Supplement to "Likelihood-ratio-based confidence sets for the timing of structural breaks"

(*Quantitative Economics*, Vol. 6, No. 2, July 2015, 463–497)

Yunjong Eo
School of Economics, University of Sydney

James Morley
School of Economics, University of New South Wales

This supplement gives proofs for propositions and corollaries in the main text.

Appendix: Proofs

Proof of Proposition 1. Following Qu and Perron (2007a), Qu and Perron (2007b), we consider the jth break date τ_j without loss of generality. The log-profile likelihood ratio subject to the restrictions $g(\beta, \Sigma) = 0$ under the null hypothesis $H_0 : \tau_j = \tau_j^0$ and the alternative hypothesis $H_1 : \tau_j \neq \tau_j^0$ is given by

$$\begin{aligned} \mathrm{LR}_j(\tau_j^0) &= -2\left[l_j^r\left(\tau_j^0, \hat{\beta}(\tau_j^0), \hat{\Sigma}(\tau_j^0)\right) - l_j^r(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma})\right] \\ &= \underbrace{-2\left\{l_j^r\left(\tau_j^0, \hat{\beta}(\tau_j^0), \hat{\Sigma}(\tau_j^0)\right) - l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right)\right\}}_{-\max_{\beta_j, \Sigma_j} lr_j^r(\tau_j^0, \beta_j, \Sigma_j)} \\ &\quad + \underbrace{2\left\{l_j^r(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma}) - l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right)\right\}}_{\max_{\tau_j, \beta_j, \Sigma_j} lr_j^r(\tau_j, \beta_j, \Sigma_j)} \\ &= \max_{\tau_j} lr_j\left(\tau_j, \beta_j^0, \Sigma_j^0\right) + o_p(1), \end{aligned} \tag{A.1}$$

where the maximization is taken over C_M. The second and third lines in (A.1) result from adding and subtracting the log-likelihood at the true values $l_j(\tau_j^0, \beta_j^0, \Sigma_j^0)$ to the first line.[1] The equality of the second and third lines and the fourth line in (A.1) follows from Theorem 1 in Qu and Perron (2007a).

Yunjong Eo: yunjong.eo@sydney.edu.au
James Morley: james.morley@unsw.edu.au

[1]Note that $l_j^r(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma}) = l_j^r(\hat{\tau}_j)$ in (4).

DOI: 10.3982/QE186

We focus on the term $lr_j(\tau_j, \beta_j^0, \Sigma_j^0) = -2[l_j(\tau_j^0, \beta_j^0, \Sigma_j^0) - l_j(\tau_j, \beta_j^0, \Sigma_j^0)]$ in the fourth line of (A.1) so as to find the asymptotic distribution of $\mathrm{LR}_j(\tau_j^0)$. Letting $lr_j(\tau_j, \beta_j^0, \Sigma_j^0) = lr_j(\tau_j - \tau_j^0)$ and $r = \tau_j - \tau_j^0$,

$$
\begin{aligned}
lr_j(r) &= 0 \quad \text{for } r = 0,\\
lr_j(r) &= 2\Bigg(-\frac{r}{2}(\log|\Sigma_j^0| - \log|\Sigma_{j+1}^0|)\\
&\qquad -\frac{1}{2}\sum_{t=\tau_j^0+r}^{\tau_j^0}(y_t - x_t'\beta_{j+1}^0)(\Sigma_{j+1}^0)^{-1}(y_t - x_t'\beta_{j+1}^0)\\
&\qquad -(y_t - x_t'\beta_j^0)(\Sigma_j^0)^{-1}(y_t - x_t'\beta_j^0)\Bigg) \quad \text{for } r < 0,\\
lr_j(r) &= 2\Bigg(-\frac{r}{2}(\log|\Sigma_j^0| - \log|\Sigma_{j+1}^0|)\\
&\qquad -\frac{1}{2}\sum_{t=\tau_j^0+1}^{\tau_j^0+r}(y_t - x_t'\beta_j^0)(\Sigma_j^0)^{-1}(y_t - x_t'\beta_j^0)\\
&\qquad -(y_t - x_t'\beta_{j+1}^0)(\Sigma_{j+1}^0)^{-1}(y_t - x_t'\beta_{j+1}^0)\Bigg) \quad \text{for } r > 0.
\end{aligned}
$$

