Stochastic Expansions and Moment Approximations for Three Indirect Estimators Revised (Extended Appendix)

Stelios Arvanitis and Antonis Demos
Athens University of Economics and Business

June 7, 2012

Abstract

This is an extended appendix for the revision of the paper Stochastic Expansions and Moment Approximations for Three Indirect Estimators.

1 Definition of Estimators

In what follows, when A is a matrix $\|A\|$ denotes a *submultiplicative* matrix norm, such as the Frobenius one (i.e. $\|A\| = \sqrt{trA'A}$). $\mathcal{O}_{\varepsilon}(\theta)$ denotes the open ε-ball around θ in a relevant metric space and $\overline{\mathcal{O}}_{\varepsilon}(\theta)$ its closure. We denote with $\mathcal{PD}(k, \mathbb{R})$ the vector space of positive definite matrices of dimension $k \times k$ endowed with the topology of the Frobenius norm. Consider the following real function from $\mathbb{R}^k \times \mathcal{PD}(k, \mathbb{R})$ for $k \in \mathbb{N}$

$$\|x\|_A \to \left(x^{/}Ax\right)^{1/2}.$$

For a given matrix the previous function defines a norm on $\mathbb{R}^k$. For $s^*, s \in \mathbb{N}^*$ with $s^* \geq s$, let $a^* = \frac{s^*-1}{2}$ and $a = \frac{s-1}{2}$.

Assumption A.1 *For a measurable space* $(\Omega, \mathcal{F})$*, the statistical model (SM) is a family of probability distributions on* $\mathcal{F}$ *parameterized by* $\operatorname{par}(\cdot)$ *a function that is onto a compact subset* $\Theta \subset \mathbb{R}^p$ *for some* $p \in \mathbb{N}$*.*

We abbreviate with $\theta_0 = \operatorname{par}(P_0) \in \operatorname{Int}(\Theta)$, for P_0 in SM. The auxiliary estimator is denoted in the paper by β_n whereas θ_n is the collective notation for the indirect ones. We also employ $b(\theta)$ to denote the binding function

Assumption A.2 *For B a compact subset of $\mathbb{R}^q$, $Q_n : \Omega \times B \to \mathbb{R}$ is jointly measurable. Moreover Q_n is continuous on B for $P_{\theta_0}-$almost every $\omega \in \Omega$.*

We suppress the dependence of the random elements involved on Ω, for notational simplicity.

Definition D.1 *The auxiliary estimator is defined as*

$$\beta_n = \arg\min_{\beta \in B} Q_n(\beta)$$

Q_n could be a likelihood function, a GMM or more generally, a distance type criterion like the ones appearing in the following definitions (see also section 4).

Assumption A.3 *The binding function $b : \Theta \to B$ is injective and continuous on Θ.*

The initial estimators are denoted by θ_n^*.

Assumption A.4 *$W_n^* : \Omega \times \Theta \to \mathbb{R}^q$ and $\theta_n^* : \Omega \to B$ are jointly measurable.*

Definition D.2 *The* GMR1 *estimator is defined as*

$$\theta_n = \arg\min_{\theta \in \Theta} \|\beta_n - b(\theta)\|_{W_n^*(\theta_n^*)}$$

Lemma 1.1 *Under assumptions A.1 and A.2, $\|E_\theta \beta_n\| < \infty$ on Θ.*

Proof. $\|E_\theta \beta_n - b(\theta)\| \leq E_\theta \|\beta_n - b(\theta)\| \leq M_1$, where M_1 denotes the diameter of B, finite due to the compactness of B. ■

Definition D.3 *The* GMR2 *estimator is defined as*

$$\theta_n = \arg\min_{\theta \in \Theta} \|\beta_n - E_\theta \beta_n\|_{W_n^*(\theta_n^*)}$$

Assumption A.5 *Let Q_n be differentiable on B for $P_\theta-$almost every $\omega \in \Omega$. We denote with c_n the derivative of Q_n except for the case where $Q_n = \|c_n(\beta)\|_{W_n(\beta_n^*)}$, where $c_n : \Omega \times B \to \mathbb{R}^l$, $W_n : \Omega \times B \to \mathcal{PD}(l, \mathbb{R})$, and $\beta_n^* : \Omega \to B$ are jointly measurable. Moreover c_n is continuous on B for $P_{\theta_0}-$almost every $\omega \in \Omega$, $c_n(\beta)$ is $P_\theta-$integrable on $\Theta \times B$ and $E_\theta(c_n(\beta))$ is continuous on $\Theta \times B$. Also $W_n^{**} : \Omega \times \Theta \to \mathbb{R}^l$ is jointly measurable.*

$E_\theta(c_n(\beta_n))$ denotes the quantity $E_\theta(c_n(\beta))|_{\beta=\beta_n}$ for notational simplicity.

Definition D.4 *The* GT *estimator is defined as*

$$\theta_n = \arg\min_{\theta\in\Theta} \|E_\theta\left(c_n\left(\beta_n\right)\right)\|_{W_n^{**}\left(\theta_n^*\right)}$$

When $p = q = l$ and $Q_n(\beta) = \|c_n(\beta)\|$, $c_n(\beta) = h_n - E_\beta h_n = h_n - g(\beta)$ with $h_n : \Omega \to \mathbb{R}^p$, integrable on Θ and B, $g(\beta)$ and $m(\theta) = E_\theta h_n$ invertible, it is easy to see that a) the GMR1 estimator is a GMM estimator and b) g is linear GMR1 = GMR2. Notice that a) would be valid even if $\beta_n = r \circ g^{-1} \circ h_n$ for r a bijection. Hence the GMR1 can be a GMM estimator even in cases that the auxiliary is an appropriate transformation of a GMM estimator.

2 Validity of Edgeworth Approximations

Assumptions Specific to the Validity of the Edgeworth Approximations

We denote with D^r, the r-derivative operator and with $D^r\left(f\left(x_0\right)\right)\left(x^r\right)$ the r^{th}-linear function defined by the evaluation of $D^r f$ at x_0 evaluated at $\underbrace{(x,...,x)}_{r\ \ times}$.

Let M denote a universal positive constant, independent of n and θ, not necessarily taking the same value across and inside assumptions proofs and results. $\mathrm{pr}_{i,j}(x)$ denotes the transformation of an r^{th} dimensional vector, say $x = (x_1, x_2, ..., x_r)'$, to a vector containing only the elements of x from the i^{th} to the j^{th} coordinate, i.e. $\mathrm{pr}_{i,j}(x) = (x_i, x_{i+1}, ..., x_j)'$, where naturally $1 \leq i \leq j \leq r$. Finally whenever the assertion "local locally independent of θ" appears in the sequel it implies "independent of θ for $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$" unless otherwise specified. Notice that due to the fact that the spaces Θ and B are separable and closed, suprema of real random elements over these spaces are typically measurable (see van der Vaart and Wellner [7], example 1.7.5 p. 47 due to the theorem of measurable projections, completeness of the underlying probability space, the compactness of Θ and the continuity of b).

Assumption A.6 *β_n is uniformly consistent for $b(\theta)$ with rate $o\left(n^{-a^*}\right)$, i.e.*

$$\sup_{\theta\in\Theta} P_\theta\left(\|\beta_n - b(\theta)\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0.$$

Moreover θ_n^* is uniformly consistent for θ with rate $o\left(n^{-a^*}\right)$.

Assumption A.7 *For $j = *, **$, suppose that there exists a sequence of random elements $x_n : \Omega \to \mathbb{R}^m$, such that $W_n^j(\theta) = \frac{1}{n}\sum W^j\left(x_i(\omega), \theta\right)$ for* *measurable* $W^* : \mathbb{R}^m \times \Theta \to \mathcal{PD}(q, \mathbb{R})$, $W^{**} : \mathbb{R}^m \times \Theta \to \mathcal{PD}(l, \mathbb{R})$ *integrable with respect to* P_{θ^*}, *such that a)*

$$\sup_{\theta^* \in \Theta} P_{\theta^*} \left(\left\| W_n^j(\theta) - E_{\theta^*} W^j(\theta) \right\| > \varepsilon \right) = o\left(n^{-a^*}\right), \forall \varepsilon > 0$$

$E_{\theta^*} W^j(\theta)$ *is continuous when w.r.t.* θ *when* $\theta^* = \theta$, *it is Lipschitz w.r.t.* θ, *for any* θ^* *and the analogous Lipschitz coefficient (say)* $\kappa^j(\theta^*) \sup_{\theta^* \in \Theta} \kappa^j(\theta^*) < +\infty$. *b) Moreover* $W^j(x, \theta)$ *is* s^**-differentiable on* $\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)$ *for* $\varepsilon_0 > \varepsilon$ *and*

$$\sup_{\theta^* \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_{\theta^*} \left(\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} \left\| D^{s^*+1} W_n^j(\theta) \right\| > M \right) = o\left(n^{-a^*}\right)$$

Let $f(x, \theta)$ denote the vector that contains stacked all the distinct components of $W^*(x, \theta)$ and $W^{**}(x, \theta)$ as well as their derivatives up to the order $s^* - 1$. Furthermore $\Psi_{n,s^*}(\theta)$ denotes an Edgeworth measure of order s^* (see for example equations (3.7) and (3.8) of Magdalinos [5]), and with $\pi_{i-1}(z, \theta)$ the polynomial in the density of $\Psi_{n,s^*}(\theta)$ coefficient $\frac{1}{n^{\frac{i-1}{2}}}$, for $i = 1, \ldots, s^*$ (notice that $\pi_0 = 1$).

Assumption A.8 $\sqrt{n} m_n(\theta)$ *has an Edgeworth expansion of order* s^* *uniformly on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ *where*

$$m_n(\theta) = \begin{pmatrix} \beta_n - b(\theta) \\ \theta_n^* - \theta \\ \frac{1}{n} \sum f(x_i, \theta) - E_\theta \frac{1}{n} \sum f(x_i, \theta) \end{pmatrix}$$

or

$$m_n(\theta) = \beta_n - b(\theta)$$

when $W^*(x, \theta)$ *and* $W^{**}(x, \theta)$ *are independent of* x *and* θ. *Furthermore* $\pi_i(z, \theta)$ *is equicontinuous on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ $\forall z \in \mathbb{R}^q$, *for* $i = 1, \ldots, a^*$, *and if* $V(\theta)$ *denotes the variance matrix in the density of* $\Psi_{n,s}(\theta)$ *then it is continuous on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ *and positive definite.*

The proof of the following theorem can be found in Arvanitis and Demos [1] (Proof of Theorem 3.2).

