

## Supplement to “Inference in dynamic stochastic general equilibrium models with possible weak identification”

(Quantitative Economics, Vol. 5, No. 2, July 2014, 457–494)

ZHONGJUN QU

Department of Economics, Boston University

### APPENDIX B

PROOF OF LEMMA A.1. The proof uses similar arguments as in [Dunsmuir \(1979\)](#), but allowing for weak identification and selecting a subset of frequencies using  $W(\omega)$ . It consists of two steps. Step 1 proves asymptotic normality and Step 2 verifies that the limiting covariance matrix is an identity matrix.

STEP 1. First consider  $\xi_{1T}$ . Rewrite it as

$$\xi_{1T} = \frac{1}{2\sqrt{T}} \sum_{j=1}^{T-1} \phi_T(\omega_j)^* \text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) \quad (\text{B.1})$$

$$+ \frac{1}{2\sqrt{T}} \sum_{j=1}^{T-1} \phi_T(\omega_j)^* \text{vec}(\mathbb{E}I_T(\omega_j) - f_{\theta_0}(\omega_j)). \quad (\text{B.2})$$

The term (B.2) is asymptotically negligible. Specifically,  $\mathbb{E}I_T(\omega)$  can be expressed as

$$\mathbb{E}I_T(\omega) = \sum_{s=-T+1}^{T-1} \left(1 - \frac{|s|}{T}\right) \Gamma(s) \exp(-is\omega)$$

with

$$\Gamma(s) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\theta_0}(\omega) \exp(is\omega) d\omega.$$

Using the property of the Cesaro sum and that  $f_{\theta_0}(\omega)$  belongs to the Lipschitz class of degree  $\beta$  with respect to  $\omega$ , we have ([Hannan \(1970, p. 513\)](#))

$$\sup_{\omega \in [-\pi, \pi]} \|\text{vec}(\mathbb{E}I_T(\omega) - f_{\theta_0}(\omega))\| = O(T^{-\beta}).$$

---

Zhongjun Qu: [qu@bu.edu](mailto:qu@bu.edu)

Copyright © 2014 Zhongjun Qu. Licensed under the [Creative Commons Attribution-NonCommercial License 3.0](https://creativecommons.org/licenses/by-nc/3.0/). Available at <http://www.qeconomics.org>.

DOI: 10.3982/QE287

The term (B.2) is, therefore, bounded by

$$\begin{aligned} \frac{1}{2} T^{1/2} \sup_{\omega \in [-\pi, \pi]} \|\phi_T(\omega)\| \sup_{\omega \in [-\pi, \pi]} \|\text{vec}(\mathbb{E}I_T(\omega) - f_{\theta_0}(\omega))\| &= O(T^{-\beta+1/2}) \\ &= o(1), \end{aligned}$$

where the first equality is because  $\phi_T(\omega)$  is finite by Assumption W and the last equality follows because  $\beta > 1/2$ . Thus, to derive the limiting distribution of  $\xi_{1T}$ , it suffices to consider (B.1) only.

Let  $\phi_{TM}(\omega)$  denote the  $(M-1)$ th order Cesaro sum of the Fourier series for  $\phi_T(\omega)$ :

$$\phi_{TM}(\omega) = \sum_{s=-M+1}^{M-1} \left(1 - \frac{|s|}{M}\right) \eta_T(s) \exp(-is\omega)$$

with

$$\eta_T(s) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_T(\omega) \exp(is\omega) d\omega.$$

Then

$$\begin{aligned} \text{(B.1)} &= \frac{1}{2\sqrt{T}} \sum_{j=1}^{T-1} \phi_{TM}(\omega_j)^* \text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) \\ &\quad + \frac{1}{2\sqrt{T}} \sum_{j=1}^{T-1} (\phi_T(\omega_j) - \phi_{TM}(\omega_j))^* \text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)). \end{aligned} \tag{B.3}$$

The second term will be asymptotically negligible if, because of conjugacy,

$$\frac{1}{2\sqrt{T}} \sum_{j=1}^{\lfloor T/2 \rfloor} (\phi_T(\omega_j) - \phi_{TM}(\omega_j))^* \text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) = o_p(1). \tag{B.4}$$