Then letting $s = v_T^2(\tau_j - \tau_j^0)$, with v_T defined in Assumption 7, the proof of Theorem 3 in Qu and Perron (2007b) shows that for $s \le 0$,

$$lr_j\left(\left[\frac{s}{v_T^2}\right]\right) \Rightarrow 2\left(-\frac{|s|}{2}\Xi_{1,j} + \Lambda_{1,j}W_{1,j}(s)\right), \tag{A.2}$$

and for $s > 0$,

$$lr_j\left(\left[\frac{s}{v_T^2}\right]\right) \Rightarrow 2\left(-\frac{|s|}{2}\Xi_{2,j} + \Lambda_{2,j}W_{2,j}(s)\right), \tag{A.3}$$

where

$$\Lambda_{1,j} = \left(\frac{1}{4}\operatorname{vec}(A_{1,j})'\Omega_{1,j}^0\operatorname{vec}(A_{1,j}) + \delta_j'\Pi_{1,j}\delta_j\right)^{1/2}, \tag{A.4}$$

$$\Lambda_{2,j} = \left(\frac{1}{4}\operatorname{vec}(A_{2,j})'\Omega_{2,j}^0\operatorname{vec}(A_{2,j}) + \delta_j'\Pi_{2,j}\delta_j\right)^{1/2}, \tag{A.5}$$

$$\Xi_{1,j} = \left(\frac{1}{2}\operatorname{tr}(A_{1,j}^2) + \delta_j'Q_{1,j}\delta_j\right), \tag{A.6}$$

$$\Xi_{2,j} = \left(\frac{1}{2}\operatorname{tr}(A_{2,j}^2) + \delta_j'Q_{2,j}\delta_j\right). \tag{A.7}$$

Note that $W_{1,j}(0) = W_{2,j}(0) = 0$ because $W_{1,j}(s)$ and $W_{2,j}(s)$ are independent and starting at $s = 0$.

Qu and Perron (2007a) derive a Bai-type distribution of $\hat{\tau} - \tau_0$ by taking the arg max of (A.2) and (A.3) over C_M and using the continuous mapping theorem. Here, instead, we are deriving the distribution of the likelihood ratio by taking the max of (A.2) and (A.3) over C_M. Thus, under the null hypothesis $H_0 : \tau_j = \tau_j^0$, we have

$$\mathrm{LR}_j(\tau_j^0) \Rightarrow \max_s \begin{cases} 2\left(-\frac{|s|}{2}\varXi_{1,j} + \varLambda_{1,j} W_j(s)\right) & \text{for } s \le 0, \\ 2\left(-\frac{|s|}{2}\varXi_{2,j} + \varLambda_{2,j} W_j(s)\right) & \text{for } s > 0, \end{cases}$$

where we can simplify this expression to relate it to a known distribution from Bhattacharya and Brockwell (1976). Let $\mathrm{LR}_j(\tau_j^0) = \xi = \max[\xi_1, \xi_2]$, where $\xi_1 = \max_{s\le 0} 2(-\frac{|s|}{2}\varXi_{1,j} + \varLambda_{1,j}W_j(s))$ and $\xi_2 = \max_{v>0} 2(-\frac{|s|}{2}\varXi_{2,j} + \varLambda_{2,j}W_j(s))$. By a change in variables $s = (\varLambda_{1,j}^2/\varXi_{1,j}^2)v$ and the distributional equality with $W(a^2x) \equiv aW(x)$, for $s \le 0$,

$$\begin{aligned} \xi_1 &= \sup_{s\le 0} 2\left(-\frac{|s|}{2}\varXi_{1,j} + \varLambda_{1,j} W_j(s)\right) \\ &= \max_{v\le 0} \frac{\varLambda_{1,j}^2}{\varXi_{1,j}} 2\left(-\frac{|v|}{2} + W_j(v)\right) = 2\omega_{1,j} \times \overline{\xi}_1, \end{aligned} \tag{A.8}$$

where $\overline{\xi}_1 = \max_{v\le 0}(-\frac{|v|}{2} + W_j(v))$ and

$$\frac{\varLambda_{1,j}^2}{\varXi_{1,j}} = \frac{\varLambda_{1,j}^2 v_T^2}{\varXi_{1,j} v_T^2} = \frac{\varGamma_{1,j}^2}{\varPsi_{1,j}} \equiv \omega_{1,j}.$$