Theorem 2.1 *Suppose that:*
-POLFOC $M_n(\theta)$ *satisfies* $0_{p \times 1} = \sum_{i=0}^{s-1} \frac{1}{n^{i/2}} \sum_{j=0}^{i+1} C_{ij_n}(\theta) \left(M_n(\theta)^j, S_n(\theta)^{i+1-j} \right) + R_n(\theta)$ *with probability* $1 - o\left(n^{-\frac{s-1}{2}}\right)$ *independent of* θ *where* $C_{ij_n} : \overline{\mathcal{O}}_\varepsilon(\theta_0) \times \mathbb{R}^{q^{i+1}} \to \mathbb{R}^p$ *is* $(i+1)$*-linear* $\forall \theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$, $C_{00_n}(\theta), C_{01_n}(\theta)$ *are independent*

of n *and have rank* p $\forall\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, C_{ij_n} *are equicontinuous on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, $\forall x^{i+1}$,
-LUE $S_n(\theta)$ *admits a locally uniform Edgeworth expansion the polynomials of the density of which are equicontinuous on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ *and the variance matrix is continuous on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ *and positive definite,*
-UAT $\sup_{\theta\in\Theta} P\left(\|M_n(\theta)\| > C\ln^{1/2} n\right) = o\left(n^{-\frac{s-1}{2}}\right)$ *for some* $C > 0$ *independent of* θ,
-USR $\sup_{\theta\in\Theta} P\left(\|R_n(\theta)\| > \gamma_n\right) = o\left(n^{-\frac{s-1}{2}}\right)$ *for some real sequence* $\gamma_n = o\left(n^{-\frac{s-1}{2}}\right)$ *independent of* θ.
Then $M_n(\theta)$ *admits a locally uniform Edgeworth expansion he polynomials of the density of which are equicontinuous on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ *and the variance matrix is continuous on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ *and positive definite.*

Existence of Edgeworth Expansions for the GMR-type Estimators

The GMR1 Case

Assumption A.9 $b(\theta)$ *is* s^*+1 *continuously differentiable and* $\operatorname{rank} Db(\theta) = p$, *for all* θ *in* $\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)$ *and* $\varepsilon_0 > \varepsilon$.

Lemma 2.2 *i) Under the assumptions A.1, A.2, A.3, A.4, A.6 and A.7.a) the* GMR1 *is uniformly consistent for* θ *with rate* $o(n^{-a})$. *ii) If additionally assumptions A.7b), A.8 and A.9 hold then,* $\sqrt{n}\,(\text{GMR1} - \theta)$ *has an Edgeworth expansion of order* s^* *uniformly on* $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, *for* $\varepsilon < \varepsilon_0$., *where* ε_0 *as in the above assumption.*

Proof: i) Due to the triangle inequality and assumption A.6 we have that for $\varepsilon > 0$

$$\begin{aligned} &\sup_{\theta^*\in\Theta} P_{\theta^*}\left(\sup_{\theta\in\Theta} \left| \|\beta_n - b(\theta)\| - \|b(\theta^*) - b(\theta)\| \right| > \varepsilon\right) \\ \leq\ &\sup_{\theta^*\in\Theta} P_{\theta^*}\left(\|\beta_n - b(\theta^*)\| > \varepsilon\right) = o\left(n^{-a^*}\right) \end{aligned}$$

Hence for $q_n(\theta) = \beta_n - b(\theta)$, $q(\theta^*, \theta) = b(\theta^*) - b(\theta)$ and by assumption A.7.a) lemma AL.3 applies. Hence for $\gamma(\theta) = \theta$ due to assumption A.3 lemma AL.1 also applies implying the result.
ii) Given i), we have that $\theta_n \in \mathcal{O}_{\epsilon}(\theta)$ with P_{θ}-probability $1 - o\left(n^{-a^*}\right)$ that is locally independent of θ for any $\epsilon > 0$. For some ϵ small enough, such that $\mathcal{O}_{\epsilon}(\theta) \subset \mathcal{O}_{\varepsilon_0}(\theta_0)$ (which exists due to the fact that $\varepsilon_0 > \varepsilon$) due to assumption A.9, we have that condition FOC of the appendix lemmas AL.4 and AL.5 is satisfied by the GMR1 estimator with $Q_n \doteqdot \frac{\partial b'}{\partial\theta}$. Furthermore assumption

A.9 implies conditions HUB ($\gamma(\theta) = \theta$ hence set $\delta = \varepsilon_0$) and RANK of the same lemma. Condition TIGHT follows from A.8, as under this assumption there is $C^* > 0$ locally independent of θ such that

$$\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\sqrt{n} \left\| \beta_n - b(\theta) \right\| > C^* \ln^{1/2} n \right) = o\left(n^{-a^*}\right) \tag{1}$$

(see lemma AL.2 of Arvanitis and Demos [1]). Hence lemma AL.4 applies ensuring that

$$\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\sqrt{n} \left\| \text{GMR1} - \theta \right\| > C \ln^{1/2} n \right) = o\left(n^{-a^*}\right)$$

for some $C > 0$ locally independent of θ. Hence condition UTIGHT of lemma AL.5 holds. Moreover assumption A.8 implies condition UEDGE of the same lemma for $M_n(\theta) = \sqrt{n} m_n(\theta)$. Due to assumption A.9 for any $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ and any θ_* sufficiently close to θ, $\frac{\partial b'}{\partial \theta}(\theta_*)$ admits a Taylor expansion of order $s^* - 1$ around θ of the form

$$\begin{aligned} \frac{\partial b'}{\partial \theta}(\theta_*) &= \sum_{i=0}^{s^*-1} \frac{1}{i!} D^i \frac{\partial b'}{\partial \theta}(\theta) \left((\theta_* - \theta)^i \right) \\ &+ \frac{1}{(s^*-1)!} \left(D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta^+) - D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta) \right) \left((\theta_* - \theta)^{s^*-1} \right) \end{aligned}$$

where θ^+ lies between θ_* and θ. This implies that for any $\theta_n = \text{GMR1}$ due to condition UTIGHT we have that with P_θ-probability $1 - o\left(n^{-a^*}\right)$ locally independent of θ

$$\frac{\partial b'}{\partial \theta}(\theta_n) = \sum_{i=0}^{s^*-1} \frac{1}{i!} \frac{1}{n^{i/2}} D^i \frac{\partial b'}{\partial \theta}(\theta) \left(\left(\sqrt{n}(\theta_n - \theta) \right)^i \right) + R_n^*(\theta_n, \theta)$$

where $R_n^*(\theta_n, \theta) = \frac{1}{(s^*-1)!} \frac{1}{n^{(s^*-1)/2}} \left(D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta_n^+) - D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta) \right) \left(\left(\sqrt{n}(\theta_n - \theta) \right)^{s^*-1} \right)$, and θ_n^+ lies between θ_n and θ. Now by assumption A.9 $\frac{\partial b'}{\partial \theta}(\theta)$ has full rank for any $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ and by submultiplicativity, the relation of θ_n^+ to θ_n and condition UTIGHT

$$\begin{aligned} &\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\left\| \begin{array}{c} \frac{1}{(s^*-1)!} \frac{1}{n^{(s^*-1)/2}} \left(D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta_n^+) - D^{s^*-1} \frac{\partial b'}{\partial \theta}(\theta) \right) \times \\ \left(\left(\sqrt{n}(\theta_n - \theta) \right)^{s^*-1} \right) \end{array} \right\| > \gamma_n^* \right) \\ \leq &\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\frac{1}{(s^*-1)!} \frac{1}{n^{(s^*-1)/2}} \sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)} \left\| D^{s^*} \frac{\partial b'}{\partial \theta}(\theta) \right\| \left\| \theta_n^+ - \theta \right\| \left\| \sqrt{n}(\theta_n - \theta) \right\|^{s^*-1} > \gamma_n^* \right) \\ \leq &\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\frac{M}{(s^*-1)!} \frac{C^{s^*}}{n^{s^*/2}} \ln^{s^*/2} n > \gamma_n^* \right) + o\left(n^{-a^*}\right) \end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^* = \frac{M}{(s^*-1)!}\frac{C^{s^*}}{n^{s^*/2}}\ln^{s^*/2} n = o\left(n^{-a^*}\right)$ and locally independent of θ. Analogously, due to assumption A.9 for any $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ and any θ_* sufficiently close to θ, $b(\theta_*)$ admits a Taylor expansion of order s^*-1 around θ of the form

$$\begin{aligned} q_n &= \beta_n - b(\theta_*) = \beta_n - b(\theta_*) - \sum_{i=1}^{s^*}\frac{1}{i!}D^i b(\theta)\left((\theta_*-\theta)^i\right) \\ &\quad -\frac{1}{s^*!}\left(D^{s^*}b\left(\theta^+\right) - D^{s^*}b(\theta)\right)\left((\theta_*-\theta)^{s^*}\right) \end{aligned}$$

where θ^+ lies between θ_* and θ. This implies that for θ_n we have that with P_θ-probability $1-o\left(n^{-a^*}\right)$

$$\begin{aligned} &\sqrt{n}\left(\beta_n - b(\theta_n)\right) \\ = &\sqrt{n}\left(\beta_n - b(\theta)\right) \\ &+\sum_{i=0}^{s^*-1}\frac{1}{(i+1)!}\frac{1}{n^{i/2}}D^{i+1}b(\theta)\left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{i+1}\right) + R_n^{\#}(\theta_n,\theta) \end{aligned}$$

where $R_n^{\#}(\theta_n,\theta) = \frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\left(D^{s^*}b\left(\theta_n^+\right) - D^{s^*}b(\theta)\right)\left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{s^*}\right)$, and θ_n^+ lies between θ_n and θ. Now by assumption A.9 $\frac{\partial b'}{\partial\theta}(\theta)$ has full rank for any $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ and so does the identity matrix in front of $\sqrt{n}\left(\beta_n - b(\theta)\right)$, and thereby due to submultiplicativity, the relation of θ_n^+ to θ_n and condition UTIGHT

$$\begin{aligned} &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|\begin{array}{c}\frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\left(D^{s^*}b\left(\theta_n^+\right) - D^{s^*}b(\theta)\right)\times \\ \left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{s^*}\right)\end{array}\right\| > \gamma_n^{\#}\right) \\ \leq &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\begin{array}{c}\frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^{s^*+1}b(\theta)\right\|\left\|\theta_n^+-\theta\right\| \\ \times\left\|\sqrt{n}\left(\theta_n-\theta\right)\right\|^{s^*} > \gamma_n^{\#}\end{array}\right) \\ \leq &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{s^*!}\frac{C^{s^*+1}}{n^{s^*/2}}\ln^{(s^*+1)/2} n > \gamma_n^{\#}\right) + o\left(n^{-a^*}\right) \end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^{\#} = \frac{M}{s^*!}\frac{C^{s^*+1}}{n^{s^*/2}}\ln^{(s^*+1)/2} n = o\left(n^{-a^*}\right)$ and locally independent of θ. Finally due to lemma AL.6 which applies by assumptions A.8 and A.7 condition EXPAND holds and the result follows by the same lemma.■

The GMR2 Case

Assumption A.10 $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^{s^*}E_\theta\beta_n\right\| < M$.

