1,t'+1,s'+1}(0,0)E\left(Z_i|x_{it+1}=x_{is+1}; x_{it'+1}=x_{is'+1}\right) + O\left(\sigma_n^2\right)$$

- (Corollary 2.2, Newey (1991)) If $\mu_n(\theta)$ is a sequence of random differentiable function such that for all $\theta\in\Theta$, $\mu_n(\theta) \underset{P}{\to} \mu(\theta)$, and if the derivative of $\mu_n(\theta)$ are dominated by a random variable U_n such that $U_n = O_p(1)$ and $E(U_n)<\infty$, the convergence in probability is uniform on the compact Θ. For $j=0,1,2$, note that the sequence of random function $\frac{1}{n\sigma_n^k}\sum_{i=1}^n\sum_{\substack{2\le t<s\le T-1\\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right)h^{(j)}_{itsml}(\theta)$ verify the condition of domination.

Let $m_i(\sigma_n,\theta) = \sum_{2\le t<s\le T-1, m\neq l} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h_{itsml}(\theta)$ and $M_n(\theta) = \frac{1}{n\sigma_n^k}\sum_{i=1}^n m_i(\sigma_n,\theta)$.
We have : $E\left[M_n(\theta)\right] \to \sum_{2\le t<s\le T-1, m\neq l} f_{t+1,s+1}(0)E\left[h_{itsml}(\theta)|x_{i,t+1}=x_{i,s+1}\right] = M(\theta)$

To prove consistency we use Theorem 5.7 of Van Der Vaart (1998), the first assumption we need to verify is a stochastic uniform convergence $\sup_{\theta\in\Theta}|M_n(\theta)-M(\theta)| \xrightarrow{P} 0$.

Because

$$\begin{array}{rcl} V\left(M_n\left(\theta\right)\right) & = & \frac{1}{n\sigma_n^{2k}} V\left(m_i(\sigma_n,\theta)\right) \\ & \leq & \frac{1}{n}\sum_{\substack{2\leq s<t\leq T-1 \\ 2\leq s'<t'\leq T-1 \\ (s,t)\neq(s',t')}} \sum_{\substack{m\neq l \\ m'\neq l'}} \frac{1}{\sigma_n^{2k}} E\left(K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) K\left(\frac{x_{it'+1}-x_{is'+1}}{\sigma_n}\right) h_{itsml}(\theta) h_{it's'm'l'}(\theta)\right) \\ & \leq & \frac{1}{n}\left(O(\frac{1}{\sigma_n^k}) + O(1)\right) = O\left(\frac{1}{n\sigma_n^k}\right) = o(1) \end{array}$$

$M_n(\theta) \overset{L_2}{\longrightarrow} M(\theta)$ for all $\theta \in \Theta$, so $M_n(\theta) \overset{P}{\longrightarrow} M(\theta)$ for all $\theta \in \Theta$. Uniformity of the convergence on Θ is ensured by domination.

The second assumption of Theorem 5.7 of Van Der Vaart (1998) is that θ_0 is a well-separated point of maximum of M : $\sup_{\theta:d(\theta,\theta_0)\geq\varepsilon} M(\theta) < M(\theta_0)$

$$M(\theta) = \sum_{2\leq t<s\leq T-1, m\neq l} f_{t+1,s+1}(0) P\left(\{y_{it}, y_{is}\} = \{l,m\} | x_{t+1} = x_{s+1}\right) g_{tsml}(\theta)$$

with $g_{tsml}(\theta) = E\left(ln\left(\frac{\exp(y_{mit}Z\theta)}{1+\exp(Z\theta)}\right) | x_{t+1} = x_{s+1}; \{y_{it}, y_{is}\} = \{l,m\}\right)$.

If $s = t+1$, $\theta \mapsto E\left(ln\left(\frac{\exp(y_{mit}Z\theta)}{1+\exp(Z\theta)}\right) | x_{t+1} = x_{s+1}; \{y_{it}, y_{is}\} = \{l,m\}; (y_{i\tau})_{\tau\neq t,s}\right)$ is well-separated for the component $\beta_m, \beta_l, \gamma_{y_{i,t-1}m}, \gamma_{y_{i,t-1}l}, \gamma_{my_{i,s+1}}, \gamma_{ly_{i,s+1}}, \gamma_{lm}, \gamma_{ml}$ and does not depend on other component of θ. It is well-separated for the component β_m, β_l, $\gamma_{y_{i,t-1}m}$, $\gamma_{y_{i,t-1}l}$, $\gamma_{my_{i,s+1}}$, $\gamma_{ly_{i,s+1}}$, $\gamma_{my_{i,t-1}}$, $\gamma_{ly_{i,t-1}}$, $\gamma_{y_{i,s+1}m}$, $\gamma_{y_{i,s+1}l}$ and does not depend on other component of θ if $s > t+1$. Then, $g_{tsml}(\theta)$ is well-separated for the component β_m, β_l and $(\gamma_{qm}, \gamma_{ql}, \gamma_{mq}, \gamma_{lq})_{q\in[1,M]}$ and does not depend on the other components. Because $f_{t+1,s+1}(0)P\left(\{y_{it}, y_{is}\} = \{l,m\} | x_{t+1} = x_{s+1}\right)$ are positive quantities for every 4-uplet s, t, l, m, $M(\theta)$ is well-separated for θ.