Similarly, for $s > 0$ with $s = (\varLambda_{2,j}^2/\varXi_{2,j}^2)v$,

$$\begin{aligned} \xi_2 &= \max_{s>0} 2\left(-\frac{|s|}{2}\varXi_{2,j} + \varLambda_{2,j} W_j(s)\right) \\ &= \max_{v>0} \frac{\varLambda_{2,j}^2}{\varXi_{2,j}} 2\left(-\frac{|v|}{2} + W_j(v)\right) = 2\omega_{2,j} \times \overline{\xi}_2, \end{aligned} \tag{A.9}$$

where $\overline{\xi}_2 = \max_{v<0}(-\frac{|v|}{2} + W_j(v))$ and

$$\frac{\varLambda_{2,j}^2}{\varXi_{2,j}} = \frac{\varLambda_{2,j}^2 v_T^2}{\varXi_{2,j} v_T^2} = \frac{\varGamma_{2,j}^2}{\varPsi_{2,j}} \equiv \omega_{2,j}.$$

Thus, we have the simplified expression for the distribution of the likelihood ratio under the null hypothesis:

$$\mathrm{LR}_j(\tau_j^0) \Rightarrow \max_s \begin{cases} 2\omega_{1,j}\left(-\frac{|v|}{2} + W_j(v)\right) & \text{for } v \le 0 \\ 2\omega_{2,j}\left(-\frac{|v|}{2} + W_j(v)\right) & \text{for } v > 0. \end{cases}$$

Bhattacharya and Brockwell (1976) show that $\bar{\xi}_1$ and $\bar{\xi}_2$ in (A.8) and (A.9) are independent and identically distributed exponential random variables with respective distribution functions $P(\bar{\xi}_1 \leq x) = 1 - \exp(-x)$ for $x \leq 0$ and $P(\bar{\xi}_2 \leq x) = 1 - \exp(-x)$ for $x > 0$. Thus,

$$\begin{aligned} P(\xi \leq x) &= P\big(\max[2\omega_{1,j}\bar{\xi}_1, 2\omega_{2,j}\bar{\xi}_2] \leq x\big) \\ &= P(2\omega_{1,j}\bar{\xi}_1 \leq x)P(2\omega_{2,j}\bar{\xi}_2 \leq x) \\ &= P\left(\bar{\xi}_1 \leq \frac{x}{2\omega_{1,j}}\right)P\left(\bar{\xi}_2 \leq \frac{x}{2\omega_{2,j}}\right) \\ &= \left(1 - \exp\left(-\frac{x}{2\omega_{1,j}}\right)\right)\left(1 - \exp\left(-\frac{x}{2\omega_{2,j}}\right)\right). \end{aligned}$$

Then using the distribution of the profile likelihood ratio for the break date τ_j, we can construct a $1-\alpha$ confidence set $C_{j,1-\alpha} = \{\tau_j | \mathrm{LR}_j(\tau_j) \leq \kappa_{\alpha,j}\}$ by inverting the α-level likelihood ratio test. The probability of coverage $C_{j,1-\alpha}$ for any τ_j^0 is given by $P_{\tau_j^0}(\tau_j^0 \in C_{j,1-\alpha})$, where we can easily calculate a critical value $\kappa_{\alpha,j}$ such that

$$\begin{aligned} P_{\tau_j^0}\big(\tau_j^0 \in C_{j,1-\alpha}\big) &= \big(1 - \exp(-\kappa_{\alpha,j}/2\omega_{1,j})\big)\big(1 - \exp(-\kappa_{\alpha,j}/2\omega_{2,j})\big) \\ &= 1 - \alpha. \end{aligned} \tag{A.10}$$

Note that $\kappa_{\alpha,j}$ will be unique because for all $\kappa > 0$, the CDF is a strictly increasing function $\frac{d(1-\exp(-\kappa/2\omega_{1,j}))(1-\exp(-\kappa/2\omega_{2,j}))}{d\kappa} > 0$. □

LEMMA 1. *Under the null hypothesis* $H_0 : \tau = \tau_0$, *if* $lr(\hat{\tau} - \tau_0) \Rightarrow \bar{\xi} = \max_v(-\frac{1}{2}|v| + W(v))$ *for* $v \in (-\infty, \infty)$, *then* $E_{\tau_0}[\lambda\{\tau | lr(\hat{\tau} - \tau) \leq x\}] = 4(1 - \exp(-x))\{x - \frac{1}{2}(1 - \exp(-x))\}$, *where* λ *denotes a Lebesgue measure.*

PROOF. As shown in Bhattacharya and Brockwell (1976), the CDF of $\bar{\xi} = \max_v(-\frac{1}{2}|v| + W(v))$ is given by $P(\bar{\xi} \leq x) = (1 - \exp(-x))^2$. Then letting $C_{1-\alpha} = \{\tau | lr(\hat{\tau} - \tau) \leq \kappa_\alpha\}$, Siegmund (1986) shows that the expected length for a $1-\alpha$ confidence set $C_{1-\alpha}$ is given by

$$\begin{aligned} E_{\tau^0}\big[\lambda\{C_{1-\alpha}\}\big] &= E_{\tau^0}\big[\lambda\{\tau | \tau \in C_{1-\alpha}\}\big] \\ &= \int_{-\infty}^{\infty} P_{\tau^0}(\tau \in C_{1-\alpha})\, d\tau \\ &= 4(1-\alpha)^{1/2}\left\{-\log\big[1 - (1-\alpha)^{1/2}\big] - \frac{1}{2}(1-\alpha)^{1/2}\right\}. \end{aligned} \tag{A.11}$$

See Siegmund (1986) for more details.