Lemma 2.3 *i) Under the assumptions A.1, A.2, A.3, A.4, A.6 and A.7.a) the* GMR2 *is uniformly consistent for* θ *with rate* $o\left(n^{-a^*}\right)$. *ii) If additionally assumptions A.7.b), A.8, A.9 and A.10 hold then* $\sqrt{n}\,(\mathrm{GMR2} - \theta)$ *has an Edgeworth expansion of order* $s^* - 1$ *uniformly on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$.

Proof: For $\varepsilon > 0$, let $E(\varepsilon, \theta) = \left\{\omega \in \Omega : \|\beta_n - b(\theta)\| > \frac{\varepsilon}{2}\right\} \in \mathcal{F}$, then

$$\sup_{\theta\in\Theta} \|E_\theta \beta_n - b(\theta)\| \le \sup_{\theta\in\Theta} E_\theta \|\beta_n - b(\theta)\| \, 1_{E(\varepsilon,\theta)} + \frac{\varepsilon}{2}.$$

As B is bounded, due to assumption A.2 and by assumption A.6 there exists an n^* such that $\sup_{\theta\in\Theta} P_\theta\left(\|\beta_n - b(\theta)\| > \frac{\varepsilon}{3}\right) \le \frac{\varepsilon}{2M}$ where M denotes the diameter of B. Hence

$$\sup_{\theta\in\Theta} \|E_\theta \beta_n - b(\theta)\| \le \varepsilon \text{ for all } n \ge n^*$$

and since ε is arbitrary

$$\sup_{\theta\in\Theta} \|E_\theta \beta_n - b(\theta)\| = o(1) \tag{2}$$

Due to the triangle inequality and assumption A.6 we have that for $\varepsilon > 0$

$$\begin{aligned} &\sup_{\theta^*\in\Theta} P_{\theta^*}\left(\sup_{\theta\in\Theta} \left|\|\beta_n - E_\theta\beta_n\| - \|b(\theta^*) - b(\theta)\|\right| > \varepsilon\right) \\ \le\ &\sup_{\theta^*\in\Theta} P_{\theta^*}\left(\|\beta_n - b(\theta^*)\| + \sup_{\theta\in\Theta}\|E_\theta\beta_n - b(\theta)\| > \varepsilon\right) = o\left(n^{-a^*}\right) \end{aligned}$$

For $q_n(\theta) = \beta_n - E_\theta\beta_n$, $q(\theta^*, \theta) = b(\theta^*) - b(\theta)$ and by assumption A.7.a) lemma AL.3 applies. Hence for $\gamma(\theta) = \theta$ due to assumption A.3 lemma AL.1 also applies implying the result.

ii) Given i), we have that $\theta_n \in \mathcal{O}_\epsilon(\theta)$ with P_θ-probability $1 - o\left(n^{-a^*}\right)$ that is locally independent of θ for any $\epsilon > 0$. For some ϵ small enough, such that $\mathcal{O}_\epsilon(\theta) \subset \mathcal{O}_{\varepsilon_0}(\theta_0)$ (which exists due to the fact that $\varepsilon_0 > \varepsilon$) due to assumption A.10, we have that condition FOC of the appendix lemmas AL.4 and AL.5 is satisfied by the GMR1 estimator with $Q_n \doteqdot \frac{\partial E_\theta \beta_n'}{\partial\theta}$. Furthermore assumption A.10 and A.9 imply conditions HUB ($\gamma(\theta) = \theta$ hence set $\delta = \varepsilon_0$) and RANK of the same lemma due to the fact that since $D^2 E_\theta\beta_n$ is uniformly bounded on $\mathcal{O}_{\varepsilon_0}(\theta_0)$, $DE_\theta\beta_n$ converges uniformly to $Db(\theta)$ due to lemma AL.7 and therefore the rank condition is implied by A.10 for large enough n. Now as $a^* > a \ge 0$ we have that $a^* > 0$ and there exists a $C_2 > 0$ locally independent of θ such that for $E^* = \left\{ \omega \in \Omega : \|\beta_n - b(\theta)\| > C_2 \frac{\ln^{1/2} n}{n^{1/2}} \right\} \in \mathcal{F}$

$$
\begin{aligned}
& \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} \|E_\theta \beta_n - b(\theta)\| \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} E_\theta \left[\|\beta_n - b(\theta)\| 1_{E^*} \right] + \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} E_\theta \left[\|\beta_n - b(\theta)\| 1_{E^*} \right] \\
\leq & \; M \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - b(\theta)\| > C_2 \frac{\ln^{1/2} n}{n^{1/2}} \right) + C_2 \frac{\ln^{1/2} n}{n^{1/2}} \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} E_\theta 1_{E^*} \\
= & \; M \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - b(\theta)\| > C_2 \frac{\ln^{1/2} n}{n^{1/2}} \right) \\
& + C_2 \frac{\ln^{1/2} n}{n^{1/2}} \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - b(\theta)\| \leq C_2 \frac{\ln^{1/2} n}{n^{1/2}} \right) \\
= & \; o\left(n^{-a^*}\right) + C_2 \frac{\ln^{1/2} n}{n^{1/2}} \left(1 - o\left(n^{-a^*}\right)\right) \\
= & \; o\left(n^{-a^*}\right) + C_2 \frac{\ln^{1/2} n}{n^{1/2}} = O\left(\frac{\ln^{1/2} n}{n^{1/2}} \right),
\end{aligned}
$$

where the penultimate line comes from equation 1, above. Hence

$$
\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} \|E_\theta \beta_n - b(\theta)\| = O\left(\frac{\ln^{1/2} n}{n^{1/2}} \right)
$$

and therefore

$$
\begin{aligned}
& \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - E_\theta \beta_n\| > C_1 \frac{\ln^{1/2} n}{n^{1/2}} \right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - b(\theta)\| + \|E_\theta \beta_n - b(\theta)\| > C_1 \frac{\ln^{1/2} n}{n^{1/2}} \right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - b(\theta)\| + O\left(\frac{\ln^{1/2} n}{n^{1/2}} \right) > C_1 \frac{\ln^{1/2} n}{n^{1/2}} \right) = o\left(n^{-a^*}\right)
\end{aligned}
$$

Hence due to A.8 and lemma AL.2 of Arvanitis and Demos [1] there exist $C_1 > 0$ large enough and locally independent of θ for which

$$
\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta \left(\|\beta_n - E_\theta \beta_n\| > C_1 \frac{\ln^{1/2} n}{n^{1/2}} \right) = o\left(n^{-a^*}\right).
$$

Hence lemma AL.4 applies ensuring that

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\sqrt{n}\,\|\mathrm{GMR2}-\theta\| > C\ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

for some $C > 0$ locally independent of θ, hence condition UTIGHT of lemma AL.5 holds. Moreover assumption A.8 along with the fact that the support of $\beta_n - b(\theta)$ is uniformly bounded by $\overline{\mathcal{O}}_{3\eta}(0)$ for any η greater or equal than the diameter of B, and the fact that $\sqrt{n}\,(\beta_n - E_\theta\beta_n)$ admits a locally uniform Edgeworth expansion of order s^*-1 (see lemma 4.1 of Arvanitis and Demos [1]) imply condition UEDGE of lemma AL.5 for $M_n(\theta) = \sqrt{n}\begin{pmatrix} m_n(\theta) \\ \beta_n - E_\theta\beta_n \end{pmatrix}$ with order s^*-1. Due to assumption A.10 for any $\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ and any θ_* sufficiently close to θ, $\frac{\partial E_\theta\beta'_n}{\partial\theta}(\theta_*)$ admits a Taylor expansion of order s^*-1 around θ of the form

$$\begin{aligned}\frac{\partial E_{\theta_*}\beta'_n}{\partial\theta} &= \sum_{i=0}^{s^*-2}\frac{1}{i!}D^i\frac{\partial E_\theta\beta''_n}{\partial\theta}\left((\theta_*-\theta)^i\right) \\ &+\frac{1}{(s^*-2)!}\left(D^{s^*-2}\frac{\partial E_{\theta^+}\beta''_n}{\partial\theta} - D^{s^*-1}\frac{\partial E_\theta\beta''_n}{\partial\theta}\right)\left((\theta_*-\theta)^{s^*-1}\right)\end{aligned}$$

where θ^+ lies between θ_* and θ. This implies that for $\theta_n = \mathrm{GMR2}$ due to condition UTIGHT we have that with P_θ-probability $1-o\left(n^{-a^*}\right)$ locally independent of θ

$$\frac{\partial E_{\theta_n}\beta'_n}{\partial\theta} = \sum_{i=0}^{s^*-2}\frac{1}{i!}\frac{1}{n^{i/2}}D^i\frac{\partial E_{\theta_n}\beta'_n}{\partial\theta}(\theta)\left(\left(\sqrt{n}\,(\theta_n-\theta)\right)^i\right) + R^*_n(\theta_n,\theta)$$

where $R^*_n(\theta_n,\theta) = \frac{1}{(s^*-2)!}\frac{1}{n^{(s^*-2)/2}}\left(D^{s^*-2}\frac{\partial E_{\theta_n^+}\beta'_n}{\partial\theta} - D^{s^*-1}\frac{\partial E_\theta\beta'_n}{\partial\theta}\right)\left(\left(\sqrt{n}\,(\theta_n-\theta)\right)^{s^*-2}\right)$, and θ_n^+ lies between θ_n and θ. Now by assumption A.10, by submultiplicativity, the relation of θ_n^+ to θ_n and condition UTIGHT

$$\begin{aligned}&\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|\begin{matrix}\frac{1}{(s^*-2)!}\frac{1}{n^{(s^*-2)/2}}\left(D^{s^*-2}\frac{\partial E_{\theta_n^+}\beta'_n}{\partial\theta} - D^{s^*-1}\frac{\partial E_\theta\beta'_n}{\partial\theta}\right)\times \\ \left(\left(\sqrt{n}\,(\theta_n-\theta)\right)^{s^*-2}\right)\end{matrix}\right\| > \gamma^*_n\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\frac{1}{(s^*-2)!}\frac{1}{n^{(s^*-2)/2}}\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^{s^*-1}\frac{\partial E_\theta\beta'_n}{\partial\theta}\right\|\left\|\theta_n^+-\theta\right\|\left\|\sqrt{n}\,(\theta_n-\theta)\right\|^{s^*-2} > \gamma^*_n\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\frac{M}{(s^*-2)!}\frac{C^{s^*}}{n^{(s^*-1)/2}}\ln^{(s^*-1)/2} n > \gamma^*_n\right) + o\left(n^{-a^*}\right)\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^* = \frac{M}{(s^*-2)!}\frac{C^{s^*}}{n^{(s^*-1)/2}}\ln^{(s^*-1)/2} n = o\left(n^{-a^*}\right)$ and locally independent of θ. Analogously, due to assumption A.9 for any $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ and any θ_* sufficiently close to θ, $E_{\theta_*}\beta_n$ admits a Taylor expansion of order s^*-1 around θ of the form

