Structural change tests for GEL criteria: Supplementary material

Alain Guay[a], Jean-François Lamarche[b,†]

[a] Département de sciences économiques, Université du Québec à Montréal, Québec, Canada, CIRPÉE and CIREQ

[b] Department of Economics, Brock University, St. Catharines, Ontario, Canada

February 26, 2015

Keywords: Generalized empirical likelihood, generalized method of moments, parameter instability, weak identification, structural change

JEL codes: C12, C13, C15, C32

† The first author gratefully acknowledges financial support of Fonds pour la Formation de Chercheurs et l'aide à la Recherche (FCAR). The first and second authors are thankful for the financial support of the Social Sciences and Humanities Research Council of Canada (SSHRC). Corresponding author: Alain Guay, Department of Economics, Université du Québec à Montréal, Case postale 8888, succursale Centre-ville, Montréal (Québec), Canada, H3C 3P8, E-mail: guay.alain@uqam.ca.

This supplementary material contains lemmas and a longer version of the proofs.

1 Assumptions

We consider triangular arrays because they are required to derive asymptotic results under the Pitman drift alternatives. Define $\mathbf{X}$ to be the domain of $g(\cdot, \theta)$ which includes the support of $x_{T,t}, \forall t, \forall T$. Let B_0 and Δ_0 denote compact subsets of R^r and R^ν that contains neighborhoods of β_0 and δ_0 in the parameter spaces B and Δ. Finally, let $\mu_{T,t}$ denote the distribution of $x_{T,t}$ and let $\bar{\mu}_T = (1/T)\sum_{t=1}^T \mu_{T,t}$. Throughout the Appendix, w.p.a.1 means with probability approaching one; p.s.d. denotes positive semi-definite; $\|\cdot\|$ denotes the Euclidean norm of a vector or matrix; $\overset{p}{\to}$ and $\overset{d}{\to}$ denote respectively convergence in probability and in distribution and $\Rightarrow$ denotes weak convergence as defined by Pollard (1984, pp. 64-66). Finally, C denotes a generic positive constant that may differ according to its use.

Assumption 1.1. *$\{x_{T,t} : t \leq T, T \geq 1\}$ is a triangular array of $\mathbf{X}$-valued rv's that is L^0-near epoch dependent (NED) on a strong mixing base $\{Y_{T,t} : t = \ldots, 0, 1, \ldots; T \geq 1\}$, where $\mathbf{X}$ is a Borel subset of R^k, and $\{\mu_{T,t} : T \geq 1\}$ is tight on $\mathbf{X}$*[1].

Define the smoothed moment conditions as:[2]

$$g_{tT}(\beta, \delta) = \frac{1}{M_T} \sum_{m=t-T}^{t-1} k\left(\frac{m}{M_T}\right) g\left(x_{T,t-m}, \beta, \delta\right)$$

for an appropriate kernel and M_T is a bandwidth parameter. From now on, we consider the uniform kernel proposed by Kitamura and Stutzer (1997):

$$g_{tT}(\beta, \delta) = \frac{1}{2K_T + 1} \sum_{m=-K_T}^{K_T} g\left(x_{T,t-m}, \beta, \delta\right).$$

Assumption 1.2. *$K_T/T^2 \to 0$ and $K_T \to \infty$ as $T \to \infty$ and $K_T = O_p\left(T^{\frac{1}{2\eta}}\right)$ for some $\eta > 1$*[3].

Assumption 1.3. *For some $d > \max\left(2, \frac{2\eta}{\eta-1}\right)$, $\{g\left(x_{T,t}, \beta_0, \delta_0\right) : t \leq T, T \geq 1\}$ is a triangular array of mean zero R^q-valued rv's that is L^2-near epoch dependent of size $-\frac{1}{2}$ on a strong mixing base $\{Y_{T,t} : t = \ldots, 0, 1, \ldots; T \geq 1\}$, of size $-d/(d-2)$ and $\sup \|g\left(x_{T,t}, \beta_0, \delta_0\right)\|^d < \infty$.*

Assumption 1.4. *$Var\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{Ts} g\left(x_{T,t}, \beta_0, \delta_0\right)\right) \to s\Omega \;\; \forall s \in [0, 1]$ for some positive definite $q \times q$ matrix Ω.*

The above assumptions are sufficient to yield weak convergence of the standardized partial sum of the smoothed moment conditions under the null and the alternatives (see Lemmas 1.1 and 1.2). In the following, x_t is used to denote $x_{T,t}$ for notational simplicity.

[1]For a definition of L^p-near epoch dependence and tightness, see Andrews (1993, p. 829-830). For a presentation of the concept of near epoch dependence, we refer the reader to Gallant and White (1988) (chapters 3 and 4).

[2]Note here that g_{tT} denotes the smoothed moment conditions and $x_{T,t}$ a triangular array of random variables.

[3]This assumption is slightly different than that in Smith, 2011 but facilitates the proofs at no real cost.

Assumption 1.5. *$\tilde{g}(\beta_0,\delta_0)=0$ with $(\beta_0,\delta_0)\in B\times\Delta$ where $\tilde{g}(\beta_0,\delta_0)=\lim_{T\to\infty}\sum_{t=1}^{T}Eg(x_t,\beta,\delta)$ and B and Δ are bounded subsets of R^r and R^ν, $g(x_t,\beta,\delta)$ is continuous in x for all $(\beta,\delta)\in B\times\Delta$ and is continuous in (β,δ) uniformly over $(\beta,\delta,x)\in B\times\Delta\times\zeta$ for all compact sets $\zeta\subset\mathbf{X}$.*

Assumption 1.6. *For every neighborhood $\Theta_0\subset\Theta$ of θ_0, $\inf_{s\in S}\left(\inf_{\theta\in\Theta/\Theta_0}\|g(\theta,s)\|\right)>0$ where $g(\theta,s)=(s\tilde{g}(\beta_1,\delta)',(1-s)\tilde{g}(\beta_2,\delta)')'$.*

Assumption 1.7. ***(a)** $\rho(\cdot)$ is twice continuously differentiable and concave on its domain, an open interval Φ containing 0, $\rho_1=\rho_2=-1$; **(b)** $\lambda(s)\in\widehat{\Lambda}_T(s)$ where $\widehat{\Lambda}_T(s)=\{\lambda(s):\|\lambda(s)\|\leq D\left(T/((2K_T+1)^2\right)^{-\zeta}\}$ for some $D>0$ with $\frac{1}{2}>\zeta>\frac{1}{d(1-1/\eta)}$.*

Assumption 1.7 (b) parallels the assumption in Newey and Smith, 2011 and Smith, 2011 but for $\lambda(s)=(\lambda_1',\lambda_2')'$. It specifies bounds on $\lambda(s)$ and with the existence of higher than second moments in Assumption 1.3 leads to the arguments $\lambda(s)'g_{tT}(\theta,s)$ being in the domain Φ of $\rho(\cdot)$ w.p.a.1 in the first subsample for all β_1,δ and $1\leq t\leq[Ts]$ and in the second subsample for all β_2,δ and $[Ts]+1\leq t\leq T$ (see Lemma 1.3).

Under Assumptions 1.1, 1.2, 1.3, 1.5, 1.6 and 1.7, we show for the partial-sample GEL estimator that $\sup_{s\in S}\|\hat{\theta}_T(s)-\theta_0\|\overset{p}{\to}0$, $\sup_{s\in S}\|\hat{\lambda}_T(s)\|\overset{p}{\to}0$, $\|\hat{\lambda}_T(s)\|=O_p\left(T/(2K_T+1)^2\right)^{-1/2}$ and $\sup_{s\in S}\|\frac{1}{T}\sum_{t=1}^{T}g_{tT}(\hat{\theta}_T(s),s)\|=O_p(T^{-1/2})$.

The consistency of the full-sample GEL estimator is obtained by slight modifications of Assumptions 1.6 and 1.7 (b). Assumption 1.6 must be modified by a simplified version with $\tilde{g}(\beta,\delta)$ instead of $g(\theta,s)$. Assumption 1.7 (b) holds but for the full-sample Lagrange multiplier λ. The consistency result that $\tilde{\theta}_T\overset{p}{\to}\theta_0$ is then derived under weaker conditions than in Smith, 2011.

The following high level assumptions are sufficient to derive the weak convergence under the null of the PS-GEL estimators $\hat{\theta}_T(s)$ and $\hat{\lambda}_T(s)$. These assumptions are similar to the ones in Andrews (1993).

Assumption 1.8. $\sup_{s\in S}\|\widehat{\Omega}_{iT}(s)-\Omega\|\overset{p}{\to}0$ *where Ω is defined in Section 2.1 and S whose closure lies in $(0,1)$ for $i=1,2$.*

Assumption 1.8 holds under conditions given in Andrews (1991) and Lemma A.3 in Smith, 2011. To respect these conditions, Assumption 1.3 can be replaced by the following assumption:

Assumption 6.3′. $\{g(x_t,\beta_0,\delta_0):t\leq T,T\geq 1\}$ is a triangular array of mean zero R^q-valued rv's that is α-mixing with mixing coefficients $\sum_{j=1}^{\infty}j^2\alpha(j)^{(\nu-1)/\nu}<\infty$ for some $\nu>1$ and $sup_{t\leq T,T\geq 1}E\|g(x_t,\beta_0,\delta_0)\|^d<\infty$ for some $d>\max\left(4\nu,\frac{2\eta}{\eta-1}\right)$.