Establishing this result faces some difficulty because  $\phi_T(\omega)$  has a finite number of discontinuities within  $[0, \pi]$  due to the presence of  $W(\omega)$ , implying  $\phi_T(\omega_j) - \phi_{TM}(\omega_j)$  does not converge uniformly to zero over  $[0, \pi]$  (the Gibbs phenomenon). However, results in Hannan (1970, pp. 506–507) imply that  $\phi_{TM}(\omega)$  converges uniformly to  $\phi_T(\omega)$  over all closed intervals excluding the jumps. At the jumps, the approximation errors remain bounded. Assume the jumps occur at  $\tilde{\omega}^k$  ( $k = 1, \dots, K$ ). Then, for any  $\varepsilon > 0$ , there exist finite constants  $M > 0$  and  $C > 0$  independent of  $T$ , such that

$$\|\phi_{TM}(\omega) - \phi_T(\omega)\| \leq \begin{cases} C, & \text{if } \omega \in I_1 \equiv \bigcup_{k=1}^K [\tilde{\omega}^k - \varepsilon, \tilde{\omega}^k + \varepsilon], \\ \varepsilon, & \text{if } \omega \in [0, \pi] \text{ but } \omega \notin I_1. \end{cases}$$

By applying the above partition, (B.4) can be decomposed into

$$\frac{1}{2\sqrt{T}} \sum_{j=1}^{[T/2]} 1(\omega_j \in I_1) (\phi_T(\omega_j) - \phi_{TM}(\omega_j))^* \text{vec}(I_T(\omega_j) - EI_T(\omega_j)) \quad (\text{T1})$$

$$+ \frac{1}{2\sqrt{T}} \sum_{j=1}^{[T/2]} 1(\omega_j \notin I_1) (\phi_T(\omega_j) - \phi_{TM}(\omega_j))^* \text{vec}(I_T(\omega_j) - EI_T(\omega_j)). \quad (\text{T2})$$

For the first term,

$$\begin{aligned} \|\text{Var}(\text{T1})\| &\leq \frac{C^2}{T} \sum_{j=1}^{[T/2]} 1(\omega_j \in I_1) \|\text{Var}\{\text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j))\}\| \\ &\quad + \frac{C^2}{T} \sum_{j=1}^{[T/2]} \sum_{h=1, h \neq j}^{[T/2]} 1(\omega_j \in I_1) \\ &\quad \times \|\mathbb{E}\{\text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) \text{vec}(I_T(\omega_h) - \mathbb{E}I_T(\omega_h))^*\}\| \\ &\leq \frac{C^2}{T} \sum_{j=1}^{[T/2]} 1(\omega_j \in I_1) D + \frac{C^2}{T^2} \sum_{j=1}^{[T/2]} \sum_{h=1}^{[T/2]} 1(\omega_j \in I_1) D, \end{aligned}$$

where  $D$  is some finite constant and the second inequality follows from Theorem 11.7.1 in Brockwell and Davis (1991), that is, for any  $\omega_j$  and  $\omega_h$  in  $[0, \pi]$ ,

$$\mathbb{E}\{\text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) \text{vec}(I_T(\omega_h) - \mathbb{E}I_T(\omega_h))^*\} = \begin{cases} O(1), & \text{if } h = j, \\ O(T^{-1}), & \text{otherwise.} \end{cases}$$

Because the length of  $I_1$  can be made arbitrarily small by choosing a small  $\varepsilon$  and a large  $M$ , we have  $\text{Var}(\text{T1}) = o(1)$ . Similar arguments can be applied to (T2),

$$\|\text{Var}(\text{T2})\| \leq T^{-1} \varepsilon^2 \sum_{j=1}^{[T/2]} 1(\omega_j \notin I_1) D + \varepsilon^2 T^{-2} \sum_{j=1}^{[T/2]} \sum_{h=1}^{[T/2]} 1(\omega_j \notin I_1) D \leq 2D\varepsilon^2,$$

which can again be made small by choosing a small  $\varepsilon$  and a large  $M$ . Thus,  $\text{Var}(\text{T2}) = o(1)$ . Combining the above results, we have proved (B.4).