Theorem 5.7 of Van Der Vaart (1998) implies the consistency of the estimate.

Now let's focus on the asymptotic normality. For that, we use Taylor expansion of the first order condition :

$$\begin{array}{rcl} 0 & = & \frac{1}{\sqrt{n\sigma_n^k}} \sum_{i=1}^n \left\{\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta} - E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}\right)\right\} \\ & & + \frac{1}{\sqrt{n\sigma_n^k}} \sum_{i=1}^n E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}\right) \\ & & + \frac{1}{\sqrt{n\sigma_n^k}} \sum_{i=1}^n \frac{\partial^2 m_i(\sigma_n,\theta_n^\star)}{\partial\theta\partial\theta'} (\widehat{\theta}_n - \theta_0) \end{array}$$

Let $\xi_{in} = \frac{1}{\sqrt{n\sigma_n^k}} \frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}$. We use the Lindeberg-Feller central limit theorem (see Van Der Vaart (1998), proposition 2.27), to show that $\sum_{i=1}^n \left(\xi_{in} - E\left(\xi_{in}\right)\right)$ converge in distribution to $\mathcal{N}(0, V)$.

$$\begin{array}{rcl} \sum_{i=1}^n Cov(\xi_{in}) & = & \frac{1}{\sigma_n^k} E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta} \frac{\partial m_i'(\sigma_n,\theta_0)}{\partial\theta'}\right) - \frac{1}{\sigma_n^k} E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}\right) E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}\right)' \\ & = & \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} f_{t+1,s+1}(0) E\left(h_{itsml}^{(1)}(\theta_0) h_{itsml}^{(1)'}(\theta_0) | x_{it+1} = x_{is+1}\right) \int K^2(u)du \\ & & + \sigma_n^k \sum_{\substack{2\leq s<t\leq T-1 \\ 2\leq s'<t'\leq T-1 \\ (s,t)\neq(s',t')}} \sum_{\substack{m\neq l \\ m'\neq l'}} f_{(t+1,s+1,t'+1,s'+1)}(0,0) E\left(h_{itsml}^{(1)}(\theta_0) h_{it's'm'l'}^{(1)'}(\theta_0) | x_{it+1} = x_{is+1}; x_{it'+1} = x_{is'+1}\right) \\ & & + O(\sigma_n^2) + O(\sigma_n^{k+2}) + O(\sigma_n^{k+4}) \\ & = & \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} f_{t+1,s+1}(0) E\left(h_{itsml}^{(1)}(\theta_0) h_{itsml}^{(1)'}(\theta_0) | x_{it+1} = x_{is+1}\right) \int K^2(u)du + O(\sigma_n^{\min(2,k)}) \end{array}$$

$$\begin{aligned}
\sum_{i=1}^{n} E(||\xi_{in}||^2 \mathbb{1}_{\{||\xi_{in}||>\varepsilon\}}) &\leq n\varepsilon^{-\delta} E\left[||\xi_{in}||^{2+\delta}\right] && \text{(Markov)} \\
&\leq \frac{n^{-\delta/2}\varepsilon^{-\delta}}{(\sigma_n^k)^{(2+\delta)/2}} \left(\frac{(T-2)(T-3)}{2}\right)^{1+\delta} \\
& E\left[\sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right)^{2+\delta} ||h_{itsml}^{(1)}(\theta_0)||^{2+\delta}\right] && \text{(Hölder)} \\
&\leq O\left(\frac{1}{\sqrt{n\sigma_n^k}^{\delta}}\right) = o(1)
\end{aligned}$$

It follows from the Lindeberg-Feller theorem (see, for instance Van Der Vaart (1998), chapter 2), that $\sum_{i=1}^{n} \xi_{in} - E(\xi_{in}) \to \mathcal{N}(0, V)$ with

$$V = \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} f_{t+1,s+1}(0) E\left(h_{itsml}^{(1)}(\theta_0) h_{itsml}^{(1)'}(\theta_0) | x_{it} = x_{is}\right) \int K^2(u)du.$$

Let $b_n = \frac{1}{n\sigma_n^k} \sum_{i=1}^{n} E\left(\frac{\partial m_i(\sigma_n,\theta_0)}{\partial\theta}\right)$, we have : $b_n = O\left(\sigma_n^2\right)$

And then $\sqrt{n\sigma_n^k} b_n = o_p(1)$.