Assumptions 6.3′ and 1.8 guarantee for the full-sample and partial-sample GEL that

$$\widetilde{\Omega}_T=\frac{2K+1}{T}\sum_{t=1}^{T}g_{tT}(\tilde{\beta}_T,\tilde{\delta}_T)g_{tT}(\tilde{\beta}_T,\tilde{\delta}_T)'\overset{p}{\to}\Omega$$

and

$$\widehat{\Omega}_T(s)=\frac{2K+1}{T}\sum_{t=1}^{T}g_{tT}(\hat{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'\overset{p}{\to}\Omega(s).$$

Now, let $G(\beta,\delta) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left[\partial g(x_t,\beta,\delta)/\partial(\beta',\delta')\right]$ and $G = G(\beta_0,\delta_0)$.

Assumption 1.9. *$g(x,\beta,\delta)$ is differentiable in $(\beta,\delta), \forall(\beta,\delta)\in B_0\times\Delta_0$ $\forall x\in \mathbf{X_0}\subset \mathbf{X}$ for a Borel measurable set $\mathbf{X_0}$ that satisfies $P(x_t\in\mathbf{X_0}) = 1 \forall t\le T, T\ge 1$, $g(x,\beta,\delta)$ is Borel measurable in x $\forall(\beta,\delta)\in B_0\times\Delta_0$, $\partial g(x_t,\beta,\delta)/\partial(\beta',\delta')$ is continuous in (x,β,δ) on $\mathbf{X}\times B_0\times\Delta_0$,*

$$\sup_{1\le t\le T} E\left[\sup_{(\beta,\delta)\in B_0\times\Delta_0} \|\partial g(x_t,\beta,\delta)/\partial(\beta',\delta')\|^{d/(d-1)}\right] < \infty$$

and $rank(G) = r+\nu$.

Assumption 1.10. *$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{Ts} Eg_{tT}(\beta,\delta)$ exists uniformly over $(\beta,\delta,s)\in B\times\Delta\times S$ and equals $s\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} Eg(x_t,\beta,\delta) = s\tilde{g}(\beta,\delta)$.*

Assumption 1.11. *$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{Ts} E\partial g_{tT}(\beta_0,\delta_0)/\partial(\beta',\delta')$ exists uniformly over $s\in S$ and equals sG $\forall s\in S$ and S whose closure lies in $(0,1)$.*

Assumption 1.12. *$G(s)'\Omega(s)^{-1}G(s)$ is nonsingular $\forall s\in S$ and has eigenvalues bounded away from zero $\forall s\in S$ and S whose closure lies in $(0,1)$.*

Assumptions 1.10 and 1.11 are asymptotic covariance stationary conditions and follow directly from $Eg_{tT}(\beta,\delta) = Eg(x_t,\beta,\delta)+o_p(1)$ and $E\partial g_{tT}(\beta_0,\delta_0)/\partial(\beta',\delta') = E\partial g(x_t,\beta_0,\delta_0)/\partial(\beta',\delta')+o_p(1)$ for the uniform kernel. Assumption 1.12 guarantees that the partial-sample GEL estimators $\hat{\theta}_T(s)$ has a well defined asymptotic variance $\forall s\in S$ and holds if G^β and G^δ are full rank.

2 Lemmas

Lemma 2.1. *Under Assumptions 1.1 to 1.4, the asymptotic distribution of the smoothed moment conditions under the null is given by:*

$$\Omega^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0,\delta_0) \Rightarrow B(s),$$

where $B(s)$ is a q-dimensional vector of standard Brownian motion.

Proof of Lemma 2.1

First, under Assumptions 1.1, 1.3 and 1.4 , Lemma A4 in Andrews (1993) implies:

$$\Omega^{-1/2}\sum_{t=1}^{[Ts]} g(x_t,\beta_0,\delta_0) \Rightarrow B(s)$$

where $B(s)$ is a q-vector of standard Brownian motion.

Second, the smoothed moment condition are defined as:

$$\Omega^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{2K_T+1}\sum_{m=-K_T}^{K_T} g(x_{t-m},\beta_0,\delta_0).$$

Considering the "endpoint effect" introduced by the extra K_T terms, we have:

$$\begin{aligned}
\sum_{t=1}^{[Ts]} \sum_{m=-K_T}^{K_T} \frac{1}{2K_T+1} g\left(x_{t-m}, \beta_0, \delta_0\right) &= \sum_{t=1}^{[Ts]} \frac{1}{2K_T+1} \sum_{m=\max\{t-[Ts],-K_T\}}^{\min\{t-1,K_T\}} g\left(x_{t-m}, \beta_0, \delta_0\right) \\
&= \sum_{t=K_T+1}^{[Ts]-K_T} g(x_t, \beta_0, \delta_0) + \sum_{t=1}^{K_T} \frac{t+K_T}{2K_T+1} g(x_t, \beta_0, \delta_0) \\
&\quad + \sum_{t=[Ts]-K_T+1}^{[Ts]} \frac{[Ts]-t+K_T+1}{2K_T+1} g(x_t, \beta_0, \delta_0) \\
&= \sum_{t=1}^{[Ts]} g(x_t, \beta_0, \delta_0) + \sum_{t=1}^{K_T} \frac{t-K_T-1}{2K_T+1} g(x_t, \beta_0, \delta_0) \\
&\quad + \sum_{t=[Ts]-K_T+1}^{[Ts]} \frac{[Ts]-t-K_T}{2K_T+1} g(x_t, \beta_0, \delta_0)
\end{aligned}$$

which implies that

$$\sum_{t=1}^{[Ts]} g\left(x_t, \beta_0, \delta_0\right) = \sum_{t=1}^{[Ts]} g_{tT}(\beta_0, \delta_0) + O_p\left(\frac{K_T^2}{2K_T+1}\right).$$

Under the Assumptions that $\max_{1 \le t \le T} \|g(x_t, \beta_0, \delta_0)\| = o_p(T^{1/2})$ and $K_T^2/T \to 0$, we get

$$\Omega^{-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} g\left(x_t, \beta_0, \delta_0\right) = \Omega^{-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} g_{tT}(\beta_0, \delta_0) + o_p(1)$$

which yields the asymptotic equivalence.

The following Lemma provides the asymptotic distribution of the smoothed moment condition under the general sequence of local alternatives appearing in (**??**).

Lemma 2.2. *Under the alternative (**??**), Assumptions 1.1, 1.2, 1.4 and replacing $g(x_t, \beta_0, \delta_0)$ by $g(x_t, \beta_0, \delta_0) - h(\eta, \tau, \frac{t}{T})/\sqrt{T}$ in Assumption 1.3, then*

$$\frac{1}{\sqrt{T}} \Omega^{-1/2} \sum_{t=1}^{[Ts]} g_{tT}(\beta_0, \delta_0) \Rightarrow B(s) + \Omega^{-1/2} H(s)$$

where $H(s) = \int_0^s h(\eta, \tau, u) du$ and $B(s)$ is a q-dimensional vectors of standard Brownian motion.

Proof of Lemma 2.2

Under the alternative (**??**) and Lemma 2.1, the sample smoothed moments satisfy:

$$\Omega^{-1/2} \sqrt{T} \left(\frac{1}{T} \sum_{t=1}^{[Ts]} g_{tT}(\beta_0, \delta_0) - E \frac{1}{T} \sum_{t=1}^{[Ts]} g_{tT}(\beta_0, \delta_0) \right) \Rightarrow B(s)$$

which implies under the alternative (??)

$$\Omega^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{2K_T+1}\sum_{m=-K_T}^{K_T}\left(g\left(x_{t-m},\beta_0,\delta_0\right)-\frac{h\left((t-m)/T\right)}{\sqrt{T}}\right)\Rightarrow B(s),$$

where $h(t/T)\equiv h(\eta,\tau,\frac{t}{T})$ to reduce the notation. Now we need to show

$$\Omega^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{2K_T+1}\sum_{m=-K_T}^{K_T}\frac{h\left((t-m)/T\right)}{\sqrt{T}}\rightarrow\Omega^{-1/2}H(s).$$

Let us examine this expression,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{2K_T+1}\sum_{m=-K_T}^{K_T}\frac{h\left((t-m)/T\right)}{\sqrt{T}}=\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{2K_T+1}\sum_{m=\max\{t-[Ts],-K_T\}}^{\min\{t-1,K_T\}}\frac{h\left((t-m)/T\right)}{\sqrt{T}}$$

which equals

$$\begin{aligned}&\frac{1}{T}\sum_{t=K_T+1}^{[Ts]-K_T}h\left(t/T\right)+\frac{1}{T}\sum_{t=1}^{K_T}\frac{t+K_T}{2K_T+1}h\left(t/T\right)+\frac{1}{T}\sum_{t=[Ts]-K_T+1}^{[Ts]}\frac{[Ts]-t+K_T+1}{2K_T+1}h\left(t/T\right)\\&=\frac{1}{T}\sum_{t=1}^{[Ts]}\frac{h\left(t/T\right)}{\sqrt{T}}+\frac{1}{T}\sum_{t=1}^{K_T}\frac{t-K_T-1}{2K_T+1}h\left(t/T\right)+\frac{1}{T}\sum_{t=[Ts]-K_T+1}^{[Ts]}\frac{[Ts]-t-K_T}{2K_T+1}h\left(t/T\right).\end{aligned}$$

The first term of the last equality converges to $\int_0^s h(\nu)d\nu$. Under the assumption that $\frac{K_T^2}{T}\rightarrow 0$, the last two terms converge to zero and using $\Omega^{-1/2}\sqrt{T}\frac{1}{T}\sum_{t=1}^{[Ts]}g_{tT}(\beta_0,\delta_0)=\Omega^{-1/2}\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{[Ts]}g_{tT}(\beta_0,\delta_0)-E\frac{1}{T}\sum_{t=1}^{[Ts]}g_{tT}(\beta_0,\delta_0)\right)+$ $\Omega^{-1/2}\sqrt{T}E\frac{1}{T}\sum_{t=1}^{[Ts]}g_{tT}(\beta_0,\delta_0)$, the result follows.

