SUPPLEMENT TO “COMMENTS ON ‘CONVERGENCE PROPERTIES
OF THE LIKELIHOOD OF COMPUTED DYNAMIC MODELS’”

BY DANIEL ACKERBERG, JOHN GEWEKE, AND JINYONG HAHN

APPENDIX A: PROOFS OF (3), (4), AND (5)

DEFINITION 2: Let \( c \neq 0 \), and define

\[
\chi(c) = \frac{1}{\sigma^2 \sqrt{2\pi}} \max_z \left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z-c)^2}{2}\right)}{\frac{c}{\sigma}} \right|.
\]

REMARK 1: In the definition above, it was implicitly assumed that \( \max_z |\exp(-z^2/2) - \exp(-(z - \frac{c}{\sigma})^2/2)/\frac{c}{\sigma}| \) is well defined. To confirm that it indeed is, define

\[
\varphi(z) = \left| \exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z-c)^2}{2}\right)/\frac{c}{\sigma} \right|.
\]

Note that (i) \( \varphi(z) \to 0 \) as \( |z| \to \infty \) and (ii) \( \varphi(0) > 0 \). Therefore, we can find \( B > 0 \) sufficiently large that \( \varphi(z) < \varphi(0)/2 \) for all \( |z| > B \). Now, over the compact set \( \mathbb{B} = \{ z : |z| \leq B \} \), the function is continuous and, therefore, there is some \( z^* \) at which the function \( \varphi(\cdot) \) is maximized over \( \mathbb{B} \). In other words,

\[
\varphi(z^*) \geq \varphi(z) \quad \forall z \in \mathbb{B}.
\]

Because \( 0 \in \mathbb{B} \), we should have \( \varphi(z^*) \geq \varphi(0) \). But, for all \( z \notin \mathbb{B} \), we have \( \varphi(z) < \varphi(0)/2 < \varphi(0) \leq \varphi(z^*) \). In other words,

\[
\varphi(z^*) \geq \varphi(z) \quad \forall z \notin \mathbb{B}.
\]

Combining (8) and (9), we conclude that \( \varphi(z^*) \geq \varphi(z) \) for all \( z \). In other words, the maximum is attained.

Given the definition, we can write

\[
|p(y; \gamma) - p_j(y; \gamma)| = \frac{1}{\sigma^2 \sqrt{2\pi}} \left| \exp\left(-\frac{(y_i - \gamma)^2}{2\sigma^2}\right) - \exp\left(-\frac{(y_i - \delta - \gamma)^2}{2\sigma^2}\right) \right|
\]
where \( \frac{y_i}{\sigma} - \frac{z}{\sigma} \) is interpreted as the \( z \) in the definition of \( \chi(c) \). Note that this bound is sharp by the definition of \( \chi(\cdot) \). In other words, there is a value of \( y_i \) (or analogously \( y_i/\sigma - \gamma/\sigma \)) such that the bound holds with equality.

**Lemma 1:** We have

\[
\frac{1}{\sigma \sqrt{2\pi}} \leq \lim \inf_{|c| \to \infty} |c| \chi(c) \leq \lim \sup_{|c| \to \infty} |c| \chi(c) \leq \frac{2}{\sigma \sqrt{2\pi}}.
\]

**Proof:** By definition,

\[
|c| \chi(c) = \frac{|c|}{\sigma^2 \sqrt{2\pi}} \max_z \left| \exp \left( -\frac{z^2}{2} \right) - \exp \left( -\frac{(z - c/\sigma)^2}{2} \right) \right|
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \max_z \left| \exp \left( -\frac{z^2}{2} \right) - \exp \left( -\frac{(z - c/\sigma)^2}{2} \right) \right|
\]

\[
\leq \frac{1}{\sigma \sqrt{2\pi}} \left[ \max_z \left| \exp \left( -\frac{z^2}{2} \right) \right| + \max_z \left| \exp \left( -\frac{(z - c/\sigma)^2}{2} \right) \right| \right]
\]

\[
\leq \frac{2}{\sigma \sqrt{2\pi}},
\]

from which we obtain

\[
\lim \sup_{|c| \to \infty} |c| \chi(c) \leq \frac{2}{\sigma \sqrt{2\pi}}.
\]
Next, note that