Because we can find a critical value κ_α such that

$$P(\bar{\xi} \leq \kappa_\alpha) = \big(1 - \exp(-\kappa_\alpha)\big)^2 = 1 - \alpha,$$

it implies that

$$\kappa_\alpha = -\log\big[1-(1-\alpha)^{1/2}\big]. \tag{A.12}$$

Then, by substituting (A.12) into (A.11), we can express the expected length for a $1-\alpha$ confidence set as a function of the critical value κ_α rather than the level $1-\alpha$ as

$$E_{\tau^0}\big[\lambda\{C_{1-\alpha}\}\big] = 4\big(1-\exp(-\kappa_\alpha)\big)\left\{\kappa_\alpha - \frac{1}{2}\big(1-\exp(-\kappa_\alpha)\big)\right\}. \tag{A.13}$$

□

Proof of Proposition 2. For the general case, as in our setup under Assumptions 1–8, first consider the period before the true jth break date, $\tau_j - \tau_j^0 \le 0$ (i.e., $v \le 0$). Given a critical value $\kappa_{\alpha,j}$, the expected length of a $1-\alpha$ confidence set in the segment $\tau_j - \tau_j^0 \le 0$ can be shown to be

$$\begin{aligned}
&E_{\tau_j^0}\big[\lambda\{\tau_j \,|\, \mathrm{LR}_j(\tau_j) \le \kappa_{\alpha,j}, \hat{\tau}_j - \tau_j \le 0\}\big]\\
&= E_{\tau_j^0}\left[\lambda\left\{\tau_j \;\middle|\; \frac{\mathrm{LR}_j(\tau_j)}{2\omega_{1,j}} \le (\kappa_{\alpha,j}/2\omega_{1,j}), \hat{\tau}_j - \tau_j \le 0\right\}\right]\\
&= \underbrace{\big(\Gamma_{1,j}^2/\Psi_{1,j}^2\big)}_{\text{(i)}}\\
&\quad\times \underbrace{2\big(1-\exp(-\kappa_{\alpha,j}/2\omega_{1,j})\big)\left\{\kappa_{\alpha,j}/2\omega_{1,j} - \frac{1}{2}\big(1-\exp(-\kappa_{\alpha,j}/2\omega_{1,j})\big)\right\}}_{\text{(ii)}}.
\end{aligned} \tag{A.14}$$

The expression (i) in the third line of (A.14) is used for rescaling because the expected length of the confidence set is measured on $v \in (-\infty, 0]$ and

$$\begin{aligned}
\tau_j - \tau_j^0 = r &= s/v_T^2\\
&= \big(\Lambda_{1,j}^2/\Xi_{1,j}^2\big)v/v_T^2\\
&= \big(\Lambda_{1,j}^2 v_T^2/\Xi_{1,j}^2 v_T^4\big)v\\
&= \big(\Gamma_{1,j}^2/\Psi_{1,j}^2\big)v.
\end{aligned} \tag{A.15}$$

Note that from Proposition 1, the second line in (A.14) implies that

$$\frac{\mathrm{LR}_j(\tau_j)}{2\omega_{1,j}} \Rightarrow \overline{\xi} = \max_v\left(-\frac{1}{2}|v| + W_j(v)\right) \quad \text{for } v \le 0. \tag{A.16}$$

Thus, the expression (ii) in the fourth line of (A.14) is calculated for $P(\overline{\xi} \le \frac{\kappa_{\alpha,j}}{2\omega_{1,j}})$ by substituting the critical value $\kappa_{\alpha,j}/2\omega_{1,j}$ into half of the expected length in Lemma 1 given that we are considering $v \le 0$. The expected length for $v > 0$ is calculated in a similar fashion such that the expected length for the entire $1-\alpha$ likelihood-ratio-based confidence set is given by

$$2\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)\left(1-\exp(-\kappa_{\alpha,j}/2\omega_{1,j})\right)\left\{\kappa_{\alpha,j}/2\omega_{1,j}-\frac{1}{2}\left(1-\exp(-\kappa_{\alpha,j}/2\omega_{1,j})\right)\right\}$$
$$+2\left(\Gamma_{2,j}^2/\Psi_{2,j}^2\right)\left(1-\exp(-\kappa_{\alpha,j}/2\omega_{2,j})\right)$$
$$\times\left\{\kappa_{\alpha,j}/2\omega_{2,j}-\frac{1}{2}\left(1-\exp(-\kappa_{\alpha,j}/2\omega_{2,j})\right)\right\}.$$