Lemma 2.3. *Under the null and Assumptions 1.3 and 1.7,*

$$\sup_{s\in S}\sup_{\theta\in\Theta,\lambda(s)\in\widehat{\Lambda}_T(s),1\leq t\leq T}\left|\lambda(s)'g_{tT}\left(\theta,s\right)\right|\xrightarrow{p}0.$$

Also w.p.a.1 $\widehat{\Lambda}_T(s)\subseteq\widehat{\Lambda}_T\left(\theta,s\right)$ *where* $\widehat{\Lambda}_T\left(\theta,s\right)=\{\lambda(s)=(\lambda_1',\lambda_2')'\;:\;\lambda(s)'g_{tT}\left(\theta,s\right)\in\Phi,(t=1,\ldots,T)\}$.

Proof of Lemma 2.3

We first show that the results hold for both subsambles for a given s. Let $\widehat{\Lambda}_T(s)=\widehat{\Lambda}_{1T}(s)$ for $t=1,\ldots,[Ts])$ and $\widehat{\Lambda}_T(s)=\widehat{\Lambda}_{2T}(s)$ for $t=[Ts]+1,\ldots,T)$. So, we have

$$\begin{aligned}\sup_{\theta\in\Theta_T,\lambda\in\widehat{\Lambda}_T(s),1\leq t\leq T}\left|\lambda(s)'g_{tT}\left(\theta,s\right)\right|&\leq\sup_{\beta\in B,\delta\in\Delta,\lambda_1\in\widehat{\Lambda}_{1T}(s),1\leq t\leq[Ts]}\left|\lambda_1'g_{tT}\left(\beta,\delta\right)\right|\\&+\sup_{\beta\in B,\delta\in\Delta,\lambda_2\in\widehat{\Lambda}_{2T}(s),[Ts]+1\leq t\leq T}\left|\lambda_2'g_{tT}\left(\beta,\delta\right)\right|.\end{aligned}$$

Consider the first subsample, by the Cauchy-Schwarz inequality and Assumption 1.7 (b)

$$\sup_{\beta\in B,\lambda_1\in\widehat{\Lambda}_{1T}(s),1\leq t\leq[Ts]}\left|\lambda_1'g_{tT}\left(\beta,\delta\right)\right|\leq D\left(T/(2K_T+1)^2\right)^{-\zeta}\sup_{\beta\in B,\delta\in\Delta,1\leq t\leq[Ts]}\left\|g_{tT}\left(\beta,\delta\right)\right\|.$$

For the last term on the RHS, we get

$$
\begin{aligned}
\sup_{\beta\in B,\delta\in\Delta,1\leq t\leq[Ts]}\|g_{tT}(\beta,\delta)\| &\leq \frac{1}{2K+1}\sup_{\beta\in B,\delta\in\Delta,1\leq t\leq[Ts]}\left\|\sum_{m=\max\{t-[Ts],-K_T\}}^{\min\{t-1,K_T\}} g(x_{t-m},\beta,\delta)\right\| \\
&\leq \sup_{\beta\in B,\delta\in\Delta,1\leq t\leq[Ts]}\|g(x_t,\beta,\delta)\|
\end{aligned}
$$

uniformly in t. Using Assumption 1.3 and by Markov's inequality:

$$
\sup_{\beta\in B,\delta\in\Delta,1\leq t\leq T}\|g(x_t,\beta,\delta)\| = O_p\left(T^{1/d}\right).
$$

Hence

$$
\sup_{\beta\in B,\delta\in\Delta,\lambda_1\in\widehat{\Lambda}_{1T}(s),1\leq t\leq[Ts]}|\lambda_1' g_{tT}(\beta,\delta)| \quad\leq\quad D\left(T/(2K_T+1)^2\right)^{-\zeta} O_p\left(T^{1/d}\right)\xrightarrow{p}0
$$

by Assumption 1.7 (b). This also holds for the second subsample.

Therefore under the null

$$
\sup_{\beta\in B,\delta\in\Delta,\lambda_1\in\widehat{\Lambda}_{1T}(s),1\leq t\leq[Ts]}|\lambda_1' g_{tT}(\beta,\delta)|\xrightarrow{p}0
$$

and $\lambda_1' g_{tT}\in\Phi$ for $t=1,\ldots,[Ts]$ w.p.a.1 for all $\beta\in B$, $\delta\in\Delta$ which implies that $\lambda_1\in\widehat{\Lambda}_{1T}(\beta,\delta,s)$. For the second subsample,

$$
\sup_{\beta\in B,\delta\in\Delta,\lambda_2\in\widehat{\Lambda}_{2T}(s),[Ts]+1\leq t\leq T}|\lambda_2' g_{tT}(\beta,\delta)|\xrightarrow{p}0
$$

and $\lambda_2' g_{tT}\in\Phi$ for $t=[Ts]+1,\ldots,T$ w.p.a.1 for all $\beta\in B$, $\delta\in\Delta$ which implies that $\lambda_2\in\widehat{\Lambda}_{2T}(\beta,\delta,s)$. Finally, these results holds uniformly $\forall s\in S$.

Lemma 2.4. *Under Assumptions 1.1, 1.2, 1.3, 1.5 and 1.10*

$$
\sup_{s\in S}\sup_{\theta\in\Theta}\|\hat{g}_T(\theta,s)-g(\theta,s)\|\xrightarrow{p}0
$$

where

$$
\hat{g}_T(\theta,s)=\frac{1}{T}\sum_{t=1}^{T}g_{tT}(\theta,s)=\frac{1}{T}\sum_{t=1}^{[Ts]}\begin{bmatrix} g_{tT}(\beta,\delta)\\ 0\end{bmatrix}+\frac{1}{T}\sum_{t=[Ts]+1}^{T}\begin{bmatrix}0\\ g_{tT}(\beta,\delta)\end{bmatrix}
$$

and $g(\theta,s)=(s\tilde{g}(\beta,\delta)',(1-s)\tilde{g}(\beta,\delta)')'$.

Proof of Lemma 2.4

Using $\sum_{[Ts]+1}^{T}=\sum_1^T-\sum_1^{[Ts]}$, the result of the Lemma holds if

$$
\sup_{\beta\in B,\delta\in\Delta}\ \sup_{T\underline{s}\leq R\leq T}\left|\frac{1}{T}\sum_{t=1}^{R}\left[g_{tT}(\beta,\delta)-\tilde{g}(\beta,\delta)\right]\right|\xrightarrow{p}0
$$

where $\underline{s} = inf\{s \ : \ s \in S\}$.

By the triangular inequality

$$\sup_{\theta\in\Theta}\sup_{T\underline{s}\leq R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[g_{tT}(\beta,\delta)-\tilde{g}(\beta,\delta)\right]\right\| \leq \sup_{\theta\in\Theta}\sup_{T\underline{s}\leq R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[g_{tT}(\beta,\delta)-Eg_{tT}(\beta,\delta)\right]\right\| + \sup_{\theta\in\Theta}\sup_{T\underline{s}\leq R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[Eg_{tT}(\beta,\delta)-\tilde{g}(\beta,\delta)\right]\right\|.$$

We show that both terms on the right-hand side converge in probability to zero. For the first term, we first show that $\frac{1}{T}\sum_{t=1}^{R}\left[g_{tT}(\beta,\delta)-Eg_{tT}(\beta,\delta)\right] = \frac{1}{T}\sum_{t=1}^{R}\left[g(x_t,\beta,\delta)-Eg(x_t,\beta,\delta)\right]+o_p(1)$. By the proof similar to the one in Lemma 2.1, we can show that: $\frac{1}{T}\sum_{t=1}^{R}g(x_t,\beta,\delta) = \frac{1}{T}\sum_{t=1}^{R}g_{tT}(\beta,\delta)+o_p(1)$. This also holds for the partial sum of the expectation, the result follows. Now using Lemma A3 in Andrews with Assumptions 1.1 and 1.7 guarantees the UWL for $\sup_{R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[g(x_t,\beta,\delta)-Eg(x_t,\beta,\delta)\right]\right\|$. This yields

$$\sup_{\theta\in\Theta}\sup_{T\underline{s}\leq R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[g(x_t,\beta,\delta)-Eg(x_t,\beta,\delta)\right]\right\| \xrightarrow{p} 0$$

which directly implies

$$\sup_{\theta\in\Theta}\sup_{T\underline{s}\leq R\leq T}\left\|\frac{1}{T}\sum_{t=1}^{R}\left[g_{tT}(\beta,\delta)-Eg_{tT}(\beta,\delta)\right]\right\| \xrightarrow{p} 0.$$

The second term holds by Assumption 1.10.