$$\begin{aligned} q_n &= \beta_n - E_{\theta_*}\beta_n = \beta_n - E_\theta\beta_n - \sum_{i=1}^{s^*-1}\frac{1}{i!}D^i E_\theta\beta_n\left((\theta_*-\theta)^i\right) \\ &\quad -\frac{1}{(s^*-1)!}\left(D^{s^*-1}E_{\theta^+}\beta_n - D^{s^*-1}E_\theta\beta_n\right)\left((\theta_*-\theta)^{s^*}\right) \end{aligned}$$

where θ^+ lies between θ_* and θ. This implies that for θ_n we have that with P_θ-probability $1-o\left(n^{-a^*}\right)$

$$\begin{aligned} &\sqrt{n}\left(\beta_n - E_{\theta_n}\beta_n\right) \\ =\ &\sqrt{n}\left(\beta_n - E_\theta\beta_n\right) \\ &+\sum_{i=0}^{s^*-2}\frac{1}{(i+1)!}\frac{1}{n^{i/2}}D^{i+1}E_\theta\beta_n\left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{i+1}\right) + R_n^{\#}\left(\theta_n,\theta\right) \end{aligned}$$

where $R_n^{\#}\left(\theta_n,\theta\right) = \frac{1}{(s^*-1)!}\frac{1}{n^{(s^*-2)/2}}\left(D^{s^*-1}E_{\theta^+}\beta_n - D^{s^*-1}E_\theta\beta_n\right)\left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{s^*-1}\right)$, and θ_n^+ lies between θ_n and θ. Now by the previous for large enough n $\frac{\partial E_\theta\beta_n'}{\partial\theta}(\theta)$ has full rank for any $\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)$ and so does the identity matrix in front of $\sqrt{n}\left(\beta_n - E_\theta\beta_n\right)$, and thereby due to submultiplicativity, the relation of θ_n^+ to θ_n and condition UTIGHT

$$\begin{aligned} &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|\begin{array}{c}\frac{1}{(s^*-1)!}\frac{1}{n^{(s^*-2)/2}}\left(D^{s^*-1}E_{\theta^+}\beta_n - D^{s^*-1}E_\theta\beta_n\right)\times \\ \left(\left(\sqrt{n}\left(\theta_n-\theta\right)\right)^{s^*-1}\right)\end{array}\right\| > \gamma_n^{\#}\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\begin{array}{c}\frac{1}{(s^*-1)!}\frac{1}{n^{(s^*-2)/2}}\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^{s^*-1}E_\theta\beta_n\right\|\left\|\theta_n^+-\theta\right\| \\ \times\left\|\sqrt{n}\left(\theta_n-\theta\right)\right\|^{s^*-1} > \gamma_n^{\#}\end{array}\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{(s^*-1)!}\frac{C^{s^*}}{n^{(s^*-1)/2}}\ln^{s^*/2} n > \gamma_n^{\#}\right) + o\left(n^{-a^*}\right) \end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^{\#} = \frac{M}{(s^*-1)!}\frac{C^{s^*}}{n^{(s^*-1)/2}}\ln^{s^*/2} n = o\left(n^{-a^*}\right)$ and locally independent of θ. Finally due to lemma AL.6 which applies by assumptions A.8 and A.7 condition EXPAND holds and the result follows by the same lemma.■

We denote with $k_{i_\beta}(z,\theta) = z\pi_{i-1}(z,\theta)$ and with $\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right) = \int_{\mathbb{R}^q} k_{i_\beta}(z,\theta)\,\varphi_{V(\theta)}(z)\,dz$ where $\pi_{i-1}(z,\theta)$ and $V(\theta)$ as in assumption A.8.

Assumption A.11 *$\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)$ is s^* continuously differentiable for $i = 1,\ldots,s^*-1$ over $\mathcal{O}_\varepsilon(\theta_0)$.*

Lemma 2.4 *If assumptions A.8, A.9 and A.11 hold for $s^* > s$ then for any sequence θ_n^+ for which*

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\theta_n^+ - \theta\right\| > M\ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

we have that for any $\varepsilon_* < \varepsilon$

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)} P_\theta\left(\left\|\sqrt{n}\left(E_{\theta_n^+}\beta_n - E_\theta\beta_n\right) - A_n(\theta)\right\| > \gamma_n\right) = o\left(n^{-a^*}\right)$$

where

$$A_n(\theta) = \sum\nolimits_{i=1}^s \frac{1}{n^{\frac{i-1}{2}} i!} D^i\left(b(\theta) + \sum\nolimits_{j=1}^{s-i}\frac{\mathcal{I}_V\left(k_{j_\beta}(z,\theta)\right)}{n^{\frac{j}{2}}}\right)\left(\sqrt{n}\left(\theta_n^+ - \theta\right)^i\right)$$

$\gamma_n = o\left(n^{-a}\right)$ independent of θ, using the convention that when $s-i=0$, then $\sum_{j=1}^{s-i}$ is empty.

Proof. By assumption A.8, lemma 3.1, below, adding subtracting $\sqrt{n}\left(b(\theta) + \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)}{n^{\frac{i}{2}}}\right)$ and $\sqrt{n}\left(b\left(\theta_n^+\right) + \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{i_\beta}\left(z,\theta_n^+\right)\right)}{n^{\frac{i}{2}}}\right)$, we get
$\sqrt{n}\left(E_{\theta_n^*}\beta_n - E_\theta\beta_n\right) - A_n(\theta) =$
$\sqrt{n}\left(E_{\theta_n^*}\beta_n - b\left(\theta_n^+\right) - \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{j_\beta}\left(z,\theta_n^+\right)\right)}{n^{\frac{i}{2}}}\right) - \sqrt{n}\left(E_\theta\beta_n - b(\theta) - \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)}{n^{\frac{i}{2}}}\right) +$
$\sqrt{n}\left(b\left(\theta_n^+\right) - b(\theta) - \sum_{i=1}^s \frac{1}{i!}D^i b(\theta)\left(\left(\theta_n^+ - \theta\right)^i\right)\right) + B_n$ where
$B_n = \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{i_\beta}\left(z,\theta_n^+\right)\right)}{n^{\frac{i-1}{2}}} - \sum_{i=1}^s \frac{\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)}{n^{\frac{i-1}{2}}} - \sum_{i=1}^s \frac{1}{i!}\sum_{j=1}^{s-i} D^i \frac{\mathcal{I}_V\left(k_{j_\beta}(z,\theta)\right)}{n^{\frac{j-1}{2}}}\left(\left(\theta_n^+ - \theta\right)^i\right).$
Employing the mean value theorem for $\mathcal{I}_V\left(k_{i_\beta}\left(z,\theta_n^+\right)\right)$, and for θ_n^{++} such that $\left\|\theta_n^{++} - \theta\right\| < \left\|\theta_n^+ - \theta\right\|$, we get $B_n =$
$= \sum_{i=1}^s \left(\frac{1}{n^{\frac{i-1}{2}}}\sum_{m=1}^{s-i}\frac{1}{m!}D^m\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\left(\left(\theta_n^+ - \theta\right)^m\right) - \frac{1}{i!}\sum_{j=1}^{s-i} D^i\frac{\mathcal{I}_V\left(k_{j_\beta}(z,\theta)\right)}{n^{\frac{j-1}{2}}}\left(\left(\theta_n^+ - \theta\right)^i\right)\right) +$
$\sum_{i=1}^s \frac{1}{n^{\frac{i-1}{2}}}\frac{1}{(s-i+1)!}D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\left(\left(\theta_n^{++} - \theta\right)^{s-i+1}\right)$. Collecting terms we get:

$$B_n = \sum\nolimits_{i=1}^s \frac{1}{n^{\frac{i-1}{2}}}\frac{1}{(s-i+1)!}D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\left(\left(\theta_n^{++} - \theta\right)^{s-i+1}\right).$$

Taking into account that $\theta_n^+ \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$ with probability $1 - o\left(n^{-a^*}\right)$ and employing the triangular inequality we have that, for $s < s^*$,

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|\sqrt{n}\left(E_{\theta_n^+}\beta_n - E_\theta\beta_n\right) - A_n(\theta)\right\| > \gamma_n\right) \\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} \sqrt{n}\left\|E_\theta\beta_n - b(\theta) - \textstyle\sum_{i=1}^s \dfrac{\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)}{n^{\frac{i}{2}}}\right\| > \dfrac{\gamma_n}{6}\right) \\
&+\textstyle\sum_{i=1}^s \displaystyle\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{1}{n^{\frac{i-1}{2}}}\left\|B_n\right\| > \frac{\gamma_n}{3s}\right) \\
&+\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|b\left(\theta_n^+\right) - b(\theta) - \textstyle\sum_{i=1}^s \dfrac{1}{i!} D^i b(\theta)\left(\left(\theta_n^+ - \theta\right)^i\right)\right\| > \dfrac{\gamma_n}{3}\right) + o\left(n^{-a^*}\right).
\end{aligned}$$

Now we have that

$$a_n = \sqrt{n}\left\|E_\theta\beta_n - b(\theta) - \textstyle\sum_{i=1}^s \dfrac{\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)}{n^{\frac{i}{2}}}\right\| = o\left(n^{-a}\right)$$

independent of θ, due to lemma 3.1.
Now, due to the continuity of $D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)$, assumption A.11, and the assumption of the asymptotic behavior of θ_n^+ we get

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{1}{n^{\frac{i-1}{2}}}\left\|B_n\right\| > \frac{\gamma_n}{3s}\right) \\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{1}{n^{\frac{i-1}{2}}}\frac{1}{(s-i+1)!}\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\right\|\left\|\theta_n^+ - \theta\right\|^{s-i+1} > \frac{\gamma_n}{3s}\right) \\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{\ln^{\frac{s-i+1}{2}} n}{n^{\frac{s}{2}}}\frac{1}{(s-i+1)!}\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\right\| > \frac{\gamma_n}{3s}\right) + o\left(n^{-a^*}\right) \\
=\ &o\left(n^{-a^*}\right)
\end{aligned}$$

provided that $\gamma_n \geq \frac{\ln^{\frac{s-i+1}{2}} n}{n^{\frac{s}{2}}}\frac{3s\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\right\|}{(s-i+1)!}$.
Furthermore using the same reasoning as above

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|b\left(\theta_n^+\right) - b(\theta) - \textstyle\sum_{i=1}^s \dfrac{1}{i!} D^i b(\theta)\left(\left(\theta_n^* - \theta\right)^i\right)\right\| > \dfrac{\gamma_n}{3}\right) \\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\theta_n^+ - \theta\right\|^{s+1} > \frac{(s+1)!\gamma_n}{3\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s+1}b(\theta)\right\|}\right) + o\left(n^{-a^*}\right) \\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{\ln^{\frac{s+1}{2}} n}{n^{\frac{s}{2}}} > \frac{(s+1)!\gamma_n}{3\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s+1}b(\theta)\right\|}\right) + o\left(n^{-a^*}\right) = o\left(n^{-a^*}\right)
\end{aligned}$$

when $\gamma_n \geq \frac{3\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s+1}b(\theta)\right\|}{(s+1)!}\frac{\ln^{\frac{s+1}{2}}n}{n^{\frac{s}{2}}}$. Hence for