\[
|c| \chi(c) = \frac{|c|}{\sigma^2 \sqrt{2\pi}} \max_z \left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right)}{\frac{c}{\sigma}} \right|
\]

\[
\geq \frac{1}{\sigma \sqrt{2\pi}} \left| \exp\left(-\frac{0^2}{2}\right) - \exp\left(-\frac{(0 - c)^2}{2}\right) \right|
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \left| 1 - \exp\left(-\frac{(c)^2}{2}\right) \right|
\]

from which we obtain

\[
\frac{1}{\sigma \sqrt{2\pi}} \leq \lim \inf |c| \chi(c).
\]

Q.E.D.

**Lemma 2:** We have

\[
\chi(\delta) \leq \frac{\exp\left(-\frac{1}{2}\right)}{\sigma \sqrt{2\pi}}.
\]

**Proof:** By the mean value theorem, we have

\[
\left| \exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right) \right| = |c| \left| \exp\left(-\frac{(z - c)^2}{2}\right)(z - c^*) \right|
\]

where \(c^*\) is on the line segment adjoining 0 and \(c\). Note that the function \(s \mapsto |\exp(-s^2/2)|\) is bounded by \(\exp(-\frac{1}{2})\) (it is maximized at \(s = 1\)). It follows that

\[
\left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right)}{c} \right| \leq \exp\left(-\frac{1}{2}\right),
\]

from which we obtain

\[
\max_z \left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right)}{c} \right| \leq \exp\left(-\frac{1}{2}\right).
\]
It follows that

\[
\chi(c) = \frac{1}{\sigma^2 \sqrt{2\pi}} \max_z \left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right)}{c} \right|
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \max_z \left| \frac{\exp\left(-\frac{z^2}{2}\right) - \exp\left(-\frac{(z - c)^2}{2}\right)}{c} \right|
\]

\[
\leq \frac{\exp\left(-\frac{1}{2}\right)}{\sigma \sqrt{2\pi}}. \quad Q.E.D.
\]

**APPENDIX B: PROOF OF (6)**

For the joint likelihood, we have

\[
\prod_{t=1}^{T} p(y_t; \gamma) = \left(\frac{1}{\sigma \sqrt{2\pi}}\right)^T \exp\left(-\frac{\sum_{t=1}^{T}(y_t - \gamma)^2}{2\sigma^2}\right)
\]

\[
= \left(\frac{1}{\sigma \sqrt{2\pi}}\right)^T \exp\left(-\frac{\sum_{t=1}^{T}(y_t - \bar{y})^2 + T(\bar{y} - \gamma)^2}{2\sigma^2}\right)
\]

\[
= \left(\frac{1}{\sigma \sqrt{2\pi}}\right)^T \exp\left(-\frac{\sum_{t=1}^{T}(y_t - \bar{y})^2}{2\sigma^2}\right) \exp\left(-\frac{T(\bar{y} - \gamma)^2}{2\sigma^2}\right)
\]

and, likewise,

\[
\prod_{t=1}^{T} p_j(y_t; \gamma) = \left(\frac{1}{\sigma \sqrt{2\pi}}\right)^T \times \exp\left(-\frac{\sum_{t=1}^{T}(y_t - \bar{y})^2}{2\sigma^2}\right) \exp\left(-\frac{T(\bar{y} - \gamma)^2}{2\sigma^2}\right).
\]
Therefore,

\[
\left| \prod_{i=1}^{T} p(y_i; \gamma) - \prod_{i=1}^{T} p(y_i; \gamma) \right| = \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^T \exp \left(- \frac{\sum_{i=1}^{T} (y_i - \bar{y})^2}{2\sigma^2} \right) \\
\times \left| \exp \left(- \frac{T(\bar{y} - \gamma)^2}{2\sigma^2} \right) - \exp \left(- \frac{T(\bar{y} - \delta - \gamma)^2}{2\sigma^2} \right) \right| \\
= \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^{T-1} \exp \left(- \frac{\sum_{i=1}^{T} (y_i - \bar{y})^2}{2\sigma^2} \right) \left| \sqrt{T} \delta \right| \frac{1}{\sigma^2 \sqrt{2\pi}} \\
\times \left| \exp \left(- \frac{(\sqrt{T} \bar{y} - \sqrt{T} \gamma)^2}{2} \right) - \exp \left(- \frac{(\sqrt{T} \bar{y} - \sqrt{T} \gamma - \sqrt{T} \delta)^2}{2} \right) \right| \\
\leq \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^{T-1} \left| \sqrt{T} \delta \right| \chi(\sqrt{T} \delta) \\
\leq \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^{T-1} \left| \sqrt{T} \delta \right| \chi(\sqrt{T} \delta),
\]