Now define

$$\begin{aligned}\widehat{P}\left(\theta(s),\lambda(s),s\right) &= \sum_{t=1}^{T}\frac{\left[\rho(\lambda(s)'g_{tT}(\theta,s))-\rho_0\right]}{T}\\ &= \sum_{t=1}^{[Ts]}\frac{\left[\rho(\lambda_1'g_{tT}(\beta_1,\delta))-\rho_0\right]}{T}+\sum_{t=[Ts]+1}^{T}\frac{\left[\rho(\lambda_2'g_{tT}(\beta_2,\delta))-\rho_0\right]}{T}\end{aligned}$$

and $\hat{g}_T(\theta_0,s) = \frac{1}{T}\sum_{t=1}^{T}g_{tT}(\theta_0,s)$.

Lemma 2.5. *Under Assumptions 1.3, 1.7 and 1.8, there is a constant C such that w.p.a.1.*

$$\frac{1}{2K_T+1}\sup_{s\in S}\sup_{\lambda(s)\in\widehat{\Lambda}_T(\theta_0,s)}\widehat{P}\left(\theta_0,\lambda(s),s\right) = \sup_{s\in S}C\|\hat{g}_T(\theta_0,s)\|^2.$$

Proof of Lemma 2.5

By a proof similar to the one of Lemma A.5 in Smith, 2011[4], we can show that

$$\frac{1}{2K_T+1}\sup_{\lambda(s)\in\widehat{\Lambda}_T(\theta_0,s)}\widehat{P}\left(\theta_0,\lambda(s),s\right) = C\|\hat{g}_T(\theta_0,s)\|^2$$

[4]In his proof, Smith, 2011 uses the fact that $(2K_T+1)\sum_{t=1}^{T}\rho_2\left(\dot{\lambda}'g_{tT}(\beta_0,\delta_0)\right)g_{tT}(\beta_0,\delta_0)g_{tT}(\beta_0,\delta_0)'/T \xrightarrow{p} -\Omega$ in our notation. This needs more restrictive assumptions than those imposed here. In fact, we only need that $(2K_T+1)\sum_{t=1}^{T}\rho_2\left(\dot{\lambda}'g_{tT}(\beta_0,\delta_0)\right)g_{tT}(\beta_0,\delta_0)g_{tT}(\beta_0,\delta_0)'/T \leq -CI_q$ in the p.s.d. sense w.p.a.1 which holds by the fact that the outer product of smoothed moment conditions is automatically positive semi-definite.

for a given $s \in S$ w.p.a.1. Since this holds for all $s \in S$, this holds for s which achieves the supremum.

2.1 Proofs of Theorems

Proof of Theorem 2.1

The outline of the proof is similar to that of Lemma A.6 and Theorem 2.2 in Smith, 2011 except that the results have to be established uniformly in $s \in S$ and by taking into account of the differences in Assumptions 1.2, 1.3 and 1.7 with respect to the corresponding assumptions in Smith, 2011.

First, we show that $\sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s), s)\|^2 = O_p(T^{-1})$ which allows us to show that $\sup_{s\in S} \|\hat{\theta}(s) - \theta_0\| \xrightarrow{p} 0$. By arguments similar to Smith, 2011, we can show that $\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T(s), s) g_{tT}(\hat{\theta}_T(s), s)'/T = O_p(1)$. Following Newey and Smith (2001) and Smith, 2011, let

$$\bar{\lambda}_T(\hat{\theta}_T(s), s) = \left(\bar{\lambda}_{1T}(\hat{\beta}_{1T}(s), \hat{\delta}_T(s), s)', \bar{\lambda}_{2T}(\hat{\beta}_{2T}(s), \hat{\delta}_T(s), s)'\right)' = -\frac{1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T(s), s)\kappa_T / \|\hat{g}_T(\hat{\theta}_T(s), s)\|$$

with $\kappa_T = D\left(T/(2K_T+1)^2\right)^{-\zeta}$ and

$$\begin{aligned}
\bar{\lambda}_{1T}(\hat{\beta}_{1T}(s), \hat{\delta}_T(s), s) &= -\frac{1}{T}\sum_{t=1}^{[Ts]} g_{tT}(\hat{\beta}_{1T}(s), \hat{\delta}_T(s))\kappa_T / \|\hat{g}_T(\hat{\theta}_T(s), s)\|, \\
\bar{\lambda}_{2T}(\hat{\beta}_{2T}(s), \hat{\delta}_T(s), s) &= -\frac{1}{T}\sum_{t=[Ts]+1}^{T} g_{tT}(\hat{\beta}_{2T}(s), \hat{\delta}_T(s))\kappa_T / \|\hat{g}_T(\hat{\theta}_T(s), s)\|
\end{aligned}$$

and writing $\bar{\lambda}_T(s) = \bar{\lambda}_T(\hat{\theta}_T(s), s)$, $\bar{\lambda}_{1T}(s) = \bar{\lambda}_{1T}(\hat{\beta}_{1T}(s), \hat{\delta}_T(s), s)$ and $\bar{\lambda}_{2T}(s) = \bar{\lambda}_{2T}(\hat{\beta}_{2T}(s), \hat{\delta}_T(s), s)$ to simplify the notation. By Lemma 2.3, $\sup_{s\in S} \max_{1\le t\le T} |\bar{\lambda}(s)' g_{tT}\left(\hat{\theta}_T(s), s\right)| \xrightarrow{p} 0$ and $\bar{\lambda}_T(s) \in \Lambda_T(\hat{\theta}_T(s), s)$ w.p.a.1. Thus, for a given s, $\dot{\lambda}_T(s) = \left(\dot{\lambda}_{1T}', \dot{\lambda}_{2T}'\right)'$ with $\dot{\lambda}_{1T} = \tau_1 \bar{\lambda}_{1T}, 0 \le \tau_1 \le 1$ and $\dot{\lambda}_{2T} = \tau_2 \bar{\lambda}_{2T}, 0 \le \tau_2 \le 1$,

$$\sup_{s\in S} \sum_{t=1}^{T} \left[\rho_2(\dot{\lambda}_T(s)' g_{tT}(\hat{\theta}_T(s), s)) - \rho_2(0)\right] g_{tT}(\hat{\theta}_T(s), s) g_{tT}(\hat{\theta}_T(s), s)'/T \xrightarrow{p} 0$$

and therefore $\sup_{s\in S}(2K_T+1)\sum_{t=1}^{T} \rho_2(\dot{\lambda}_T(s)' g_{tT}(\hat{\theta}_T(s), s)) g_{tT}(\hat{\theta}_T(s), s) g_{tT}(\hat{\theta}_T(s), s)'/T \ge -C I_{2q}$ in the p.s.d. sense w.p.a.1. Hence, by a second-order Taylor expansion

$$\begin{aligned}
\frac{1}{2K_T+1}\widehat{P}(\hat{\theta}_T(s), \bar{\lambda}_T(s), s) &= -\left(\frac{\bar{\lambda}_T(s)}{2K_T+1}\right)' \hat{g}_T(\hat{\theta}_T(s), s) \\
&+ \left(\frac{\bar{\lambda}_T(s)}{2K_T+1}\right)' \left(\sum_{t=1}^{T} \rho_2(\dot{\lambda}_T(s)' g_{tT}(\hat{\theta}_T(s), s) g_{tT}(\hat{\theta}_T(s), s) g_{tT}(\hat{\theta}_T(s), s)'/T\right) \bar{\lambda}_T(s)/2 \\
&\ge -\left(\frac{\bar{\lambda}_T(s)}{2K_T+1}\right)' \hat{g}_T(\hat{\theta}_T(s), s) - C\left(\frac{\bar{\lambda}_T(s)}{2K_T+1}\right)' \left(\frac{\bar{\lambda}_T(s)}{2K_T+1}\right) \\
&= \|\hat{g}_T(\hat{\theta}_T(s), s)\| \left(\frac{\kappa_T}{2K_T+1}\right) - C\left(\frac{\kappa_T}{2K_T+1}\right)^2
\end{aligned}$$

w.p.a.1 and this holds $\forall s \in S$. Now using Lemma 2.5, we get w.p.a.1

$$\begin{aligned} \sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s),s)\| \left(\frac{\kappa_T}{2K_T+1}\right) - C\left(\frac{\kappa_T}{2K_T+1}\right)^2 &\leq \sup_{s\in S} \frac{1}{2K_T+1} \widehat{P}\left(\hat{\theta}_T(s), \bar{\lambda}_T(s), s\right) \\ &\leq \sup_{s\in S} \sup_{\lambda(s)\in\widehat{\Lambda}_T(\hat{\theta}_T(s),s)} \frac{1}{2K_T+1} \widehat{P}\left(\hat{\theta}_T(s), \lambda(s), s\right) \\ &\leq \sup_{s\in S} \sup_{\lambda(s)\in\widehat{\Lambda}_T(\theta_0,s)} \frac{1}{2K_T+1} \widehat{P}\left(\theta_0, \lambda(s), s\right) \\ &\leq \sup_{s\in S} C\|\hat{g}_T(\theta_0,s)\|^2 = O_p\left(T^{-1}\right) \end{aligned}$$

as $\|\hat{g}_T(\theta_0,s)\| = O_p(T^{-1/2})$ by CLT (Corollary 3.1 of Wooldridge and White, 1988). This yields

$$\sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s),s)\| \leq C\left(\frac{\kappa_T}{2K_T+1}\right) + \sup_{s\in S} C\|\hat{g}_T(\theta_0,s)\|^2 \left(\frac{\kappa_T}{2K_T+1}\right) = O_p\left(\kappa_T/(2K_T+1)\right),$$

which implies $\sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s),s)\| = O_p(T^{-1/2})$ by Assumption 1.2 for all $\eta > 1$. By the result that $\sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s),s\| = O_p(T^{-1/2})$ we have $\sup_{s\in S} \hat{g}_T(\hat{\theta}_T(s),s) \xrightarrow{p} 0$. By Lemma 2.4, $\sup_{s\in S}\sup_{\theta\in\Theta} \|\hat{g}_T(\theta,s) - g(\theta,s)\| \xrightarrow{p} 0$ and $\tilde{g}(\beta,\delta)$ is continuous by Assumption 1.5. The triangular inequality then gives that $\sup_{s\in S} g(\hat{\theta}_T(s),s) \xrightarrow{p} 0$. Since $\tilde{g}(\beta,\delta) = 0$ has a unique zero at β_0 and δ_0 (by Assumption 1.6), for every neighborhood $\Theta_0(\in\Theta)$ of θ_0, $\inf_{s\in S}\left(\inf_{\theta\in\Theta/\Theta_0} \|g(\theta,s)\|\right) > 0$, then $\sup_{s\in S} \|\hat{\theta}_T(s) - \theta_0\| \xrightarrow{p} 0$.