$\gamma_n = \max\left(\frac{3\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s+1}b(\theta)\right\|}{(s+1)!}\frac{\ln^{\frac{s+1}{2}}n}{n^{\frac{s}{2}}}, 6a_n, \frac{\ln^{\frac{s-i+1}{2}}n}{n^{\frac{s}{2}}}\frac{3s\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^{s-i+1}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right)\right\|}{(s-i+1)!}, i=1,\ldots,s\right)$
the result follows for large enough n. ■

Lemma 2.5 *Suppose that $p = q$ and assumptions A.1, A.2, A.3, A.4, A.6, A.7, A.8, A.9 and A.11 hold for $s^* > s$. i) If $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^2E_\theta\beta_n\right\| < M$ then $\sqrt{n}\,(\mathrm{GMR2}-\theta)$ has an Edgeworth expansion of order s uniformly on $\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)$ for any $\varepsilon_* < \varepsilon$. ii) if $\beta_n = b\,(\mathrm{GMR1})$ with probability $1 - o\,(n^{-a})$ uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ and $\beta_n = E_{\mathrm{GMR2}}\beta_n$ with probability $1-o\,(n^{-a})$ uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ then $\sqrt{n}\,(\mathrm{GMR2}-\theta)$ has an Edgeworth expansion of order s uniformly on $\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)$ for any $\varepsilon_* < \varepsilon$.*

Proof. i) Notice that the uniform consistency follow for the GMR1 and GMR2 as in the first parts of lemmas 2.2, 2.3. Assumption A.9 along with i) imply that for $r = 1, 2$, $\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|D^r\left(E_\theta\beta_n - b\,(\theta)\right)\right\| < M$, which in turn means that $D^{r-1}\left(E_\theta\beta_n - b\,(\theta)\right)$ are uniformly Lipschitz on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$, and therefore uniformly equicontinuous on the same ball. This implies the commutativity of the limit, with respect to n and the derivative operator, uniformly over $\overline{\mathcal{O}}_\varepsilon(\theta_0)$. This along with the second part of assumption A.9, i.e. $\operatorname{rank} Db\,(\theta) = p$ for all θ in $\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)$, and continuity imply that $\operatorname{rank} DE_\theta\beta_n = p$, for all θ in $\mathcal{O}_{\varepsilon_0}(\theta_0)$ for n large enough. As now $p = q$, by the definition of GMR2 we get that $\beta_n = E_{\mathrm{GMR2}}\beta_n$ with probability $1 - o\left(n^{-a^*}\right)$ uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$. This implies condition FOC of lemma AL.5. Furthermore by the second part of lemma 2.2 we have that

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\mathrm{GMR1}-\theta\right\| > M\ln^{1/2}n\right) = o\left(n^{-a^*}\right) \tag{3}$$

Hence with probability $1 - o\left(n^{-a^*}\right)$ locally independent of θ, applying the mean value theorem we have that

$$b\,(\mathrm{GMR1}) = b\,(\mathrm{GMR2}) + \frac{\partial b^{/}\left(\theta_n^+\right)}{\partial\theta}\left(\mathrm{GMR1}-\mathrm{GMR2}\right),$$

where θ_n^+ is such that $\left\|\theta_n^+ - \mathrm{GMR2}\right\| < \left\|\mathrm{GMR1}-\mathrm{GMR2}\right\|$. It follows that with P_θ-probability $1 - o\left(n^{-a^*}\right)$ locally independent of θ

$$\mathrm{GMR1}-\mathrm{GMR2} = \left(\frac{\partial b^{/}\left(\theta_n^+\right)}{\partial\theta}\right)^{-1}\left(b\,(\mathrm{GMR1}) - b\,(\mathrm{GMR2})\right).$$

As now $p = q$, by the definition of GMR1 we get that $b(\text{GMR1}) = \beta_n$ with probability $1 - o\left(n^{-a^*}\right)$ uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$. Hence with probability $1 - o\left(n^{-a^*}\right)$ uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$

$$\begin{aligned}
\|\text{GMR1} - \text{GMR2}\| &\leq M\|\beta_n - b(\text{GMR2})\| \\
&\leq M\left(\|\beta_n - E_{\text{GMR2}}\beta_n\| + \|E_{\text{GMR2}}\beta_n - b(\text{GMR2})\|\right) \\
&\leq M\|E_{\text{GMR2}}\beta_n - b(\text{GMR2})\| = O\left(\frac{1}{n}\right)
\end{aligned}$$

and the last equality is true (as β_n has a uniform Edgeworth expansion on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$, assumption A.8, and apply lemma 3.1). Taking into account equation 3 we get that, for some $C > 0$, locally independent of θ

$$\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\|\text{GMR2} - \theta\| > C\ln^{1/2} n\right) = o\left(n^{-a^*}\right).$$

This implies condition UTIGHT of lemma AL.5. It also, along with lemmas 2.4 and 3.1, implies that for any $\varepsilon_* < \varepsilon$

$$\sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)} P_\theta\left(\left\|\sqrt{n}\left(\beta_n - E_{\text{GMR2}}\beta_n\right) - \Gamma_n(\theta)\right\| > \gamma_n\right) = o\left(n^{-a}\right)$$

where $\gamma_n = o(n^{-a})$ independent of θ and

$$\begin{aligned}
\Gamma_n(\theta) &= \sqrt{n}\left(\beta_n - E_\theta\beta_n\right) - \textstyle\sum_{i=1}^{s-1}\dfrac{1}{n^{\frac{i}{2}}}\mathcal{I}_V\left(k_{i_\beta}(z,\theta)\right) \\
&\quad - \textstyle\sum_{i=1}^{s}\dfrac{1}{n^{\frac{i-1}{2}}i!}D^i\left(b(\theta) + \displaystyle\sum_{j=1}^{s-i}\frac{\mathcal{I}_V\left(k_{j_\beta}(z,\theta)\right)}{n^{\frac{j}{2}}}\right)\left(\sqrt{n}\left(\text{GMR2} - \theta\right)^i\right)
\end{aligned}$$

which validates condition EXPAND lemma AL.5 of for $Q_n = W_n^j = \text{Id}_{p\times p}$. Moreover assumption A.8 along with the fact that the support of $\beta_n - b(\theta)$ is uniformly bounded by $\overline{\mathcal{O}}_{3\eta}(0)$ for any η greater or equal than the diameter of B, and lemma 4.1 of Arvanitis and Demos [1] imply condition UEDGE of the same lemma for $M_n(\theta) = \sqrt{n}\begin{pmatrix} m_n(\theta) \\ \beta_n - E_\theta\beta_n \end{pmatrix}$ with order $s^* - 1$. Hence the conditions of lemma AL.5 are satisfied and the result follows. ii) follows the same way as i) except now $\|\beta_n - E_{\text{GMR2}}\beta_n\|$ is zero with probability $1 - o\left(n^{-a^*}\right)$ independent of θ. ■

Existence of Edgeworth Expansion for the GT Estimator

We first consider two cases which link the asymptotic behaviors of the GMR1 and the GT estimators.

Lemma 2.6 ***A***. *Suppose that* $p = q = l$, $E_{\text{GT}}(c_n(\beta_n)) = \mathbf{0}_l$ *with probability* $1 - o(n^{-a^*})$ *independent of* θ *and* $E_\theta(c_n(\beta)) = \mathbf{0}_l$ iff $\beta = b(\theta)$. *i) If the provisions of lemma 2.2.i) hold then the* GT *is uniformly consistent for* θ *with rate* $o(n^{-a})$. *ii) If the provisions of lemma 2.2.ii) hold then* $\sqrt{n}(\text{GT} - \theta)$ *has an Edgeworth expansion of order* s^* *uniformly on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ *which coincides with the one of lemma 2.2.* ***B***. *Suppose that* $q = l$, $c_n(\beta) = q_n - \beta$ *for* q_n *an appropriate q-dimensional random element and* $W_n^* = W_n^{**}$ *(*P_θ *almost everywhere for all* θ*). i) If the provisions of lemma 2.3.i) hold then the* GT *is uniformly consistent for* θ *with rate* $o(n^{-a})$. *ii) If the provisions of lemma 2.3.ii) or the ones of lemma 2.5 i) or ii) hold then* $\sqrt{n}(\text{GT} - \theta)$ *has an Edgeworth expansion of order* s^* *uniformly on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ *which coincides with the expansions of lemmas 2.3 or 2.5 i) or ii) respectively.*

Proof: **A**. From the assumptions we have that

$$E_{\text{GT}} c_n(\beta) = \mathbf{0}_p \text{ iff } \beta = b(\text{GT})$$

hence the GT equivalently satisfies

$$\beta_n - b(\text{GT}) = \mathbf{0}_p$$

which defines the GMR1 estimator in these special circumstances. Hence under these special assumptions we have that GMR1 = GT with probability $1 - o(n^{-a^*})$ independent of θ. The rest are trivial consequences of lemma 2.2. **B**. Similarly this special assumption implies that $\beta_n = q_n$ (P_θ almost surely for all θ). Hence $E_\theta c_n(\beta)|_{\beta=\beta_n} = E_\theta q_n - \beta_n = E_\theta \beta_n - \beta_n$. This and the assumed coincidence of the weighting matrices involved along with lemmas 2.3 or 2.5 i) or ii) imply the result.■

In a more general case, due to the definition of the particular estimator, we utilize the following two assumptions concerning the asymptotic behavior of c_n.

Assumption A.12 *Let* $Q_n = \|c_n(\beta)\|_{W_n(\beta_n^*)}$ *and*

$$\|c_n(\beta) - c_n(\beta')\| \leq \kappa_n \|\beta - \beta'\|, \text{ for all } \beta, \beta' \tag{4}$$

$\sup_{\theta\in\Theta} E_\theta \kappa_n = O(1)$ *and*

$$\sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B} \|c_n(\beta) - c(\theta,\beta)\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0 \tag{5}$$

where $c(\theta,\beta)$ *is continuous on* B *and equals zero* iff $\beta = b(\theta)$ *for any* θ. *Furthermore*

$$\sup_{\theta^*\in\Theta} \limsup_n E_{\theta^*} \|c_n(\beta)\|^2 < +\infty, \text{ for all } \beta. \tag{6}$$

Assumption A.13 *For* $\varphi = (\theta', \beta')'$, φ_0 *as before and* η *large enough for* $\overline{\mathcal{O}}_\eta(\varphi_0) \supset \overline{\mathcal{O}}_{\varepsilon_0}(\theta_0) \times \overline{\mathcal{O}}_{\varepsilon'}(b(\theta_0))$, rank $\left(\lim_{n\to\infty} \frac{\partial E_\theta c_n(b(\theta))}{\partial \theta'}\right) = p$, rank $\left(\lim_{n\to\infty} \frac{\partial E_\theta c_n(b(\theta))}{\partial \beta'}\right) = q$ *on* $\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)$, $\sup_{\varphi \in \overline{\mathcal{O}}_\eta(\varphi_0)} \left\| D^{s^*+1} E_\theta c_n(\beta) \right\| < M$.