where now \( \frac{(\sqrt{T} \bar{y} - \sqrt{T} \gamma - \sqrt{T} \delta)}{\sigma^2} \) is interpreted as the \( z \) in the definition of \( \chi(c) \). By the definition of \( \chi(c) \), the first inequality will hold with equality at some value of \( \bar{y} \). The second inequality holds with equality by setting \( y_i = \bar{y} \) for all \( t \). Hence this bound is sharp.

\( Q.E.D. \)
APPENDIX C: PROOF OF THEOREM 2

Because $Q_0(\gamma)$ is continuous, $\Gamma$ is compact, and $\gamma_0$ is the unique maximizer of $Q_0(\gamma)$, we can find $\varepsilon > 0$ such that $\sup_{|\gamma - \gamma_0| > \varepsilon} Q_0(\gamma) < Q_0(\gamma_0)$ and $Q'_0(\gamma) < 0$ for $|\gamma - \gamma_0| \leq \varepsilon$. We can then find $\eta > 0$ sufficiently small such that $\sup_{|\gamma - \gamma_0| > \varepsilon} Q_0(\gamma) < Q_0(\gamma_0) - 3\eta$ and $Q'_0(\gamma) < -3\eta$ for $|\gamma - \gamma_0| \leq \varepsilon$.

We now show that $|\gamma_j - \gamma_0| \leq \varepsilon$ for $j$ sufficiently large, say for all $j \geq J$. By NM (Lemma 2.4), for example, we have $Q_0(\gamma)$ continuous and

$$\sup_{\gamma \in \Gamma} \left| \frac{1}{T} \sum_{t=1}^{T} \log p(y_t; \gamma) - Q_0(\gamma) \right| = o_p(1).$$

Likewise, we also have $Q_j(\gamma)$ continuous and

$$\sup_{\gamma \in \Gamma} \left| \frac{1}{T} \sum_{t=1}^{T} \log p_j(y_t; \gamma) - Q_j(\gamma) \right| = o_p(1).$$

Because of the definition of the bound $\Delta_j$ and Condition 1, we then have $|Q_j(\gamma) - Q_0(\gamma)| \leq \eta$, $|Q'_j(\gamma) - Q'_0(\gamma)| \leq \eta$, and $|Q''_j(\gamma) - Q''_0(\gamma)| \leq \eta$ for $j$ sufficiently large. Because $-\eta \leq Q_j(\gamma) - Q_0(\gamma) \leq \eta$, we have $Q_j(\gamma) \leq Q_0(\gamma) + \eta$, in particular for $|\gamma - \gamma_0| > \varepsilon$. We therefore obtain

$$\sup_{|\gamma - \gamma_0| > \varepsilon} Q_j(\gamma) \leq \sup_{|\gamma - \gamma_0| > \varepsilon} Q_0(\gamma) + \eta. \tag{10}$$

We also have

$$\sup_{|\gamma - \gamma_0| > \varepsilon} Q_0(\gamma) < Q_0(\gamma_0) - 3\eta. \tag{11}$$

Combining (10) and (11), we obtain $\sup_{|\gamma - \gamma_0| > \varepsilon} Q_j(\gamma) \leq Q_0(\gamma_0) - 2\eta$ or

$$Q_0(\gamma_0) \geq \sup_{|\gamma - \gamma_0| > \varepsilon} Q_j(\gamma) + 2\eta. \tag{12}$$