Note that as either $\omega_{1,j}$ or $\omega_{2,j}$ gets larger (i.e., the magnitude of a structural break is larger), the expected length becomes shorter because there is more precise information about the timing of the structural break. □

PROOF OF COROLLARY 1. If there is no break in variance, $\Sigma_j = \Sigma$ for all j and $B_{1,j} = B_{2,j} = 0$. In addition, if the errors form a martingale difference sequence, $\Pi_{1,j} = Q_{1,j}$ and $\Pi_{2,j} = Q_{2,j}$. From these simplifications, $\omega_{1,j} = \omega_{2,j} = 1$, $(\frac{\Gamma_{1,j}}{\Psi_{1,j}})^2 = \frac{1}{\Delta\beta_j' Q_1 \Delta\beta_j}$, and $(\frac{\Gamma_{2,j}}{\Psi_{2,j}})^2 = \frac{1}{\Delta\beta_j' Q_2 \Delta\beta_j}$. Then, by substituting these values into the critical value and the expected length in Proposition 1, we can find the results in Corollary 1. The results in Remarks 1 and 2 follow in the same way. □

PROOF OF COROLLARY 2. If there is no break in conditional mean, $\Delta\beta_j = 0$ and, in addition, if the standardized errors, η_t, are identically Normally distributed, $\eta_t\eta_t'$ has a Wishart distribution with $\mathrm{var}(\mathrm{vec}(\eta_t\eta_t')) = I_{n^2} + K_n$, where K_n is the commutation matrix. Then $\Omega_{1,j} = \Omega_{2,j} = \Omega = I_{n^2} + K_n$. Furthermore, because K_n is an idempotent matrix,

$$\begin{aligned}&\mathrm{vec}(B_{1,j})'\Omega^0\,\mathrm{vec}(B_{1,j})/4\\&\quad=\mathrm{vec}(B_{1,j})'(I_{n^2}+K_n)\,\mathrm{vec}(B_{1,j})/4\\&\quad=\mathrm{vec}(B_{1,j})'\,\mathrm{vec}(B_{1,j})/2.\end{aligned}$$

Thus,

$$\begin{aligned}\omega_{1,j}&=\frac{\Gamma_{1,j}^2}{\Psi_{1,j}}\\&=\frac{\frac{1}{4}\,\mathrm{vec}(B_{1,j})'\Omega_{1,j}^0\,\mathrm{vec}(B_{1,j})}{\frac{1}{2}\,\mathrm{tr}(B_{1,j}^2)}\\&=\frac{\frac{1}{2}\,\mathrm{vec}(B_{1,j})'\,\mathrm{vec}(B_{1,j})}{\frac{1}{2}\,\mathrm{tr}(B_{1,j}^2)}\\&=1\end{aligned}$$

because $\mathrm{vec}(B_{1,j})'\,\mathrm{vec}(B_{1,j}) = \mathrm{tr}(B_{1,j}^2)$. Similarly, $\omega_{2,j} = 1$. Then $\frac{\Gamma_{1,j}^2}{\Psi_{1,j}^2} = \frac{2}{\mathrm{tr}(B_1^2)}$ and $\frac{\Gamma_{2,j}^2}{\Psi_{2,j}^2} = \frac{2}{\mathrm{tr}(B_2^2)}$. □

References

Bhattacharya, P. K. and P. J. Brockwell (1976), "The minimum of an additive process with applications to signal estimation and storage theory." *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 37, 51–75. [3, 4]

Qu, Z. and P. Perron (2007a), "Estimating and testing structural changes in multivariate regressions." *Econometrica*, 75 (2), 459–502. [1, 3]

Qu, Z. and P. Perron (2007b), "Supplement to 'Estimating and testing structural changes in multivariate regressions'." *Econometrica Supplementary Material*, 75 (2), http://dx.doi.org/10.1111/j.1468-0262.2006.00754.x. [1, 2]

Siegmund, D. (1986), "Boundary crossing probabilities and statistic applications." *The Annals of Statistics*, 14 (2), 361–404. [4]

Submitted April, 2013. Final version accepted May, 2014.