Now we need to show $\sup_{s\in S} \|\hat{\lambda}_T(s)\| = O_p\left(\left(T/(2K+1)^2\right)^{-1/2}\right)$and $\sup_{s\in S} \|\hat{\lambda}_T(s)\| \xrightarrow{p} 0$. By a second-order Taylor expansion around $\lambda(s) = 0$, for a given $s \in S$ and for any $\dot{\lambda}_T(s) = \left(\dot{\lambda}_{1T}', \dot{\lambda}_{2T}'\right)'$ with $\dot{\lambda}_{1T} = \tau_1\hat{\lambda}_{1T}, 0 \leq \tau_1 \leq 1$ and $\dot{\lambda}_{2T} = \tau_2\hat{\lambda}_{2T}, 0 \leq \tau_2 \leq 1$

$$\begin{aligned} (2K_T+1)\widehat{P}\left(\hat{\theta}_T(s),0,s\right) &\leq \sup_{\lambda(s)\in\widehat{\Lambda}_T(\hat{\theta}_T(s),s)} (2K_T+1)\widehat{P}\left(\hat{\theta}_T(s),\lambda(s),s\right) \\ &= (2K_T+1)\widehat{P}(\hat{\theta}_T(s),\hat{\lambda}_T(s),s) \\ &\leq -(2K_T+1)\hat{\lambda}_T(s)'\hat{g}_T(\hat{\theta}_T(s),s) \\ &+ \hat{\lambda}_T(s)'\left((2K_T+1)\sum_{t=1}^{T}\rho_2(\dot{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s))\hat{g}_{tT}(\hat{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'/T\right)\hat{\lambda}_T(s)/2 \\ &\leq -(2K_T+1)\hat{\lambda}_T(s)'g_T(\hat{\theta}_T(s),s) - C\hat{\lambda}_T(s)'\hat{\lambda}_T(s) \\ &\leq (2K_T+1)\|\hat{\lambda}_T(s)\|\|\hat{g}_T(\hat{\theta}_T(s),s)\| - C\|\hat{\lambda}_T(s)\|^2 \end{aligned}$$

w.p.a.1. Since $\widehat{P}\left(\hat{\theta}_T(s),0,s\right) = 0, \forall s \in S$, this implies directly that $C\|\hat{\lambda}_T(s)\| \leq (2K_T+1)\|\hat{g}_T(\hat{\theta}_T(s),s)\|$ and this holds for all $s \in S$ which implies that $\sup_{s\in S} C\|\hat{\lambda}_T(s)\| \leq \sup_{s\in S}(2K_T+1)\|\hat{g}_T(\hat{\theta}_T(s),s)\|$. Finally, considering that $\sup_{s\in S} \|\hat{g}_T(\hat{\theta}_T(s),s)\| = O_p(T^{-1/2})$ directly yields that $\sup_{s\in S} \|\hat{\lambda}_T(s)\| = O_p\left[\left(T/(2K_T+1)^2\right)^{-1/2}\right]$ and $\sup_{s\in S} \|\hat{\lambda}_T(s)\| \xrightarrow{p} 0$ by Assumption 1.2.

Proof of Theorem 2.2

The first-order conditions of the partial-sample GEL with respect to $\lambda(s)$ and $\theta(s)$ are:

$$\frac{1}{T}\sum_{t=1}^{T}\rho_1(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s))g_{tT}(\hat{\theta}_T(s),s)=0$$

$$\frac{1}{T}\sum_{t=1}^{T}\rho_1(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s))G_{tT}(\hat{\theta}_T(s),s)'\hat{\lambda}_T(s)=0.$$

By a mean-value expansion of the former first-order conditions for the partial-sample GEL where $\Xi_T=\left(\hat{\beta}_{1T}(s)',\hat{\beta}_{2T}(s)',\hat{\delta}_T(s)',\frac{\hat{\lambda}_{1T}(s)'}{2K_T+1},\frac{\hat{\lambda}_{2T}(s)'}{2K_T+1}\right)'$ and $\Xi_0=(\beta_0',\beta_0',\delta_0',0,0)'$ with the latter first-order conditions yields:

$$0=-T^{1/2}\left(\begin{array}{c}0\\ \frac{1}{T}\sum_{t=1}^{T}g_{tT}(\theta_0,s)\end{array}\right)+\bar{M}(s)T^{1/2}\left(\hat{\Xi}_T(s)-\Xi_0\right)$$

where

$$\bar{M}(s)=\frac{1}{T}\sum_{t=1}^{T}\left[\begin{array}{cc}0 & \bar{M}_{12}(s)\\ \bar{M}_{21}(s) & \bar{M}_{22}(s)\end{array}\right]$$

with $\bar{M}_{12}(s)=\rho_1\left(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)G_{tT}(\hat{\theta}_T(s),s)'$, $\bar{M}_{21}(s)=\rho_1\left(\bar{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)G_{tT}(\bar{\theta}_T(s),s)'$ and $\bar{M}_{22}(s)=(2K_T+1)\,\rho_2\left(\bar{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)g_{tT}(\bar{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'$ and $\bar{\theta}_T(s)$ is a random vector on the line segment joining $\hat{\theta}_T(s)$ and θ_0 and $\bar{\lambda}_T(s)$ is a random vector joining $\hat{\lambda}_T(s)$ to $(0',0')'$ that may differ from row to row.

Now, we need to show that $\bar{M}(s)\xrightarrow{p}M(s)$ where

$$M(s)=-\left[\begin{array}{cc}0 & G(s)'\\ G(s) & \Omega(s)\end{array}\right].$$

By Lemma 2.3; $\sup_{s\in S}\sup_{1\le t\le T}|\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)|\xrightarrow{p}0$ and $\sup_{s\in S}\sup_{1\le t\le T}|\bar{\lambda}_T(s)'g_{tT}(\bar{\theta}_T(s),s)|\xrightarrow{p}0$ which implies

$$\begin{aligned}\sup_{s\in S}\max_{1\le t\le T}|\rho_1\left(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)-\rho_1(0)| &\xrightarrow{p} 0\\ \sup_{s\in S}\max_{1\le t\le T}|\rho_1\left(\bar{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)-\rho_1(0)| &\xrightarrow{p} 0\end{aligned}$$

and $\sup_{s\in S}\max_{1\le t\le T}|\rho_2\left(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)-\rho_2(0)|\xrightarrow{p}0$. To show that

$$\sup_{s\in S}\frac{1}{T}\sum_{t=1}^{T}\rho_1\left(\bar{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)G_{tT}(\bar{\theta}_T(s),s)\xrightarrow{p}-G(s)$$

and

$$\sup_{s\in S}\frac{1}{T}\sum_{t=1}^{T}\rho_1\left(\hat{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)G_{tT}(\hat{\theta}_T(s),s)\xrightarrow{p}-G(s),$$

it remains to show that

$$\sup_{s\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\bar{\theta}_T(s),s)-G(s)\right\|\xrightarrow{p}0 \tag{1}$$

and

$$\sup_{s\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\hat{\theta}_T(s),s)-G(s)\right\|\xrightarrow{p}0. \tag{2}$$

For (1), by the triangular inequality

$$\begin{aligned}\sup_{s\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\bar{\theta}_T(s),s)-G(s)\right\| &\leq \sup_{s\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\bar{\theta}_T(s),s)-E\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\bar{\theta}_T(s),s)\right\|\\ &+ \sup_{s\in S}\left\|E\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\bar{\theta}_T(s),s)-E\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\theta_0,s)\right\|\\ &+ \sup_{s\in S}\left\|E\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\theta_0,s)-G(s)\right\|.\end{aligned}$$

The first term on the right-hand side $\xrightarrow{p} 0$ by an application of UWL given by Lemma A3 in Andrews (1993) which implies

$$\sup_{s\in S}\sup_{\theta\in\Theta}\left\|\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\theta,s)-E\frac{1}{T}\sum_{t=1}^{T}G_{tT}(\theta,s)\right\|\xrightarrow{p}0.$$

The second term $\xrightarrow{p} 0$ under the tightness of $\{\bar{\mu}_T;\ T\geq 1\}$ (Assumption 1.1), Assumption 1.9 and $\bar{\theta}_T(s)\xrightarrow{p}\theta_0$ (see equations A.13 and A.14 in Andrews, 1993). Finally, the third term $\xrightarrow{p} 0$ by Assumption 1.11 and by $\frac{1}{T}\sum_{t=1}^{[Ts]}\frac{\partial g(x_t,\beta,\delta)}{\partial(\beta',\delta')}=\frac{1}{T}\sum_{t=1}^{[Ts]}G_{tT}(\beta,\delta)+o_p(1)$. The proof for (2) is similar.