It remains to analyze the first term in (B.3). Apply the definition of  $\phi_{TM}(\omega_j)$ ,

$$\begin{aligned} &\frac{1}{2\sqrt{T}} \sum_{j=1}^{T-1} \phi_{TM}(\omega_j)^* \text{vec}(I_T(\omega_j) - \mathbb{E}I_T(\omega_j)) \\ &= \frac{1}{4\pi} \sum_{s=-M-1}^{M-1} \left(1 - \frac{|s|}{M}\right) \eta_T(s)^* \{\sqrt{T} \text{vec}(\hat{\Gamma}(s) - \mathbb{E}\hat{\Gamma}(s))\} \quad (\text{T3}) \end{aligned}$$

$$+ \frac{1}{4\pi} \sum_{s=-M-1}^{M-1} \left(1 - \frac{|s|}{M}\right) \eta_T(s)^* \{\sqrt{T} \text{vec}(\hat{\Gamma}(s-T) - \mathbb{E}\hat{\Gamma}(s-T))\}, \quad (\text{T4})$$

where the last equality uses  $\sum_{s=1}^T \exp(-is\omega_j) = 0$  unless  $s = kT$  ( $k = 0, \pm 1, \dots$ ) and

$$\hat{\Gamma}(s) = \begin{cases} T^{-1} \sum_{t=1}^{T-s} (Y_{t+s} - \mu(\theta_0))(Y_t - \mu(\theta_0))', & \text{if } 0 \leq s \leq T-1, \\ \hat{\Gamma}(-s)', & \text{if } -T+1 \leq s \leq 0. \end{cases}$$

Term (T4) converges in probability to zero. This is because  $M$  is finite,  $\eta_T(s)^*$  is uniformly bounded, and  $\sqrt{T} \text{vec}(\hat{\Gamma}(s-T) - \mathbb{E}\hat{\Gamma}(s-T)) \rightarrow^p 0$  for each  $|s| < M$  by the definition of  $\hat{\Gamma}(s-T)$  (note that the summation in the definition of  $\hat{\Gamma}(s-T)$  involves at most  $M$  terms). In (T3),  $\sqrt{T} \text{vec}(\hat{\Gamma}(s) - \mathbb{E}\hat{\Gamma}(s))$  satisfies a central limit theorem for each  $|s| \leq M$ ; see Hannan (1976). Thus, (T3) converges to a vector of normal random variables because  $M$  is finite. Therefore,  $\xi_{1T}$  has a multivariate normal limiting distribution. For  $\xi_{2T}$ ,  $\psi_T$  is finite because of Assumption W. Its asymptotic normality then follows from the central limit theorem.

STEP 2. For  $\xi_{1T}$ , it suffices to examine the covariance matrix of (T3). Apply the definition of  $\eta_T(s)$  and use the relationship between the vec and the trace operator. Its  $l$ th element can be written as

$$\begin{aligned} \xi_{TM}(l) &= \frac{1}{8\pi^2} \sum_{s=-M-1}^{M-1} \left(1 - \frac{|s|}{M}\right) \\ &\quad \times \int_{-\pi}^{\pi} W(\omega) \text{tr}\{B(l, \omega) \sqrt{T}(\hat{\Gamma}(s) - \mathbb{E}\hat{\Gamma}(s))\} \exp(-is\omega) d\omega, \end{aligned} \quad (\text{B.5})$$

where

$$B(l, \omega) = f_{\theta_0}^{-1}(\omega) \left( \sum_{k=1}^q \frac{\partial f_{\theta_0}(\omega)}{\partial \theta_k} [Q_T^c(\theta_0) \Lambda_T^c(\theta_0)^{-1/2}]_{kl} \right) f_{\theta_0}^{-1}(\omega)$$

with  $[\cdot]_{kl}$  denoting the  $(k, l)$ th element of the matrix inside the bracket. Because  $B(l, \omega)$  is an  $n_Y$ -by- $n_Y$  matrix, (B.5) can be further rewritten as

$$\begin{aligned} \xi_{TM}(l) &= \frac{1}{8\pi^2} \sum_{s=-M-1}^{M-1} \left(1 - \frac{|s|}{M}\right) \\ &\quad \times \int_{-\pi}^{\pi} W(\omega) \left\{ \sum_{a,b=1}^{n_Y} B_{ba}(l, \omega) \sqrt{T}(\hat{\Gamma}_{ab}(s) - \mathbb{E}\hat{\Gamma}_{ab}(s)) \right\} \exp(-is\omega) d\omega, \end{aligned}$$

where  $B_{ba}(l, \omega)$  is the  $(b, a)$ th element of the corresponding matrix. Therefore,

$$\begin{aligned} &\text{Cov}(\xi_{TM}(l), \xi_{TM}(k)) \\ &= \frac{1}{64\pi^4} \sum_{a,b,c,d=1}^{n_Y} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} W(r) B_{ba}(l, r) W(\lambda) B_{dc}(k, \lambda)^* \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned} & \times \sum_{s,h=-M+1}^{M-1} \left(1 - \frac{|s|}{M}\right) \left(1 - \frac{|h|}{M}\right) \\ & \times \mathbb{E}\{T(\hat{\Gamma}_{ab}(s) - \mathbb{E}\hat{\Gamma}_{ab}(s))(\hat{\Gamma}_{cd}(h) - \mathbb{E}\hat{\Gamma}_{cd}(h))\} \exp(-isr) dr \exp(ih\lambda) d\lambda. \end{aligned}$$