Let $J_n(\theta) = \frac{1}{n\sigma_n^k} \sum_{i=1}^{n} \frac{\partial^2 m_i(\sigma_n,\theta)}{\partial\theta\partial\theta'}$. For all $\theta \in \Theta$, we have :

$$\begin{aligned}
E(J_n(\theta)) &= \frac{1}{\sigma_n^k} E\left(\sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h_{itsml}^{(2)}(\theta)\right) \\
&= \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} f_{s+1,t+1}(0) E\left[h_{itsml}^{(2)}(\theta) | x_{it+1} = x_{is+1}\right] + o(1) \\
&= J(\theta) + o(1)
\end{aligned}$$

The variance of the jj''th component of $J_n(\theta)$ decrease to 0.

$$\begin{aligned}
Var(J_n(\theta)(j,l)) &\leq \frac{1}{n} E\left[\left(\frac{1}{\sigma_n^k} \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h_{itsml}^{(2)}(\theta)(j,j')\right)^2\right] \\
&\leq \frac{1}{n\sigma_n^k} \sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} f_{s+1,t+1}(0) E\left(h_{itsml}^{(2)}(\theta)(j,j')^2 | x_{it+1} = x_{is+1}\right) \int K^2(u)du \\
&+ \frac{1}{n} \sum_{\substack{2\leq s<t\leq T-1 \\ 2\leq s'<t'\leq T-1 \\ (s,t)\neq(s',t')}} \sum_{\substack{m\neq l \\ m'\neq l'}} f_{(s+1,t+1,s'+1,t'+1)}(0,0) E\left(h_{itsml}^{(2)}(j,j') h_{it's'm'l'}^{(2)}(j,j') | x_{it} = x_{is}; x_{it'} = x_{is'}\right) \\
&+ O\left(\frac{\sigma_n^2}{n\sigma_n^k}\right) + O\left(\frac{\sigma_n^2}{n}\right) \\
&= O\left(\frac{1}{n\sigma_n^k}\right) = o(1)
\end{aligned}$$

So for all $\theta \in \Theta$ $J_n(\theta) = J(\theta) + o_p(1)$. Using the second part of the preliminary remark, the convergence is uniform for $\theta \in \Theta$. We deduce that $\widehat{J}_n = J_n\left(\widehat{\theta}_n\right) = J(\theta_0) + o_p(1) = J + o_p(1)$

Similarly, if we note

$$V_n(\theta) = \frac{1}{n\sigma_n^k} \sum_{i=1}^{n} \left[\sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h_{itsml}^{(1)}(\theta)\right] \left[\sum_{\substack{2\leq t<s\leq T-1 \\ m\neq l}} K\left(\frac{x_{it+1}-x_{is+1}}{\sigma_n}\right) h_{itsml}^{(1)}(\theta)'\right]$$

or

$$V_n(\theta) = \frac{1}{n\sigma_n^k} \sum_{i=1}^{n} \left[\sum_{\substack{2 \le t < s \le T-1 \\ m \ne l}} K\left(\frac{x_{it+1} - x_{is+1}}{\sigma_n}\right)^2 h_{itsml}^{(1)}(\theta) h_{itsml}^{(1)}(\theta)' \right]$$

Then $V_n(\theta)$ converges uniformly on Θ to

$$V(\theta) = \sum_{\substack{2 \le t < s \le T-1 \\ m \ne l}} f_{t+1,s+1}(0) E\left(h_{itsml}^{(1)}(\theta) h_{itsml}^{(1)'}(\theta) | x_{it} = x_{is}\right) \int K^2(u) du$$

And we conclude that $V_n\left(\widehat{\theta}_n\right) = V(\theta_0) + o_p(1) = V + o_p(1)$

References

HONORÉ, B., AND E. KYRIAZIDOU (2000): "Panel Data Discrete Choice Models with Lagged Dependent Variables," *Econometrica*, 68(4), 839–874.

NEWEY, W. K. (1991): "Uniform Convergence in Probability and Stochastic Equicontinuity," *Econometrica*, 59(4), 1161–1167.

VAN DER VAART, A. W. (1998): *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge Unversity Press.