Lemma 2.7 *i) Under the assumptions A.1, A.2, A.3, A.4, A.6, A.7.a) and A.12 the* GT *is uniformly consistent for* θ *with rate* $o(n^{-a})$. *ii) If additionally* $c(\theta, \beta) = E_\theta c_n(\beta)$ *and assumptions A.7.b), A.8 and A.13 hold then* $\sqrt{n}(\mathrm{GT} - \theta)$ *has an Edgeworth expansion of order* s^* *uniformly on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$.

Proof: i) By assumption A.12.4, we have that for $\varepsilon > 0$

$$\begin{aligned} &\sup_{\theta^* \in \Theta} P_{\theta^*}\left(\sup_{\theta \in \Theta} \|E_\theta c_n(\beta_n) - E_\theta c_n(b(\theta^*))\| > \varepsilon\right) \\ \leq\ &\sup_{\theta^* \in \Theta} P_{\theta^*}\left(\left(\sup_{\theta \in \Theta} E_\theta \kappa_n\right) \|\beta_n - b(\theta^*)\| > \varepsilon\right) = o\left(n^{-a^*}\right) \end{aligned}$$

and the equality is due to assumption A.6. Moreover due to A.12.5-6 and uniform integrability we obtain that

$$\sup_{\theta \in \Theta} \|E_\theta c_n(b(\theta^*)) - c(\theta, b(\theta^*))\| = o(1)$$

These via the triangle inequality imply that

$$\sup_{\theta^* \in \Theta} P_{\theta^*}\left(\sup_{\theta \in \Theta} \|E_\theta c_n(\beta_n) - c(\theta, b(\theta^*))\| > \varepsilon\right) = o\left(n^{-a^*}\right)$$

Hence for $q_n(\theta) = E_\theta c_n(\beta_n)$, $q(\theta^*, \theta) = c(\theta, b(\theta^*))$ and by assumptions A.7.a) lemma AL.3 applies. Hence for $\gamma(\theta) = \theta$ due to assumptions A.3, A.12 lemma AL.1 also applies proving the result.
ii) Given i), we have that $\theta_n \in \mathcal{O}_\epsilon(\theta)$ with P_θ-probability $1 - o\left(n^{-a^*}\right)$ that is locally independent of θ for any $\epsilon > 0$. For some ϵ small enough, such that $\mathcal{O}_\epsilon(\theta) \subset \mathcal{O}_{\varepsilon_0}(\theta_0)$ (which exists due to the fact that $\varepsilon_0 > \varepsilon$) due to assumption A.13, we have that condition FOC of lemma AL.4 (in the Appendix) is satisfied by the GT estimator for $Q_n = \frac{\partial E_\theta c_n(\beta_n)'}{\partial \theta}$. Furthermore assumption A.13 implies conditions HUB ($\gamma(\theta) = \theta$ hence set $\delta = \varepsilon_0$) and RANK of the same lemma. Condition TIGHT follows from A.8 lemma AL.2 of Arvanitis

and Demos [1] and as $E_\theta c_n(b(\theta))=0$ the fact that

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\|E_\theta c_n(\beta_n)\|>C_1\frac{\ln^{1/2}n}{n^{1/2}}\right)\\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(E_\theta\|c_n(\beta_n)-c_n(b(\theta))\|>C_1\frac{\ln^{1/2}n}{n^{1/2}}\right)\\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\|\beta_n-b(\theta)\|>\frac{C_1}{\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}E_\theta(\kappa_n)}\frac{\ln^{1/2}n}{n^{1/2}}\right)
\end{aligned}$$

imply that there exist $C_1>0$ large enough locally independent of θ for which the last term in the previous display is of order $o\left(n^{-a^*}\right)$. Hence lemma AL.4 applies ensuring that

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\,\|\mathrm{GT}-\theta\|>C\ln^{1/2}n\right)=o\left(n^{-a^*}\right)$$

for some $C>0$ independent of θ. Hence condition UTIGHT of lemma AL.5 holds. Moreover assumption A.8 implies condition UEDGE of the same lemma for $M_n(\theta)=\sqrt{n}m_n(\theta)$. Due to assumption A.13 for any $\varphi=\begin{pmatrix}\theta\\ b(\theta)\end{pmatrix}$ for any $\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)$ and any $\varphi_*=\begin{pmatrix}\theta_*\\ \varphi_*\end{pmatrix}$ sufficiently close to φ, $\frac{\partial E_{\theta_*}c_n(\beta_*)'}{\partial\theta}$ admits a Taylor expansion of order s^*-1 around φ of the form

$$\begin{aligned}
&\frac{\partial E_{\theta_*}c_n(\beta_*)'}{\partial\theta}\\
=\ &\frac{\partial E_{\theta_*}c_n(b(\theta_*))'}{\partial\theta}\\
&+\sum_{i_1+i_2=1}^{s^*-1}\frac{1}{i_1!i_2!}D^{i_1,i_2}\left(\frac{\partial E_\theta c_n(b(\theta))'}{\partial\theta}\right)\left((\beta_*-b(\theta))^{i_1},(\theta_*-\theta)^{i_2}\right)\\
&+\frac{1}{(s^*-1)!}\left(D^{s^*-1}\left(\frac{\partial E_{\theta^+}c_n\left(\beta^+\right)'}{\partial\theta}\right)-D^{s^*-1}\left(\frac{\partial E_\theta c_n(b(\theta))'}{\partial\theta}\right)\right)\\
&\left((\varphi_*-\varphi)^{s^*-1}\right)
\end{aligned}$$

where $\varphi^+=\begin{pmatrix}\theta^+\\ \beta^+\end{pmatrix}$ lies between φ_* and φ. This implies that for $\theta_n=\mathrm{GT}$ due to conditions UTIGHT and EXPAND we have that with P_θ-probability

$1-o\left(n^{-a^{*}}\right)$ that is independent of θ

$$\begin{aligned}
&\frac{\partial E_{\theta_n} c_n\left(\beta_n\right)'}{\partial\theta}\\
=\ &\frac{\partial E_{\theta} c_n\left(b\left(\theta\right)\right)'}{\partial\theta}\\
&+\sum_{i_1+i_2=1}^{s^*-1}\frac{1}{i_1!i_2!}D^{i_1,i_2}\left(\frac{\partial E_{\theta} c_n\left(b\left(\theta\right)\right)'}{\partial\theta}\right)\left(\left(\beta_n-b\left(\theta\right)\right)^{i_1},\left(\theta_n-\theta\right)^{i_2}\right)+R_n^*\left(\theta_n,\theta\right)
\end{aligned}$$

where $R_n^*\left(\theta_n,\theta\right)=\frac{1}{(s^*-1)!}\left(D^{s^*-1}\left(\frac{\partial E_{\theta_n^+}c_n\left(\beta_n^+\right)'}{\partial\theta}\right)-D^{s^*-1}\left(\frac{\partial E_{\theta}c_n(b(\theta))'}{\partial\theta}\right)\right)\left(\left(\varphi_n-\varphi\right)^{s^*-1}\right)$, and θ_n^+, β_n^+ lie between θ_n and θ and β_n and $b\left(\theta\right)$ respectively. Due to assumptions A.13, A.8, lemma AL.2 of Arvanitis and Demos [1] and by submultiplicativity, the relation of θ_n^+ to θ_n and condition UTIGHT

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}P_{\theta}\left(\left\|\begin{array}{c}\frac{1}{(s^*-1)!}\frac{1}{n^{(s^*-1)/2}}\left(D^{s^*-1}\left(\frac{\partial E_{\theta_n^+}c_n\left(\beta_n^+\right)'}{\partial\theta}\right)-D^{s^*-1}\left(\frac{\partial E_{\theta}c_n(b(\theta))'}{\partial\theta}\right)\right)\times\\ \left(\left(\varphi_n-\varphi\right)^{s^*-1}\right)\end{array}\right\|>\gamma_n^*\right)\\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}P_{\theta}\left(\begin{array}{c}\frac{2^{s^*-1}}{(s^*-1)!}\frac{1}{n^{(s^*-1)/2}}\sup_{\varphi\in\overline{\mathcal{O}}_{\eta}(\varphi_0)}\left\|D^{s^*}\frac{\partial E_{\theta}c_n(\beta)'}{\partial\theta}\right\|\times\\ \left(\left\|\theta_n^+-\theta\right\|+\left\|\beta_n^+-b\left(\theta\right)\right\|\right)\left(\left\|\sqrt{n}\left(\theta_n-\theta\right)\right\|^{s^*-1}+\left\|\beta_n-b\left(\theta\right)\right\|^{s^*-1}\right)>\gamma_n^*\end{array}\right)\\
\leq\ &\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}P_{\theta}\left(\frac{2^{s^*-1}M}{(s^*-1)!}\frac{\max^{s^*}\left(C,C^+\right)}{n^{s^*/2}}\ln^{s^*/2}n>\gamma_n^*\right)+o\left(n^{-a^*}\right)
\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^*=\frac{2^{s^*-1}M}{(s^*-1)!}\frac{\max^{s^*}\left(C,C^+\right)}{n^{s^*/2}}\ln^{s^*/2}n=o\left(n^{-a^*}\right)$ and independent of θ. Furthermore, due to the same assumption and the fact that $c\left(\theta,b\left(\theta\right)\right)=\mathbf{0}$ we have that

$$\begin{aligned}
q_n\ =\ &E_{\theta_*}c_n\left(\beta_*\right)=\sum_{i_1+i_2=1}^{s^*}\frac{1}{i_1!i_2!}D^{i_1,i_2}E_{\theta}c_n\left(b\left(\theta\right)\right)\left(\left(\beta_*-b\left(\theta\right)\right)^{i_1},\left(\theta_*-\theta\right)^{i_2}\right)\\
&+\frac{1}{(s^*-1)!}\left(D^{s^*-1}\left(\frac{\partial E_{\theta^+}c_n\left(\beta^+\right)'}{\partial\theta}\right)-D^{s^*-1}\left(\frac{\partial E_{\theta}c_n\left(b\left(\theta\right)\right)'}{\partial\theta}\right)\right)\left(\left(\varphi_*-\varphi\right)^{s^*-1}\right)
\end{aligned}$$

where θ^+ lies between θ_* and θ. Hence with P_θ-probability $1-o\left(n^{-a^*}\right)$ locally independent of θ