The only random elements in (B.6) are the sample autocovariances, which satisfy (see equation (3) on p. 397 in Hannan (1976))

$$\begin{aligned} & \mathbb{E}\{T(\hat{\Gamma}_{ab}(s) - \mathbb{E}\hat{\Gamma}_{ab}(s))(\hat{\Gamma}_{cd}(h) - \mathbb{E}\hat{\Gamma}_{cd}(h))\} \\ & \rightarrow 2\pi \int_{-\pi}^{\pi} f_{ac}(\omega) \overline{f_{bd}(\omega)} \exp(-i(h-s)\omega) d\omega \end{aligned} \quad (\text{T5})$$

$$+ 2\pi \int_{-\pi}^{\pi} f_{ad}(\omega) \overline{f_{bc}(\omega)} \exp(i(s+h)\omega) d\omega, \quad (\text{T6})$$

where  $f_{ac}(\omega)$  stands for the  $(a, c)$ th element of  $f_{\theta_0}(\omega)$ . Applying (T5) to (B.6) leads to

$$\begin{aligned} & \frac{1}{8\pi} \int_{-\pi}^{\pi} \sum_{a,b,c,d=1}^q f_{ac}(\omega) \overline{f_{bd}(\omega)} \\ & \times \left( \int_{-\pi}^{\pi} W(r) B_{ba}(l, r) \left\{ \frac{1}{2\pi} \sum_{s=-M+1}^{M-1} \left(1 - \frac{|s|}{M}\right) \exp(-is(r-\omega)) \right\} dr \right) \\ & \times \left( \int_{-\pi}^{\pi} W(\lambda) B_{dc}(k, \lambda)^* \left\{ \frac{1}{2\pi} \sum_{h=-M+1}^{M-1} \left(1 - \frac{|h|}{M}\right) \exp(-ih(\omega-\lambda)) \right\} d\lambda \right) d\omega. \end{aligned}$$

The two terms inside the two curly brackets are Fejér's kernels. Therefore,

$$\begin{aligned} & \int_{-\pi}^{\pi} W(r) B_{ba}(l, r) \left\{ \frac{1}{2\pi} \sum_{s=-M+1}^{M-1} \left(1 - \frac{|s|}{M}\right) \exp(-is(r-\omega)) \right\} dr \rightarrow W(\omega) B_{ba}(l, \omega), \\ & \int_{-\pi}^{\pi} W(\lambda) B_{dc}(k, \lambda)^* \left\{ \frac{1}{2\pi} \sum_{h=-M+1}^{M-1} \left(1 - \frac{|h|}{M}\right) \exp(-ih(\omega-\lambda)) \right\} d\lambda \\ & \rightarrow W(\omega) B_{dc}(k, \omega)^* \end{aligned}$$

uniformly over all closed intervals excluding the jumps. At the jumps, the approximation error is finite and, therefore, it does not interfere with the limiting results. The effect of (T6) can be analyzed similarly. Combining the two results, we have

$$\begin{aligned} & \text{Cov}(\xi_{TM}(l) \xi_{TM}(k)) \\ & \rightarrow \frac{1}{8\pi} \int_{-\pi}^{\pi} \sum_{a,b,c,d=1}^{ny} W(\omega) \{ f_{ac}(\omega) \overline{f_{bd}(\omega)} B_{ba}(l, \omega) B_{dc}(k, \omega)^* \\ & + f_{ad}(\omega) \overline{f_{bc}(\omega)} B_{ba}(l, \omega) B_{dc}(k, -\omega)^* \} d\omega \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4\pi} \int_{-\pi}^{\pi} W(\omega) \operatorname{tr}\{f_{\theta_0}(\omega)B(k, \omega)f_{\theta_0}(\omega)B(l, \omega)\} d\omega \\
&= \frac{1}{2T} \sum_{j=1}^{T-1} W(\omega_j) \operatorname{vec}(B(l, \omega_j))^* (f_{\theta_0}(\omega_j)' \otimes f_{\theta_0}(\omega_j)) \operatorname{vec}(B(k, \omega_j)) + o(1),
\end{aligned}$$