Moreover, Assumptions 1.8 implies that

$$\frac{2K_T+1}{T}\sum_{t=1}^{[Ts]}g_{tT}(\bar{\beta}_{1T},\bar{\delta}_T)g_{tT}(\hat{\beta}_{1T},\hat{\delta}_T)'\xrightarrow{p}s\Omega$$

and

$$\frac{2K_T+1}{T}\sum_{t=[Ts]+1}^{T}g_{tT}(\bar{\beta}_{2T},\bar{\delta}_T)g_{tT}(\hat{\beta}_{2T},\hat{\delta}_T)'\xrightarrow{p}(1-s)\Omega$$

which yields

$$\frac{2K_T+1}{T}\sum_{t=1}^{T}\rho_2\left(\bar{\lambda}_T(s)'g_{tT}(\hat{\theta}_T(s),s)\right)g_{tT}(\bar{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'\xrightarrow{p}-\Omega(s).$$

By Assumption 1.12, this gives

$$M(s)^{-1}=\begin{bmatrix}-\Sigma(s) & H(s)\\ H(s)' & P(s)\end{bmatrix}$$

where $\Sigma(s) = \left(G(s)'\Omega(s)^{-1}G(s)\right)^{-1}$, $H(s) = \Sigma(s)G(s)'\Omega(s)^{-1}$ and $P(s) = \Omega(s)^{-1} - \Omega(s)^{-1}G(s)\Sigma(s)G(s)'\Omega(s)^{-1}$.

As $\bar{M}(s)$ is positive definite w.p.a.1, we obtain:

$$\begin{aligned}\sqrt{T}\left(\Xi_T(s) - \Xi_0\right) &= -\bar{M}^{-1}(s)\left(0, -\sqrt{T}\frac{1}{T}\sum_{t=1}^{T} g_{tT}(\theta_0, s)'\right) + o_p(1) \\ &= -\left(H(s)', P(s)\right)' \sqrt{T}\frac{1}{T}\sum_{t=1}^{T} g_{tT}(\theta_0, s) + o_p(1).\end{aligned}$$

We also have by Lemma 2.1, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{tT}(\theta_0, s) \Rightarrow J(s)$ for $s \in S$. Combining the results above yields:

$$\begin{aligned}\sqrt{T}\left(\hat{\theta}_T(s) - \theta_0\right) &= -\left(G(s)'\Omega(s)^{-1}G(s)\right)^{-1} G(s)'\Omega(s)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{tT}(\theta_0, s) + o_p(1) \\ &\Rightarrow -\left(G(s)'\Omega(s)^{-1}G(s)\right)^{-1} G(s)'\Omega(s)^{-1}J(s)\end{aligned}$$

and

$$\begin{aligned}\frac{\sqrt{T}}{2K_T + 1}\hat{\lambda}_T(s) &= -\left(\Omega^{-1}(s) - \Omega^{-1}(s)G(s)\left(G(s)'\Omega(s)^{-1}G(s)\right)^{-1} G(s)'\Omega(s)^{-1}\right)\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{tT}(\theta_0, s) + o_p(1) \\ &\Rightarrow -\left(\Omega^{-1}(s) - \Omega^{-1}(s)G(s)\left(G(s)'\Omega(s)^{-1}G(s)\right)^{-1} G(s)'\Omega(s)^{-1}\right) J(s).\end{aligned}$$

Proof of Theorem 2.3

This is a direct implication of Lemma 2.2 and the proof of Theorem **??**.

Proof of Theorem 3.1

Using results derived above, we get for terms in the Wald statistic:

$$\begin{aligned}\widehat{G}^{\beta}_{1,tT}(s) &= \frac{1}{[Ts]}\sum_{t=1}^{[Ts]} \frac{\partial g(x_t, \hat{\beta}_{1T}(s), \hat{\delta}_T(s))}{\partial \beta_1'} + o_p(1), \\ \widehat{G}^{\beta}_{2,tT}(s) &= \frac{1}{T - [Ts]}\sum_{t=[Ts]+1}^{T} \frac{\partial g(x_t, \hat{\beta}_{2T}(s), \hat{\delta}_T(s))}{\partial \beta_2'} + o_p(1), \\ \widehat{\Omega}_{1T}(s) &\xrightarrow{p} \Omega_1(s), \ \ \widehat{\Omega}_{2T}(s) \xrightarrow{p} \Omega_2(s)\end{aligned}$$

and terms in the LM statistic:

$$\begin{aligned}\hat{g}_{1T}(\tilde{\theta}_T, s) &= \frac{1}{T}\sum_{t=1}^{[Ts]} g(x_t, \tilde{\beta}_T, \tilde{\delta}_T) + o_p(1), \\ \widetilde{G}^{\beta}_{tT} &= \frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(x_t, \tilde{\beta}_T, \tilde{\delta}_T)}{\partial \beta'} + o_p(1), \\ \widetilde{\Omega}_T &\xrightarrow{p} \Omega.\end{aligned}$$

The asymptotic distributions for the $Wald_T(s)$ and $LM_T(s)$ under the null can then be directly derived using the expressions above from similar arguments than in the proof of Theorem 3 in Andrews (1993). The asymptotic distribution under the alternative is a direct implication of Theorem **??**. For the $LR_T(s)$ statistic, expanding the partial-sample GEL objective function evaluated at the unrestricted estimator about $\lambda = 0$ yields,

$$\begin{aligned} \frac{2T}{(2K_T+1)}\frac{1}{T}\sum_{t=1}^{T}\rho(\hat{\lambda}_T(\hat{\theta}_T(s),s)'g_{tT}(\hat{\theta}_T(s),s) &= -\frac{2T}{(2K_T+1)}\frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_T(\hat{\theta}_T(s),s)'g_{tT}(\hat{\theta}_T(s),s) - \\ &\quad \frac{T}{(2K_T+1)}\frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_T(\hat{\theta}_T(s),s)'g_{tT}(\hat{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'\hat{\lambda}_T(\hat{\theta}_T(s),s) \\ &+ \quad o_p(1) \end{aligned}$$

since $\rho_1(\cdot) \xrightarrow{p} -1$ and $\rho_2(\cdot) \xrightarrow{p} -1$.

By the fact that $\hat{\Omega}_T(s) = \frac{2K_T+1}{T}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T(s),s)g_{tT}(\hat{\theta}_T(s),s)'$ is a consistent estimator of $\Omega(s)$ and by $\sqrt{T}/(2K_T+1)\hat{\lambda}_T(s) = -\Omega(s)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g_{tT}(\hat{\theta}_T(s),s)) + o_p(1)$, we get

$$\frac{2T}{(2K_T+1)}\frac{1}{T}\sum_{t=1}^{T}\rho(\hat{\lambda}_T(\hat{\theta}_T(s),s)'g_{tT}(\hat{\theta}_T(s),s)) = Tg_T(\hat{\theta}_T(s),s)'\Omega(s)^{-1}g_T(\hat{\theta}_T(s),s) + o_p(1).$$

Similarly, the expansion of the partial-sample GEL objective function but evaluated at the restricted estimator yields:

$$\frac{2T}{(2K_T+1)}\frac{1}{T}\sum_{t=1}^{T}\rho(\hat{\lambda}_T(\tilde{\theta}_T,s)'g_{tT}(\tilde{\theta}_T,s)) = Tg_T(\tilde{\theta}_T,s)'\Omega(s)^{-1}g_T(\tilde{\theta}_T,s) + o_p(1)$$

since that $\tilde{\Omega}_T(s) = \frac{2K_T+1}{T}\sum_{t=1}^{T} g_{tT}(\tilde{\theta}_T,s)g_{tT}(\tilde{\theta}_T,s)'$ is a consistent estimator of $\Omega(s)$ under the null. The $LR_T(s)$ is then asymptotically equivalent to the LR statistic defined in Andrews (1993) for the standard GMM.

Proof of Theorem 3.3

First, for the statistic $O_T(s)$, the asymptotic equivalence between $\sum_{t=1}^{[Ts]} g_{tT}(\hat{\beta}_{1T}(s))$ with $\sum_{t=1}^{[Ts]} g(x_t,\hat{\beta}_{1T}(s))$ and $\sum_{t=[Ts]+1}^{T} g_{tT}(\hat{\beta}_{2T}(s))$ with $\sum_{t=[Ts]+1}^{T} g(x_t,\hat{\beta}_{2T}(s))$ is a direct implication of the Lemmas 2.1 and 2.2 and by the asymptotic consistency of the estimator $\hat{\Omega}_{1T}(s)$ and $\hat{\Omega}_{2T}(s)$ for Ω, the result under the null and alternative follows directly from proofs for Theorems **??** and **??** and subsection A.2 in Hall and Sen (1999).