$$\begin{aligned}&\sqrt{n}E_{\theta_n}c_n\left(\beta_n\right)\\ =\ & \sum_{i_1+i_2=0}^{s^*-1}\frac{1}{(i_1+1)!\,(i_2+1)!}\frac{1}{n^{i_1/2}}\frac{1}{n^{i_2/2}}D^{(i_1+1),(i_2+1)}E_\theta c_n\left(b\left(\theta\right)\right)\\ &\left(\left(\sqrt{n}\left(\beta_n-b\left(\theta\right)\right)\right)^{i_1+1},\left(\theta_n-\theta\right)^{i_2+1}\right)+R_n^{\#}\left(\theta_n,\theta\right)\end{aligned}$$

where $R_n^{\#}\left(\theta_n,\theta\right)=\frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\left(D^{s^*}E_{\theta_n^+}c_n\left(\beta_n^+\right)-D^{s^*}E_\theta c_n\left(b\left(\theta\right)\right)\right)\left(\left(\sqrt{n}\left(\varphi_n-\varphi\right)\right)^{s^*}\right)$, and θ_n^+ lies between θ_n and θ. Hence analogously to the previous

$$\begin{aligned}&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}P_\theta\left(\left\|\begin{array}{c}\frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\left(D^{s^*}E_{\theta_n^+}c_n\left(\beta_n^+\right)-D^{s^*}E_\theta c_n\left(b\left(\theta\right)\right)\right)\times\\ \left(\left(\sqrt{n}\left(\varphi_n-\varphi\right)\right)^{s^*}\right)\end{array}\right\|>\gamma_n^{\#}\right)\\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}P_\theta\left(\begin{array}{c}\frac{1}{s^*!}\frac{1}{n^{(s^*-1)/2}}\sup_{\varphi\in\overline{\mathcal{O}}_\eta(\varphi_0)}\left\|D^{s^*+1}E_\theta c_n\left(\beta\right)\right\|\\ \times\left(\left\|\theta_n^+-\theta\right\|+\left\|\beta_n^+-b\left(\theta\right)\right\|\right)\left(\left\|\sqrt{n}\left(\theta_n-\theta\right)\right\|^{s^*}+\left\|\beta_n-b\left(\theta\right)\right\|^{s^*}\right)>\gamma_n^{\#}\end{array}\right)\\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}P_\theta\left(\frac{2^{s^*}M}{s^*!}\frac{\max^{s^*+1}\left(C,C^+\right)}{n^{s^*/2}}\ln^{(s^*+1)/2}n>\gamma_n^{\#}\right)+o\left(n^{-a^*}\right)\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n^{\#}=\frac{2^{s^*}M}{s^*!}\frac{\max^{s^*+1}\left(C,C^+\right)}{n^{s^*/2}}\ln^{(s^*+1)/2}n=o\left(n^{-a^*}\right)$ and independent of θ. Then due to assumption A.13 and the fact that $E_\theta c_n\left(\beta\right)=c\left(\theta,\beta\right)$, $\frac{\partial E_\theta c_n(b(\theta))'}{\partial\theta}$, $\frac{\partial E_\theta c_n(b(\theta))'}{\partial\beta}$ are of full rank for any $\theta\in\overline{\mathcal{O}}_\varepsilon\left(\theta_0\right)$. Finally due to lemma AL.6 which applies by assumptions A.8 and A.7 condition EXPAND holds and the result follows by the same lemma.■

3 Validity of 1st and 2nd Moment Expansions

Lemma 3.1 *Suppose that K is a m-linear real function on $\mathbb{R}^w$, the support of an $\mathbb{R}^w$ valued random element (say) ζ_n is bounded by $\mathcal{O}_{\sqrt{n}\rho}\left(0\right)$ for some $\rho>0$, and ζ_n admits an Edgeworth expansion of order $s^*=2a+m+1$ then*

$$\left|\int_{\mathbb{R}^q}K\left(z^m\right)\left(dP_n-\left(1+\sum\nolimits_{i=1}^{s-1}\frac{\pi_i\left(z\right)}{n^{\frac{i}{2}}}\right)\varphi_V\left(z\right)dz\right)\right|=o\left(n^{-a}\right)$$

where P_n, and $\left(1+\sum\nolimits_{i=1}^{s}\frac{\pi_i(z)}{n^{\frac{i}{2}}}\right)\varphi_V\left(z\right)$ denote the distribution of ζ_n and the density of the analogous Edgeworth measure of order $s=2a+1$ respectively. Moreover if P_n depends on θ, and $\pi_i\left(z\right)$ are continuous on $\overline{\mathcal{O}}_\varepsilon\left(\theta_0\right)$ for any z, V is continuous on $\overline{\mathcal{O}}_\varepsilon\left(\theta_0\right)$ and the expansion is uniformly valid on $\overline{\mathcal{O}}_\varepsilon\left(\theta_0\right)$, the approximation holds uniformly on $\overline{\mathcal{O}}_\varepsilon\left(\theta_0\right)$.

Proof. Let Q_n denote the measure with density $\left(1+\sum_{i=1}^{s-1}\frac{\pi_i(z)}{n^{\frac{i}{2}}}\right)\varphi_V(z)$. Since $2a+m+1>2a+1$, we have that $\sup_{A\in\mathcal{B}_C}|P_n(A)-Q_n(A)|=O\left(n^{-a-\eta}\right)$, where $\eta>0$. Hence

$$
\begin{aligned}
& n^a\left|\int_{\mathbb{R}^q}K(x^m)(dP_n-dQ_n)\right|\le n^a\left|\int_{\mathcal{O}_{c(\ln n)^\epsilon}(0)}K(x^m)(dP_n-dQ_n)\right| \\
& +n^a\left|\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}K(x^m)\,dP_n\right|+n^a\left|\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}K(x^m)\,dQ_n\right| \\
\le\ & n^aM(\ln n)^{m\epsilon}\int_{\mathcal{O}_{c(\ln n)^\epsilon}(0)}|dP_n-dQ_n|+n^a\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,(dP_n+|dQ_n|) \\
\le\ & M(\ln n)^{m\epsilon}\sup_{A\in\mathcal{B}_C}n^a|P_n(A)-Q_n(A)|+n^a\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,(dP_n+|dQ_n|)
\end{aligned}
$$

Due to the hypothesis for the support of P_n

$$
\begin{aligned}
& n^a\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,dP_n \\
=\ & n^a\int_{\left[\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)\right]\cap\mathcal{O}_{\sqrt{n}\rho}(0)}|K(x^m)|\,dP_n+n^a\int_{\left[\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)\right]\cap\left(\mathcal{O}_{\sqrt{n}\rho}(0)\right)^c}|K(x^m)|\,dP_n \\
=\ & n^a\int_{\left[\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)\right]\cap\mathcal{O}_{\sqrt{n}\rho}(0)}|K(x^m)|\,dP_n=n^a\int_{\mathcal{O}_{\sqrt{n}\rho}(0)\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,dP_n \\
\le\ & n^{a+m\beta}\rho^mq^m\int_{\mathbb{R}^q}1_{\|x\|>c(\ln n)^\epsilon}dP_n
\end{aligned}
$$

Hence

$$
\begin{aligned}
& n^a\left|\int_{\mathbb{R}^q}x^m(dP_n-dQ_n)\right|\le M(\ln n)^{m\epsilon}\sup_{A\in\mathcal{B}_C}n^a|P_n(A)-Q_n(A)| \\
& +n^{a+m\beta}\rho^mq^mP\left(\|\zeta_n\|>c(\ln n)^\epsilon\right)+n^a\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,|dQ_n|\,.
\end{aligned}
$$

As $\sup_{A\in\mathcal{B}_C}n^a|P_n(A)-Q_n(A)|=O\left(n^{-\eta}\right)$ for $\eta>0$, we have that

$$(\ln n)^{2\epsilon}\sup_{A\in\mathcal{B}_C}n^a|P_n(A)-Q_n(A)|=o(1)$$

and $n^{a+\frac{m}{2}}\rho^mq^mP\left(\|\zeta_n\|>c(\ln n)^\epsilon\right)=o(1)$ if $\epsilon\ge\frac{1}{2}$ and $c\ge\sqrt{2a+m+1}$ by lemma 2 of Magdalinos [5]. Finally $n^a\int_{\mathbb{R}^q\setminus\mathcal{O}_{c(\ln n)^\epsilon}(0)}|K(x^m)|\,|dQ_n|=o(1)$

due to Gradshteyn and Ryzhik [4] formula 8.357. For the uniform case first notice that

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_{\theta}\left(\|\zeta_n\| > M\ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

This is due to the fact that the set $\left\{x\in\mathbb{R}^q : \|x\| \leq M\ln^{1/2} n\right\}$ has boundary of Lebesgue measure zero and

$$\begin{aligned}
&\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} \int_{\|x\|>M\ln^{1/2} n} \left(1+\sum\nolimits_{i=1}^{s^*} \frac{1}{n^{\frac{i}{2}}}\left|\pi_i(x,\theta)\right|\right)\varphi_{V(\theta)}(x)\,dx \\
\leq\; &\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} \int_{\|z\|>\frac{M}{\lambda_{\max}(\theta)}\ln^{1/2} n} \left(1+\sum\nolimits_{i=1}^{s^*} \frac{1}{n^{\frac{i}{2}}}\left|\pi_i\left(V^{1/2}(\theta)\,z,\theta\right)\right|\right)\varphi(z)\,dz \\
\leq\; &\int_{\|z\|>\frac{M}{\lambda_{\max}(\theta^*)}\ln^{1/2} n} \left(1+\sum\nolimits_{i=1}^{s^*} \frac{1}{n^{\frac{i}{2}}}\left|\pi_i\left(V^{1/2}(\theta_i^*)\,z,\theta_i^*\right)\right|\right)\varphi(x)\,dx
\end{aligned}$$

where $\lambda_{\max}(\theta)$ denotes the maximum absolute eigenvalue of $V^{1/2}(\theta)$ and $\theta_i^*\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ exist for all $i=1,\ldots,s^*$ due to the continuity and are independent of z due to the positivity and the fact that π_i are polynomials in x, and θ^* exists due to continuity of V and the compactness of $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$. For $M\geq s^*\lambda_{\max}(\theta^*)$ the result follows from lemma 2 of Magdalinos [5]. The rest follows in the same spirit of the first part. ■

Remark R.1 *Notice that in the case that the support of ζ_n is not bounded the previous result would hold for $s^*=2a+m+2$. This follows easily from the previous proof by letting $\rho=\ln^{\epsilon} n$ and by the fact that the Edgeworth approximation is uniform w.r.t. the Borel algebra.*

In the following we suppress the dependence on θ and z where possible for notational convenience. For the rest of this section we denote by $b=b(\theta)$, $b_{,j}$ is the j^{th} element of b, $W^*=E_{\theta}W^*(\theta)$, $W^*_{j,j'}$ is the (j,j') element of W^*, and analogously for W^{**}. Moreover, $\mathcal{C}=\frac{\partial b'}{\partial\theta}W^*\frac{\partial b}{\partial\theta'}$, $k_{i_\beta}(z,\theta)=\mathrm{pr}_{1,q}(z)\,\pi_{i-1}(z,\theta)$, $k_{i_{\theta^*}}(z,\theta)=\mathrm{pr}_{q+1,p+q}(z)\,\pi_{i-1}(z,\theta)$, $k_{i_{w^*}}(z,\theta)$ is the matrix containing the elements $\mathrm{pr}_{p+q+1,q^2}(z)\,\pi_{i-1}(z,\theta)$ and $k_{i_{w^*}}(z,\theta)$ is the matrix containing the elements of $\mathrm{pr}_{q^2+1,2q^2}(z)\,\pi_{i-1}(z,\theta)$ at the appropriate orders.