where the first equality uses  $B_{dc}(k, \omega)^* = B_{cd}(k, \omega)$ ,  $\overline{f_{bd}(\omega)} = f_{db}(\omega)$ ,  $\overline{f_{bc}(\omega)} = f_{cb}(\omega)$ ,  $B_{dc}(k, -\omega)^* = B_{dc}(k, \omega)$ , and the last equality follows because the summand belongs to  $\operatorname{Lip}(\beta)$  with  $\beta > 1/2$ . In matrix notation, the above result can be stated as

$$\begin{aligned}
&\operatorname{Var}(\xi_{1T}) \\
&= \Lambda_T^c(\theta_0)^{-1/2} Q_T^c(\theta_0)' \\
&\quad \times \left\{ \frac{1}{2T} \sum_{j=0}^{T-1} W(\omega_j) \left( \frac{\partial \operatorname{vec} f_{\theta_0}(\omega_j)}{\partial \theta'} \right)^* (f_{\theta_0}^{-1}(\omega_j)' \otimes f_{\theta_0}^{-1}(\omega_j)) \frac{\partial \operatorname{vec} f_{\theta_0}(\omega_j)}{\partial \theta'} \right\} \\
&\quad \times Q_T^c(\theta_0) \Lambda_T^c(\theta_0)^{-1/2} + o(1).
\end{aligned}$$

Now consider  $\xi_{2T}$ . It is asymptotically independent of  $\xi_{1T}$ , satisfying

$$\begin{aligned}
\operatorname{Var}(\xi_{2T}) &\rightarrow \frac{1}{2\pi} \Lambda_T^c(\theta_0)^{-1/2} Q_T^c(\theta_0)' \\
&\quad \times \left\{ W(0) \frac{\partial \mu(\theta_0)'}{\partial \theta} f_{\theta_0}^{-1}(0) \frac{\partial \mu(\theta_0)}{\partial \theta'} \right\} Q_T^c(\theta_0) \Lambda_T^c(\theta_0)^{-1/2}.
\end{aligned}$$

Therefore,  $\operatorname{Var}(\xi_{1T} + \xi_{2T}) = \Lambda_T^c(\theta_0)^{-1/2} Q_T^c(\theta_0)' M_T(\theta_0) Q_T^c(\theta_0) \Lambda_T^c(\theta_0)^{-1/2} + o(1) \rightarrow \mathbb{I}_{q_1+q_2}$ , where the last equality uses the definition of  $Q_T^c(\theta_0)$  and  $\Lambda_T^c(\theta_0)$ .  $\square$

**ADDITIONAL PROOF.** This proof shows that the confidence band covers the impulse response function with probability at least  $(1 - \alpha)$  asymptotically. Let  $C_\theta(1 - \alpha)$  denote the  $(1 - \alpha)$  confidence set for  $\theta$  obtained by inverting  $S_T(\theta)$  and let  $C_{\text{IR}}$  denote the confidence band for the impulse response function obtained from Steps 1–3. By construction, if  $\theta_0 \in C_\theta(1 - \alpha)$ , then  $\text{IR}(\theta_0) \in C_{\text{IR}}$ . Thus, if  $\text{IR}(\theta_0) \notin C_{\text{IR}}$ , then  $\theta_0 \notin C_\theta(1 - \alpha)$ . Equivalently,  $\Pr(\text{IR}(\theta_0) \notin C_{\text{IR}}) \leq \Pr(\theta_0 \notin C_\theta(1 - \alpha))$ . As  $T \rightarrow \infty$ ,  $\Pr(\theta_0 \notin C_\theta(1 - \alpha)) \rightarrow \alpha$ . Therefore,  $\lim_{T \rightarrow \infty} \Pr(\text{IR}(\theta_0) \notin C_{\text{IR}}) \leq \alpha$ .  $\square$

### *Eigenvalue conditions that correspond to other characterizations of weak identification*

We illustrate that the characterizing conditions for weak identification used in the IV and GMM literature can be stated using the curvatures of the criterion functions used for inference as in Assumption W.