Second, for the statistic $O_T(s)^{GEL}$, as in the proof of Theorem 3.2, we can show that:

$$\frac{2[Ts]}{2K_T+1}\sum_{t=1}^{[Ts]}\frac{\left[\rho(\hat{\lambda}_{1T}(\hat{\beta}_{1T}(s),s)'g_{tT}(\hat{\beta}_{1T}(s))) - \rho_0\right]}{[Ts]} = O1_T(s) + o_p(1)$$

and

$$\frac{2(T-[Ts])}{2K_T+1}\sum_{t=[Ts]+1}^{T}\frac{\left[\rho(\hat{\lambda}_{2T}(\hat{\beta}_{2T}(s),s)'g_{tT}(\hat{\beta}_{2T}(s))) - \rho_0\right]}{T-[Ts]} = O2_T(s) + o_p(1).$$

The asymptotic distribution under the null and the alternative follows directly.

Finally, for the statistic $LM_T^O(s)$, the have the following asymptotic equivalences:

$$\begin{aligned}
\frac{\sqrt{[Ts]}}{(2K_T+1)}\hat{\lambda}_{1T}(\hat{\beta}_{1T}(s),s) &= -\Omega(s)^{-1}([Ts])^{-1/2}\sum_{t=1}^{[Ts]} g_{tT}(\hat{\beta}_{1T}(s)) + o_p(1) \\
\frac{\sqrt{T-[Ts]}}{(2K_T+1)}\hat{\lambda}_{2T}(\hat{\beta}_{2T}(s),s) &= -\Omega(s)^{-1}(T-[Ts])^{1/2}\sum_{t=[Ts]+1}^{T} g_{tT}(\hat{\beta}_{2T}(s)) + o_p(1)
\end{aligned}$$

which implies directly the asymptotic distribution of this statistic under the null and the alternative.

Proof of Theorem 3.4

Since $\tilde{\theta}_T(s)$ minimizes the restricted partial sample GEL for all $s \in S$, this implies for all $s \in S$ and all T,

$$\widehat{P}(\tilde{\theta}_T(s), \hat{\lambda}_T(\tilde{\theta}_T(s), s) \leq \widehat{P}(\theta_0, \hat{\lambda}_T(\tilde{\theta}_T(s), s), s).$$

The limit for $\widehat{P}(\tilde{\theta}_T(s), \hat{\lambda}_T(\tilde{\theta}_T(s), s), s)$ is then bounded by the limit of $\widehat{P}(\theta_0, \hat{\lambda}_T(\tilde{\theta}_T(s), s), s)$. Let $\hat{\lambda}_T(\theta_0, s) = \arg\max_{\lambda_s \in \widehat{\Lambda}_T(\theta_0,s)} \widehat{P}(\theta_0, \lambda(s), s)$ and $\dot{\lambda}_T(s) = \tau\hat{\lambda}_T(s)$, $0 \leq \tau \leq 1$. Thus, $\widehat{P}(\tilde{\theta}_T(s), \hat{\lambda}_T(\tilde{\theta}_T(s), s), s) \leq \widehat{P}(\theta_0, \hat{\lambda}_T(\tilde{\theta}_T(s), s), s) \leq \widehat{P}(\theta_0, \hat{\lambda}_T(\theta_0, s), s)$. By a second-order Taylor expansion with Lagrange remainder and using $(2K_T+1)\sum_{t=1}^T \rho_2(\dot{\lambda}(s)' g_{tT}(\theta_0, s)) g_{tT}(\theta_0, s) g_{tT}\theta_0, s)'/T \xrightarrow{p} -\Omega(s)$,

$$\begin{aligned}
\frac{1}{2K_T+1}\widehat{P}(\theta_0, \hat{\lambda}_T(\theta_0,s),s) &= -\left(\frac{\hat{\lambda}_T(\theta_0,s)}{2K_T+1}\right)' \hat{g}_T(\theta_0,s) \\
&+ \left(\frac{\hat{\lambda}_T(\theta_0,s)}{2K_T+1}\right)' \left(\sum_{t=1}^T \rho_2\left(\dot{\lambda}_T(s)' g_{tT}(\theta_0,s)\right) g_{tT}(\theta_0,s) g_{tT}(\theta_0,s)'/T\right) \hat{\lambda}_T(\theta_0,s)/2 \\
&= \hat{g}_T(\theta_0,s)'\Omega(s)^{-1}\hat{g}_T((\theta_0,s) - \hat{g}_T(\theta_0,s)'\Omega(s)^{-1}\hat{g}_T(\theta_0,s)/2 + o_p(1) \\
&= \hat{g}_T(\theta_0,s)'\Omega(s)^{-1}\hat{g}_T(\theta_0,s)/2 + o_p(1)
\end{aligned}$$

w.p.a.1 where the second equality holds by $\frac{1}{2K_T+1}\hat{\lambda}_T(\theta_0, s) = -\Omega(s)^{-1}\hat{g}_T(\theta_0, s) + o_p(1)$. The asymptotic distribution of the statistic $\frac{2T}{2K_T+1}\widehat{P}(\tilde{\theta}_T(s), \hat{\lambda}_T(\tilde{\theta}_T(s), s), s)$ is then asymptotically bounded for all $s \in S$ by the asymptotic distribution of $T\hat{g}_T(\theta_0, s)'\Omega(s)^{-1}\hat{g}_T(\theta_0, s)$. By using Lemma 2.1, the result under the null follows. Lemma 2.2 yields the asymptotic distribution under the alternative. The equivalence for the statistic $LM_T^R(\tilde{\theta}_T(s), s)$ is straightforward to show.

Proof of Theorem 3.5

To prove this Theorem, additional assumptions are needed. Let

$$\Sigma(\beta_0) = \lim_{T\to\infty} var\left(\frac{1}{T}\sum_{t=1}^T (g_t(\beta_0)', vec(G_t(\beta_0))')\right)'$$

a $(q + qr) \times (q + qr)$ positive semi-definite symmetric matrix and

$$\Sigma(\beta_0) = \begin{bmatrix} \Omega(\beta_0) & \Omega_{gG}(\beta_0) \\ \Omega_{Gg}(\beta_0) & \Omega_{GG}(\beta_0) \end{bmatrix}$$

where $\Omega_{gG}(\beta_0) = \Omega_{Gg}(\beta_0)'$ is a $(q \times qr)$ matrix and $\Omega_{GG}(\beta_0)$ is a $(qr \times qr)$ matrix.

We define the estimators under the null of no structural change

$$\hat{\Sigma}_{1T}(\beta_0, s) = \frac{2K_T + 1}{[Ts]} \sum_{t=1}^{[Ts]} (g_t(\beta_0)', vec(G_t(\beta_0))')' \, (g_t(\beta_0)', vec(G_t(\beta_0) - EG_{tT}(\beta_0))')$$

$$\hat{\Sigma}_{2T}(\beta_0, s) = \frac{2K_T + 1}{T - [Ts]} \sum_{t=[Ts]+1}^{T} (g_t(\beta_0)', vec(G_t(\beta_0))')' \, (g_t(\theta_0)', vec(G_t(\beta_0) - EG_{tT}(\beta_0))') \,.$$

Assumption 6.8′. Under the true value of the parameters θ_0, $\sup_{s\in S} \|\hat{\Sigma}_{iT}(\beta_0, s) - \Sigma(\beta_0\| \xrightarrow{p} 0$ with S whose closure lies in $(0, 1)$ for $i = 1, 2$.
Assumption 6.3″. Under the true value of the parameters θ_0, $\{g\,(x_{Tt}, \beta_0)\,, vec\,(G\,(x_{Tt}, \beta_0) - EG\,(x_{Tt}, \beta_0)) : t \le T, T \ge 1\}$ is a triangular array of mean zero R^q-valued rv's that is α-mixing with mixing coefficients $\sum_{j=1}^{\infty} j^2 \alpha(j)^{(\nu-1)/\nu} < \infty$ for some $\nu > 1$ with $sup_{t\le T, T\ge 1} E\|g\,(x_{Tt}, \beta_0)\,\|^d < \infty$ and $sup_{t\le T,T\ge 1} E\|G\,(x_{Tt}, \beta_0)\,\|^d < \infty$ for some $d > \max\left(4\nu, \frac{2\eta}{\eta-1}\right)$.

Assumptions 6.3″ and 6.8′ guarantee for the restricted partial-sample GEL that

$$\hat{\Omega}_{Gg,1T}(\beta_0, s) = \frac{2K + 1}{T} \sum_{t=1}^{[Ts]} vec\,(G_{tT}(\beta_0))\, g_{tT}(\beta_0)' \xrightarrow{p} s\Omega_{Gg}(\beta_0), \tag{3}$$

$$\hat{\Omega}_{Gg,2T}(\beta_0, s) = \frac{2K + 1}{T} \sum_{t=[Ts]+1}^{T} vec\,(G_{tT}(\beta_0))\, g_{tT}(\theta_0)' \xrightarrow{p} (1 - s)\Omega_{Gg}(\beta_0), \tag{4}$$

and

$$\hat{\Omega}_{GG,1T}(\beta_0, s) = \frac{2K + 1}{T} \sum_{t=1}^{[Ts]} vec\,(G_{tT}(\beta_0))\, vec\,(G_{tT} - EG_{tT}(\beta_0))' \xrightarrow{p} s\Omega_{GG}(\beta_0),$$

$$\hat{\Omega}_{GG,2T}(\beta_0, s) = \frac{2K + 1}{T} \sum_{t=[Ts]+1}^{T} vec\,(G_{tT}(\beta_0))\, vec\,(G_{tT}(\beta_0) - EG_{tT}(\beta_0))' \xrightarrow{p} (1 - s)\Omega_{GG}(\beta_0).$$

Lemma 2.1 can be shown for the derivatives of the smoothed moment conditions under Assumptions 1.1, 1.2, 6.3″ and 6.8′ as shown for the smoothed moment conditions. Thus, the asymptotic distribution of the derivatives of the centered smoothed moment conditions under the null is given by:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} vec\,(G_{tT}(\beta_0) - EG_{tT}(\beta_0)) \Rightarrow \Omega_{GG}(\beta_0)^{1/2} B_{qr}(s) \tag{5}$$

where $B_{qr}(s)$ is a qr-dimensional vector of standard Brownian motion. Using Lemma 2.1, this yields for the whole vector $(g_{tT}(\beta_0)', (vec(G_{tT}(\beta_0) - EG_{tT}(\beta_0)))')'$

$$T^{-1/2} \sum_{t=1}^{[Ts]} (g_{tT}(\beta_0)', (vec(G_{tT}(\beta_0) - EG_{tT}(\beta_0)))')' \Rightarrow \Sigma(\beta_0)^{1/2} B_{q+qr}(s) \tag{6}$$

where $B_G(s)$ is a $((q + qr) \times 1)$-vector of standard Brownian motion.