Because $Q_j(\gamma) \geq Q_0(\gamma) - \eta$ for $|\gamma - \gamma_0| \leq \varepsilon$, we have $\sup_{|\gamma - \gamma_0| \leq \varepsilon} Q_j(\gamma) \geq \sup_{|\gamma - \gamma_0| \leq \varepsilon} Q_0(\gamma) - \eta$. But because $\sup_{|\gamma - \gamma_0| \leq \varepsilon} Q_0(\gamma) = Q_0(\gamma_0)$, we have

$$\sup_{|\gamma - \gamma_0| \leq \varepsilon} Q_j(\gamma) \geq Q_0(\gamma_0) - \eta. \tag{13}$$

Combining (12) and (13), we obtain

$$\sup_{|\gamma - \gamma_0| \leq \varepsilon} Q_j(\gamma) \geq \sup_{|\gamma - \gamma_0| > \varepsilon} Q_j(\gamma) + \eta,$$

$^1$See, for example, NM (Lemma 2.4).
and the maximizer $\gamma_j$ of $Q_j(\gamma)$ satisfies $|\gamma_j - \gamma_0| \leq \epsilon$.

We now get back to the proof of Theorem 2. By the first order condition, we have $0 = Q_j'(\gamma_j)$. By the mean value theorem, we obtain $0 = Q_j'(\gamma_0) + Q_j''(\gamma^*_j)(\gamma_j - \gamma_0)$, where $\gamma^*_j$ is on the line segment adjoining $\gamma_j$ and $\gamma_0$. We therefore have $\gamma_j - \gamma_0 = -Q_j''(\gamma^*_j)(\gamma_j - \gamma_0)$. Because $|\gamma^*_j - \gamma_0| \leq |\gamma_j - \gamma_0| \leq \epsilon$, we can see that $Q_0''(\gamma^*_j) < -3\eta$. This means that $Q_j''(\gamma^*_j) < -2\eta$ and that the division is well defined. Hence,

\begin{equation}
|\gamma_j - \gamma_0| \leq |Q_j'(\gamma_0)|/2\eta \leq \frac{\Delta_j}{2\eta}
\end{equation}

for all $j \geq J$. (Roughly speaking, this inequality indicates that when the approximation is sufficiently precise, the difference between $\gamma_j$ and $\gamma_0$ depends on the degree of approximation and the concavity of the objective function at $\gamma_0$.) For $j < J$, let

\begin{equation}
\varrho = \max_{1 \leq j < J} \left\{ \frac{|\gamma_j - \gamma_0|}{\Delta_j} : 1(\Delta_j > 0) \right\},
\end{equation}

where $1(\cdot)$ denotes the indicator function. Let $\zeta = \max(\frac{1}{2\varrho}, \varrho)$ (note that $\zeta$ does not depend on $T$).

Combining (14) and (15), we conclude that

\[ |\gamma_j - \gamma_0| \leq \zeta \cdot \Delta_j \]

for all $j$. \textit{Q.E.D.}

**APPENDIX D: PROOF OF THEOREM 3**

Note that

\[ \sup_{\gamma \in \Gamma} \left| \frac{1}{T} \sum_{t=1}^{T} \log p_j(y; \gamma) - \frac{1}{T} \sum_{t=1}^{T} \log p(y; \gamma) \right| \leq \Delta_j \]

by definition and

\[ \sup_{\gamma \in \Gamma} \left| \frac{1}{T} \sum_{t=1}^{T} \log p(y; \gamma) - Q(\gamma) \right| = o_p(1). \]

This implies that

\begin{equation}
\sup_{\gamma \in \Gamma} \left| \frac{1}{T} \sum_{t=1}^{T} \log p_j(y; \gamma) - Q(\gamma) \right| \leq \Delta_j + o_p(1) = o_p(1)
\end{equation}

by the assumption that $\Delta_j = o(1)$. Combining (16) with Conditions 2 and 4, and using NM (Theorem 2.5), we obtain the desired conclusion. \textit{Q.E.D.}
APPENDIX E: PROOF OF THEOREM 4

Recalling that Theorem 1 implies that \( \lim_{T \to \infty} \hat{\gamma}_j = \gamma_j \), consider

\[
\hat{\gamma}_j - \gamma_0 = (\hat{\gamma}_j - \gamma_j) + (\gamma_j - \gamma_0).
\]