**Linear IV (Staiger and Stock (1997)):** Consider the model  $y = Y\beta + u$ ,  $Y = ZII + v$ , where  $y$  and  $Y$  are  $T \times 1$  vectors,  $Z$  is a  $T \times K$  matrix of instruments, and  $u$  and  $v$  are

$T \times 1$  vectors of disturbances with  $\mathbb{E}uu' = \sigma_u^2 I_T$ . The objective function is  $Q(\beta) = (y - Y\beta)' P_Z(y - Y\beta)$ . Its first order derivative, normalized by  $T^{-1/2}$ , equals

$$D_T(\beta_0) = -2T^{-1/2}u'Z\Pi - 2T^{-1/2}(u'Z)(Z'Z)^{-1}(Z'v).$$

If  $\beta_0$  is strongly identified, that is,  $\Pi$  is nonzero and independent of  $T$ , then the first term in  $D_T(\beta_0)$  is of exact order  $O_p(1)$  and the second is  $O_p(T^{-1/2})$ . Therefore,

$$\lim_{T \rightarrow \infty} \mathbb{E}(D_T(\beta_0)D_T(\beta_0)') = \lim_{T \rightarrow \infty} 4T^{-1}\mathbb{E}(\Pi'Z'Z\Pi)\sigma_u^2,$$

consistent with the order of  $\Lambda_{1T}(\theta_0)$  in Assumption W. If  $\beta_0$  is weakly identified, that is,  $\Pi = T^{-1/2}C$ , then  $D_T(\beta_0)$  is of exact order  $O_p(T^{-1/2})$ . Therefore, the eigenvalue of  $\mathbb{E}(D_T(\beta_0)D_T(\beta_0)')$  is of order  $O(T^{-1})$ , consistent with the order of  $\Lambda_{2T}(\theta_0)$  in Assumption W.

Weak identification in a continuous updating GMM (CU-GMM) setting (Kleibergen (2005)): Consider inference based on the moment restriction  $\mathbb{E}\phi_t(\theta_0) = 0$  with  $\theta_0 \in R^m$ . Without loss of generality, assume  $\phi_t(\theta_0)$  is serially uncorrelated. Let  $f_T(\theta) = T^{-1/2} \sum_{t=1}^T \phi_t(\theta)$ . Then the CU-GMM criterion function is given by  $Q_T(\theta) = f_T(\theta)' \hat{V}_{ff}(\theta)^{-1} f_T(\theta)$ , where  $\hat{V}_{ff}(\theta) \rightarrow^p V_{ff}(\theta) = \lim_{T \rightarrow \infty} \text{Var}(f_T(\theta))$ . Define

$$\frac{\partial f_T(\theta_0)}{\partial \theta'} = q_T(\theta_0) = (q_{1,T}(\theta_0), \dots, q_{m,T}(\theta_0)).$$

Kleibergen (2005) characterized the strength of identification using the order of  $\mathbb{E}q_T(\theta_0)$ . Under strong identification,  $T^{-1/2}\mathbb{E}q_T(\theta_0)$  has a fixed full rank value, while under weak identification  $T^{-1/2}\mathbb{E}q_T(\theta_0) = T^{-1/2}C$ . We have

$$D_T(\theta_0)' = 2T^{-1/2}f_T(\theta_0)'\hat{V}_{ff}(\theta_0)^{-1}(\hat{R}_T(\theta_0) - \mathbb{E}q_T(\theta_0)) \quad (\text{a})$$

$$+ 2T^{-1/2}f_T(\theta_0)'\hat{V}_{ff}(\theta_0)^{-1}\mathbb{E}q_T(\theta_0), \quad (\text{b})$$

where the  $j$ th column of  $\hat{R}_T(\theta_0)$  equals  $q_{j,T}(\theta_0) - \hat{V}_{\theta f,j}(\theta_0)\hat{V}_{ff}(\theta_0)^{-1}f_T(\theta_0)$ , that is, the residual from projecting  $q_{j,T}(\theta_0)$  onto  $f_T(\theta_0)$ ;  $\hat{V}_{\theta f,j}(\theta_0)$  is the sample covariance between  $f_T(\theta_0)$  and  $q_{j,T}(\theta_0)$ . Thus,

$$\mathbb{E}(D_T(\theta_0)D_T(\theta_0)') = \mathbb{E}(a'a) + \mathbb{E}(b'b) + \mathbb{E}(a'b) + \mathbb{E}(b'a).$$

The first term  $\mathbb{E}(a'a)$  is of order  $O(T^{-1})$  irrespective of the strength of identification. The second term  $\mathbb{E}(b'b)$  is of exact order  $O(1)$  under strong and  $O(T^{-1})$  under weak identification, respectively. The order of  $\mathbb{E}(b'b) + \mathbb{E}(a'b)$  is always lower than that of  $\mathbb{E}(b'b)$ . Therefore, the eigenvalues of  $\mathbb{E}(D_T(\theta_0)D_T(\theta_0)')$  are  $O(1)$  under strong identification and  $O(T^{-1})$  under weak identification, consistent with Assumption W in the paper.