We also need the following assumptions:

Assumption 2.1. *Suppose Assumption 1.9 but for $\partial g(x_t, \beta)/\partial \beta_i$ for $i = 1, \ldots, r$.*

Let $\hat{D}_{1T}(\beta_0, s) = \left[\hat{D}_{1,1T}(\beta_0, s), \hat{D}_{2,1T}(\beta_0, s), \ldots, \hat{D}_{r,1T}(\beta_0, s)\right]$ with $\hat{D}_{i,1T}(\beta, s) = \frac{1}{T}\sum_{t=1}^{[Ts]} \rho_1(\hat{\lambda}_{1T}(\beta, s)' g_{tT}(\beta)) G_{i,tT}(\beta, s)$ for $i = 1, \ldots, p$ and respectively for $\hat{D}_{2T}(\beta_0, s)$. By a Taylor expansion of $\hat{D}_{i,1T}(\beta_0, s)$ and $\hat{D}_{i,2T}(\beta_0, s)$ around $\hat{\lambda}_{1T}(\beta_0, s) = 0$ and $\hat{\lambda}_{2T}(\beta_0, s) = 0$ respectively yields

$$\hat{D}_{i,1T}(\beta_0, s) = -\frac{1}{T}\sum_{t=1}^{[Ts]} G_{i,tT}(\beta_0) + \frac{2K+1}{T}\sum_{t=1}^{[Ts]} G_{i,tT}(\beta_0) g_{tT}(\beta_0)' \hat{\Omega}_{1T}(\beta_0, s)^{-1} \frac{1}{[Ts]}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0) + o_p(1)$$

$$\hat{D}_{i,2T}(\beta_0, s) = -\frac{1}{T}\sum_{t=[Ts]+1}^{T} G_{i,tT}(\beta_0) + \frac{2K+1}{T}\sum_{t=[Ts]+1}^{T} G_{i,tT}(\beta_0) g_{tT}(\beta_0)' \hat{\Omega}_{2T}(\beta_0, s)^{-1} \frac{1}{T-[Ts]}\sum_{t=[Ts]+1}^{T} g_{tT}(\beta_0) + o_p(1)$$

using $\frac{1}{2K_T+1}\hat{\lambda}_{1T}(\beta_0, s) = -\hat{\Omega}_{1T}(\beta_0, s)^{-1} \frac{1}{[Ts]}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0) + o_p(1)$ and $\frac{1}{2K_T+1}\hat{\lambda}_{2T}(\beta_0, s) = -\hat{\Omega}_{2T}(\beta_0, s)\frac{1}{T-[Ts]}\sum_{t=[Ts]+1}^{T} g_{tT}(\beta_0) + o_p(1)$ with $\sup_{s\in S}\max_{1\le t\le T} |\rho_2(\hat{\lambda}_{iT}(\beta_0, s)' g_{tT}(\beta_0)) - \rho_2(0)| \xrightarrow{p} 0$ for $i = 1, 2$.

Using (3), (4), (5), (6), Lemma 1.1 and with $G(\beta_0) = \lim_{T\to\infty}\left[T^{-1}\sum_{t=1}^{T} G_{tT}(\beta_0)\right]$, we obtain that

$$\begin{bmatrix} I_q & 0 \\ -\frac{2K+1}{T}\sum_{t=1}^{[Ts]} vec\left(G_{tT}(\beta_0)\right) g_{tT}(\beta_0)'\hat{\Omega}_{1T}(\beta_0)^{-1} & I_{qr} \end{bmatrix} \times \begin{bmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0) \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} vec\left(G_{tT}(\beta_0) - EG_{tT}(\beta_0)\right) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0) \\ -\sqrt{T}\left(\hat{D}_{1T}(\beta_0, s) - sG(\beta_0)\right) \end{bmatrix} \Rightarrow \begin{bmatrix} \Omega(\beta_0)^{1/2} B_q(s) \\ \Omega_D(\beta_0)^{1/2} B_{2.1}(s) \end{bmatrix}$$

with $\Omega_D(\beta_0)^{1/2} B_{2.1}(s) = \Omega_{GG}(\beta_0)^{1/2} B_{qr}(s) - \Omega_{Gg}(\beta_0)\Omega(\beta_0)^{-1}\Omega(\beta_0)^{1/2} B_q(s)$, $\Omega_D(\beta_0) = \Omega_{GG}(\beta_0) - \Omega_{Gg}(\beta_0)\Omega(\beta_0)^{-1}\Omega_{Gg}(\beta_0)$ and $B_{2.1}(s)$ is independent of $B_q(s)$. This result is true for any value of $G(\beta_0)$. Thus, $G(\beta_0)$ can be of full rank value, weak value such that $G_T(\beta_0) = \frac{C_1}{T^{1/2}}$ for $q \times r$ matrix C_1 or $G(\beta_0) = 0$ in the case of no identification.

This implies that

$$\sqrt{T}\left(\hat{D}_{1T}(\beta_0, s) - sG(\beta_0)\right) \Rightarrow -\Omega_D(\beta_0)^{1/2} B_{2.1}(s)$$

and

$$\sqrt{T}\left(\hat{D}_{2T}(\beta_0, s) - (1-s)G(\beta_0)\right) \Rightarrow -\Omega_D(\beta_0)^{1/2}(B_{2.1}(1) - B_{2.1}(s)).$$

Since $\hat{D}_{1T}(\beta_0, s)$ and $\hat{D}_{2T}(\beta_0, s)$ are respectively independent of $\frac{1}{T}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0)$ and $\frac{1}{T}\sum_{t=[Ts]+1}^{T} g_{tT}(\beta_0)$ this yields

$$\left(\hat{D}_{1T}(\beta_0)'\hat{\Omega}_{1T}(\beta_0)^{-1}\hat{D}_{1T}(\beta_0)\right)^{-1/2}\hat{D}_{1T}(\beta_0)'\hat{\Omega}_{1T}(\beta_0)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} g_{tT}(\beta_0) \Rightarrow B_r(s) \tag{7}$$

and

$$\left(\hat{D}_{2T}(\beta_0)'\hat{\Omega}_{2T}(\beta_0)^{-1/2}\hat{D}_{2T}(\beta_0)\right)^{-1}\hat{D}_{2T}(\beta_0)'\hat{\Omega}_{2T}(\beta_0)^{-1}\frac{1}{\sqrt{T}}\sum_{t=[Ts]+1}^{T} g_{tT}(\beta_0) \Rightarrow B_r(1) - B_r(s) \tag{8}$$

where $B_r(s)$ is a r-vector of standard Brownian motion.

Since $\tilde{\theta}_{K,T}(s) = \left(\tilde{\beta}_{K,T}(s)', \tilde{\beta}_{K,T}(s)'\right)'$ minimize the objective function

$$KGEL_T(s) = K_{1T}(\tilde{\beta}_{K,T(s)}, s) + K_{2T}(\tilde{\beta}_{K,T(s)}, s) \leq K_{1T}(\beta_0, s) + K_{2T}(\beta_0, s)$$

for all $s \in S$ and all T. The result follows directly under the null. The derivation under the alternative can be easily obtained. The proof for $KLM_T^R(s)$ is straightforward considering that

$$\sqrt{T}/(2K_T+1)\hat{\lambda}_{1T}(\beta, s) = -\hat{\Omega}_{1T}(\beta)\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} g_{tT}(\beta) + o_p(1),$$

and

$$\sqrt{T}/(2K_T+1)\hat{\lambda}_{2T}(\beta, s) = -\hat{\Omega}_{2T}(\beta)\frac{1}{\sqrt{T}}\sum_{t=[Ts]+1}^{T} g_{tT}(\beta) + o_p(1).$$

References

Andrews, D.W.K. (1991) 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation.' *Econometrica* 59, 817–858

Gallant, A.R., and H. White (1988) *A Unified theory of estimation and inference for nonlinear dynamic models* (Basil Blackwell)

Smith, R. J. (2011) 'GEL criteria for moment condition models.' *Econometric Theory* 27, 1192–1235

Wooldridge, J.M., and H. White (1988) 'Some invariance principles and central limit theorems for dependent heterogeneous processes.' *Econometric Theory* 4, 210–230