We first characterize the asymptotic distribution of \( \sqrt{T}(\hat{\gamma}_j - \gamma_j) \). Note that \( \gamma_j \) and \( \hat{\gamma}_j \) solve

\[
0 = E[\nabla_{\gamma} \log p_j(y_t; \gamma_j)],
\]

\[
0 = \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma} \log p_j(y_t; \hat{\gamma}_j).
\]

Expanding the second equality around \( \gamma_j \) and using the mean value theorem, we obtain

\[
0 = \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma} \log p_j(y_t; \gamma_j) + \left( \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma \gamma} \log p_j(y_t; \tilde{\gamma}_j) \right) (\hat{\gamma}_j - \gamma_j),
\]

where \( \tilde{\gamma}_j \) is on the line segment adjoining \( \hat{\gamma}_j \) and \( \gamma_j \). It follows that

\[
\sqrt{T}(\hat{\gamma}_j - \gamma_j) = -\left( \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma \gamma} \log p_j(y_t; \tilde{\gamma}_j) \right)^{-1} \times \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma} \log p_j(y_t; \gamma_j).
\]

Note that

\[
\left| \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma \gamma} \log p_j(y_t; \tilde{\gamma}_j) - \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma \gamma} \log p(y_t; \tilde{\gamma}_j) \right| \leq \Delta_j
\]

by definition. We also have

\[
\left| \frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma \gamma} \log p(y_t; \tilde{\gamma}_j) - Q_0'(\tilde{\gamma}_j) \right| = o_p(1)
\]

by NM (Lemma 2.4), for example. Finally, because \( \tilde{\gamma}_j = \gamma_0 + o_p(1) \) (since \( \tilde{\gamma}_j \) is on the line segment between \( \gamma_0 \) and \( \gamma_j \)), and because \( Q_0'(\gamma) \) is continuous by dominated convergence, we have

\[
Q_0'(\tilde{\gamma}_j) = Q_0'(\gamma_0) + o_p(1).
\]
Combining (19), (20), and (21), and the assumption that $\Delta_j \to 0$, we obtain that
\begin{equation}
\frac{1}{T} \sum_{t=1}^{T} \nabla_{\gamma_j} \log p_j(y_t; \tilde{\gamma}_j) = Q_0''(\gamma_0) + o_p(1).
\end{equation}

We also note that
\begin{equation}
E \left[ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma_j} \log p_j(y_t; \gamma_j) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma} \log p(y_t; \gamma_0) \right] = 0,
\end{equation}

since by definition $\gamma_j$ and $\gamma_0$ maximize $Q_j(\gamma)$ and $Q_0(\gamma)$, respectively. In addition, since $\Delta_j \to 0$,
\begin{equation}
\text{Var} \left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma_j} \log p_j(y_t; \gamma_j) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma} \log p(y_t; \gamma_0) \right)
= E \left[ (\nabla_{\gamma_j} \log p_j(y_t; \gamma_j) - \nabla_{\gamma} \log p(y_t; \gamma_0))^2 \right] = o(1).
\end{equation}

Combining (23) and (24) and applying Chebyshev’s inequality, we conclude that
\begin{equation}
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma_j} \log p_j(y_t; \gamma_j) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma} \log p(y_t; \gamma_0) + o_p(1).
\end{equation}

Now, (18), (22), and (25) imply that
\begin{equation}
\sqrt{T}(\hat{\gamma}_j - \gamma_j) = -Q_0''(\gamma_0)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \nabla_{\gamma} \log p(y_t; \gamma_0) + o_p(1)
\Rightarrow N(0, -Q_0''(\gamma_0)^{-1})
\end{equation}
as $T \to \infty$ and $\Delta_j \to 0$. The second line in (26) uses the central limit theorem and information equality.

Last, note that Theorem 2 implies that $\gamma_j - \gamma_0 = O(\Delta_j)$ or
\begin{equation}
\sqrt{T}(\gamma_j - \gamma_0) = O(\sqrt{T} \Delta_j).
\end{equation}

Combining this with (17) and (26), we conclude that
\begin{equation}
\sqrt{T}(\hat{\gamma}_j - \gamma_0) \Rightarrow N(0, -Q_0''(\gamma_0)^{-1})
as $T \to \infty$ and $\sqrt{T} \Delta_j \to 0$. Q.E.D.