Weak identification under a GMM setting (Stock and Wright (2000)): Consider the same setup as in the CU-GMM case, but with inference based on the GMM criterion function

$$Q_T(\theta) = f_T(\theta)'W_T f_T(\theta),$$

where  $W_T$  is some consistent estimate of the optimal weighting matrix that is, without loss of generality, assumed to be nonrandom. Then

$$D_T(\theta_0)' = 2T^{-1/2}f_T(\theta_0)'W_T(q_T(\theta_0) - \hat{R}_T(\theta_0)) \quad (c)$$

$$+ 2T^{-1/2}f_T(\theta_0)'W_T(\hat{R}_T(\theta_0) - \mathbb{E}q_T(\theta_0)) \quad (d)$$

$$+ 2T^{-1/2}f_T(\theta_0)'W_T\mathbb{E}q_T(\theta_0). \quad (e)$$

Simple algebra shows that the leading term in  $\mathbb{E}(D_T(\theta_0)D_T(\theta_0)')$  is

$$\mathbb{E}(c'c) + \mathbb{E}(c'e) + \mathbb{E}(e'c) + \mathbb{E}(d'd) + \mathbb{E}(e'e).$$

The first four terms are always of order  $O(T^{-1})$  irrespective of the strength of identification. The last term converges to a positive definite matrix under strong identification. Therefore, all the eigenvalues of  $\mathbb{E}(D_T(\theta_0)D_T(\theta_0)')$  are of order  $O(1)$  under strong identification.

Under weak identification, [Stock and Wright \(2000\)](#) assumed that  $\theta$  admits a partition  $\theta = (\alpha', \beta)'$ , such that  $\alpha$  is weakly identified while  $\beta$  is strongly identified. Specifically, let  $m_T(\alpha, \beta) = \mathbb{E}f_T(\alpha, \beta)$ , and write  $m_T(\alpha, \beta) = m_T(\alpha_0, \beta_0) + T^{-1/2}m_{1T}(\alpha, \beta) + m_{2T}(\beta)$  with  $m_{1T}(\alpha, \beta) = T^{1/2}(m_T(\alpha, \beta) - m_T(\alpha_0, \beta))$  and  $m_{2T}(\beta) = m_T(\alpha_0, \beta) - m_T(\alpha_0, \beta_0)$ . [Stock and Wright \(2000\)](#) assumed (cf. Assumption C in their paper)

$$m_{1T}(\alpha, \beta) \rightarrow m_1(\alpha, \beta) \quad \text{and} \quad m_{2T}(\beta) \rightarrow m_2(\beta).$$

Let  $C_T = \text{Diag}(T^{1/2}I_{\dim(\alpha)}, I_{\dim(\beta)})$ . Then

$$\begin{aligned} & C_T\mathbb{E}(e'e)C_T \\ &= 4T^{-1}C_T\mathbb{E}q_T(\theta_0)'W_T\mathbb{E}(f_T(\theta_0)f_T(\theta_0)')W_T\mathbb{E}q_T(\theta_0)C_T \\ &= 4T^{-1}C_T\mathbb{E}q_T(\theta_0)'(W_TV_{ff}(\theta_0)W_T)\mathbb{E}q_T(\theta_0)C_T \\ &\rightarrow 4\left[\frac{\partial m_1(\alpha_0, \beta_0)}{\partial \alpha'} \quad \frac{\partial m_2(\beta_0)}{\partial \beta'}\right]'V_{ff}^{-1}(\theta_0)\left[\frac{\partial m_1(\alpha_0, \beta_0)}{\partial \alpha'} \quad \frac{\partial m_2(\beta_0)}{\partial \beta'}\right]. \end{aligned}$$

The limit is a positive definite matrix. Therefore, in large samples,  $\mathbb{E}(e'e)$  has  $\dim(\beta)$  eigenvalues that are  $O(1)$  and  $\dim(\alpha)$  eigenvalues of order  $O(T^{-1})$ ; so does  $\mathbb{E}(D_T(\theta_0) \times D_T(\theta_0)')$ .

### *Weak identification in a two-equation model*

The model consists of two equations,

$$r_t = \gamma y_t + \beta \pi_t + u_t,$$

$$\pi_t = \rho \pi_{t-1} + v_t,$$

with  $\text{var}(u_t) = \sigma_u^2$ ,  $\text{var}(v_t) = \sigma_v^2$ ,  $\text{cov}(u_t, v_t) = \sigma_{uv}$ , and  $Eu_t u_s = Ev_t v_s = Eu_t v_s = 0$  for all  $t \neq s$ . The first equation is a monetary policy rule ([Taylor \(1993\)](#)) with  $y_t$  and  $\pi_t$  being deviations of GDP and inflation from their targets, and the second equation describes the



inflation dynamics. The parameter of interest is  $\beta$ . To simplify the derivation, we assume  $\rho$ ,  $\gamma$ , and  $\sigma_v^2$  are known. The unknown parameter vector is, therefore,  $\theta = (\beta, \sigma_u^2, \sigma_{uv})$ .

Rewrite the model as

$$\begin{aligned}\tilde{r}_t &= \beta\pi_t + u_t, \\ \pi_t &= \rho\pi_{t-1} + v_t,\end{aligned}\tag{B.7}$$

with  $\tilde{r}_t = r_t - \gamma y_t$ . It can then be viewed as a dynamic version of the limited information simultaneous equation model, in which  $\pi_t$  is the endogenous explanatory and  $\pi_{t-1}$  is the instrument. The parameter  $\beta$  is weakly identified if  $\rho$  is small. Intuitively, because there is little persistence in  $\pi_t$ , it is difficult to differentiate between systematic policy responses ( $\beta\pi_t$ ) and random disturbances ( $u_t$ ). Geometrically, it is possible to move  $\theta$  along a certain direction such that the likelihood surface changes little. In the extreme case with  $\rho = 0$ ,  $\beta$  becomes unidentified. Then there exists a path along which the likelihood is completely flat. (It turns out that changing  $\theta$  in the direction given by  $(1, -2\sigma_{uv}, -\sigma_v^2)$  yields such a non-identification curve.)

We let  $\rho = T^{-1/2}c$  with  $c > 0$ ; other parameter values are independent of  $T$ . Let  $W(\omega) = 1$  for all  $\omega \in [-\pi, \pi]$ .

LEMMA B.1. *Let  $\theta_0$  denote the true value of  $\theta = (\beta, \sigma_u^2, \sigma_{uv})$ . Then  $M_T(\theta_0)$  satisfies the following statements:*

- (i) *It has two positive eigenvalues  $\lambda_{1T}$  and  $\lambda_{2T}$  that satisfy  $T\lambda_{1T} \rightarrow \infty$  and  $T\lambda_{2T} \rightarrow \infty$ .*
- (ii) *The smallest eigenvalue  $\lambda_{3T}$  satisfies*

$$T\lambda_{3T} \rightarrow \frac{16\pi^2\sigma_v^4c^2}{(1 + \sigma_v^4 + 4\sigma_{uv}^2)(\sigma_v^2\sigma_u^2 - \sigma_{uv}^2)}.$$

- (iii) *The elements of*

$$\frac{\partial \text{vec } f_{\theta_0}(\omega)}{\partial \theta'} Q_T(\theta_0) \Lambda_T(\theta_0)^{-1/2}$$

*are bounded and Lipschitz continuous in  $\omega$ .*

Note that Lemma B.1(i) corresponds to Assumption W(i), while Lemma B.1(ii) corresponds to Assumption W(ii); Lemma B.1(iii) is a stronger result than Assumption W(iv).

#### REFERENCES

- Brockwell, P. and R. Davis (1991), *Time Series: Theory and Methods*, second edition. Springer, New York. [3]
- Dunsmuir, W. (1979), "A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise." *The Annals of Statistics*, 7, 490–506. [1]

Hannan, E. J. (1970), *Multiple Time Series*. Wiley, New York. [1, 2]

Hannan, E. J. (1976), "The asymptotic distribution of serial covariances." *The Annals of Statistics*, 4, 396–399. [4, 5]

Kleibergen, F. (2005), "Testing parameters in GMM without assuming that they are identified." *Econometrica*, 73, 1103–1124. [7]

Staiger, D. and J. H. Stock (1997), "Instrumental variables regression with weak instruments." *Econometrica*, 65, 557–586. [6]

Stock, J. H. and J. H. Wright (2000), "GMM with weak identification." *Econometrica*, 68, 1055–1096. [7, 8]

Taylor, J. (1993), "Discretion versus policy rules in practice." *Carnegie–Rochester Conference Series on Public Policy*, 39, 195–214. [8]

---

Submitted May, 2012. Final version accepted September, 2013.