3.1 Valid 2^{nd} order Bias approximation for the Indirect estimators

GMR1 Estimator

Lemma 3.2 *Let θ_n denote the* GMR1 *estimator. If assumptions A.1, A.2, A.3, A.4, A.6 and A.7, A.8, A.9 and A.10 hold with $s^* \geq 3$ then uniformly over $\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)$*

$$\left\| E_\theta \sqrt{n}\left(\theta_n - \theta\right) - \frac{\xi_1(\theta)}{\sqrt{n}} \right\| = o\left(n^{-\frac{1}{2}}\right)$$

where

$$\begin{aligned}
\xi_1(\theta) &= \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right) \\
&\quad -\frac{1}{2}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_{\varphi_{V^*}}\left(\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right)'\frac{\partial b_j}{\partial\theta\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right]_{j=1,\ldots,q}\right) \\
&\quad +\mathcal{C}^{-1}\mathcal{I}_{\varphi_{V^*}}\left(\left(\begin{array}{c}\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right)'\frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p}W^*(\theta) \\ +\frac{\partial b'}{\partial\theta}k_{1_{w^*}}+\frac{\partial b'}{\partial\theta}\left[\frac{\partial}{\partial\theta'}W^*_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q}\end{array}\right)\left(Id_q-\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\right)k_{1_\beta}\right),
\end{aligned}$$

where $\mathcal{C} = \frac{\partial b'}{\partial\theta}W^*\frac{\partial b}{\partial\theta'}$.

Proof. Our assumptions and lemmas 2.2, 3.1 ensure the validity of the mean approximation. Then from theorem 3.1 of Arvanitis and Demos [1] we have that the relevant moment approximation can be obtained if $\sqrt{n}\left(\beta_n - b\left(\theta_n\right)\right)$ is approximated by

$$\sqrt{n}\left(\beta_n - b(\theta)\right) - \frac{\partial b}{\partial\theta'}\sqrt{n}\left(\theta_n-\theta\right) - \frac{1}{2\sqrt{n}}\left[\sqrt{n}\left(\theta_n-\theta\right)'\frac{\partial b_j}{\partial\theta\partial\theta'}\sqrt{n}\left(\theta_n-\theta\right)\right]_{j=1,\ldots,q}$$

Moreover $W^*_n\left(\theta^*_n\right)$ is appropriately approximated by

$$W^*_n(\theta) + \frac{1}{\sqrt{n}}\left[\frac{\partial}{\partial\theta'}W^*_n(\theta)_{j,j'}\sqrt{n}\left(\theta^*_n-\theta\right)\right]_{j,j=1,\ldots,q}$$

that is by

$$W^*(\theta) + \frac{1}{\sqrt{n}}k_{1_{w^*}} + \frac{1}{\sqrt{n}}\left[\frac{\partial}{\partial\theta'}W^*(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q}$$

and analogously $\frac{\partial b'(\theta_n)}{\partial\theta}$ is appropriately approximated by

$$\frac{\partial b'}{\partial\theta} + \frac{1}{\sqrt{n}}\left[\sqrt{n}\left(\theta^*_n-\theta\right)'\frac{\partial^2 b_j}{\partial\theta\partial\theta'}\right]_{j=1,\ldots,l}$$

hence by

$$\frac{\partial b'}{\partial\theta} + \frac{1}{\sqrt{n}}\left[\sqrt{n}\left(\theta_n-\theta\right)'\frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p}$$

Therefore an appropriate approximation for $\sqrt{n}(\theta_n - \theta)$ is obtained by inverting

$$\left(\frac{\partial b'}{\partial \theta} W^*(\theta) + \frac{1}{\sqrt{n}} \left(\begin{array}{c} \left[\sqrt{n}(\theta_n-\theta)' \frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p} W^*(\theta) \\ + \frac{\partial b'}{\partial\theta} k_{1_{w^*}} + \frac{\partial b'}{\partial\theta} \left[\frac{\partial}{\partial\theta'} W^*(\theta)_{j,j'} k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q} \end{array} \right)\right) \times$$

$$\left(\sqrt{n}(\beta_n - b(\theta)) - \frac{\partial b}{\partial\theta'}\sqrt{n}(\theta_n-\theta) - \frac{1}{2\sqrt{n}}\left[\sqrt{n}(\theta_n-\theta)' \frac{\partial b_j}{\partial\theta\partial\theta'}\sqrt{n}(\theta_n-\theta)\right]_{j=1,\ldots,q}\right)$$

and for $\mathcal{C} = \frac{\partial b'}{\partial\theta} W^* \frac{\partial b}{\partial\theta'}$, $\sqrt{n}(\theta_n-\theta)$ is approximated by

$$\begin{array}{l} \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta} W^*(\theta)\sqrt{n}(\beta_n - b(\theta)) \\ +\frac{1}{\sqrt{n}}\mathcal{C}^{-1} \left(\begin{array}{c} \left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta} W^*(\theta) k_{1_\beta}\right)' \frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p} W^*(\theta) \\ + \frac{\partial b'}{\partial\theta} k_{1_{w^*}} + \frac{\partial b'}{\partial\theta}\left[\frac{\partial}{\partial\theta'} W^*(\theta)_{j,j'} k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q} \end{array}\right) \left(Id_q - \frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*(\theta)\right) k_{1_\beta} \\ -\frac{1}{2\sqrt{n}}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*(\theta)\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*(\theta)k_{1_\beta}\right)' \frac{\partial b_j}{\partial\theta\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*(\theta)k_{1_\beta}\right]_{j=1,\ldots,q}. \end{array}$$

Integrating with respect to $\left(1 + \frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V^*(\theta)}(z)$, noting that $k_{1_\beta}(z,\theta) = z$, $k_{2_\beta}(z,\theta) = z\pi_1(z,\theta)$ we obtain that

$$\left\| E_\theta\sqrt{n}(\theta_n - \theta) - \frac{\xi_1(\theta)}{\sqrt{n}}\right\| = o\left(n^{-\frac{1}{2}}\right)$$

where

$$\begin{array}{rcl} \xi_1(\theta) & = & \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right) \\ & & -\frac{1}{2}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_{\varphi_{V^*}}\left(\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right)'\frac{\partial b_j}{\partial\theta\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right]_{j=1,\ldots,q}\right) \\ & & +\mathcal{C}^{-1}\mathcal{I}_{\varphi_{V^*}}\left(\left(\begin{array}{c}\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right)'\frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p} W^*(\theta) \\ +\frac{\partial b'}{\partial\theta}k_{1_{w^*}} + \frac{\partial b'}{\partial\theta}\left[\frac{\partial}{\partial\theta'}W^*_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q}\end{array}\right)\left(Id_q - \frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\right)k_{1_\beta}\right), \end{array}$$

where the dependences of $W^*(\theta)$ and $b(\theta)$ on θ have been suppressed. ■

It follows trivially.

Corollary 1 *When W^* is independent of x and θ and $b(\theta)$ is affine then*

$$\xi_1(\theta) = \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_V\left(k_{2_\beta}\right)$$

GMR2 **Estimator**

Lemma 3.3 *Let θ_n denote the* GMR2 *estimator. If assumptions A.1, A.2, A.3, A.4, A.6 and A.7, A.8, A.9, A.10 and A.11 hold for $s^* \geq 4$ then uniformly over $\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)$ for any $\varepsilon_* < \varepsilon$*

$$\left\| E_\theta \sqrt{n}\left(\theta_n - \theta\right) - \frac{\xi_2(\theta)}{\sqrt{n}} \right\| = o\left(n^{-\frac{1}{2}}\right)$$

where

$$\xi_2(\theta) = \xi_1(\theta) - \mathcal{C}^{-1}\frac{\partial b'}{\partial \theta} W^* \mathcal{I}_V\left(k_{2_\beta}\right)$$

Proof. The assumptions and lemmas 2.2, 3.1 ensure the validity of the mean approximation uniformly over $\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)$. Furthermore from lemma AL.7 we get that $\sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)} \left\| DE_\theta \beta_n - Db(\theta) - \frac{1}{n} D\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)(\theta) \right\| = o(1)$ (recall that $\mathcal{I}_{\varphi_{V^*}}\left(k_{1_\beta}\right) = \mathbf{0}$). Then from theorem 3.1 of Arvanitis and Demos [1] we get that the relevant moment approximation can be obtained if $\sqrt{n}\left(\beta_n - E_{\theta_n}\beta_n\right)$ is approximated by

$$\sqrt{n}\left(\beta_n - b(\theta)\right) - \frac{\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)}{\sqrt{n}} - \left(\frac{\partial b}{\partial \theta'} + \frac{1}{n}\frac{\partial \mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)}{\partial \theta'}\right)\sqrt{n}\left(\theta_n - \theta\right)$$
$$-\frac{1}{2\sqrt{n}}\left[\sqrt{n}\left(\theta_n - \theta\right)' \frac{\partial\left(b + \frac{1}{n}\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)\right)_j}{\partial \theta \partial \theta'}\sqrt{n}\left(\theta_n - \theta\right)\right]_{j=1,\ldots,q}$$

$W_n^*\left(\theta_n^*\right)$ is the same as the proof of lemma 3.2 before and analogously $\frac{\partial E_{\theta_n}\beta_n}{\partial \theta}$ is appropriately approximated by

$$\frac{\partial\left(b + \frac{1}{n}\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)\right)'}{\partial \theta} + \frac{1}{\sqrt{n}}\left[\sqrt{n}\left(\theta_n - \theta\right)' \frac{\partial^2\left(b + \frac{1}{n}\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right)\right)'}{\partial \theta \partial \theta_j}\right]_{j=1,\ldots,p}$$

hence by

$$\frac{\partial b'}{\partial \theta} + \frac{1}{\sqrt{n}}\left[\sqrt{n}\left(\theta_n - \theta\right)' \frac{\partial^2 b'}{\partial \theta \partial \theta_j}\right]_{j=1,\ldots,p}$$

In this respect an approximation for $\sqrt{n}\left(\theta_n - \theta\right)$ is

$$\mathcal{C}^{-1}\frac{\partial b'}{\partial \theta} W^* \sqrt{n}\left(\beta_n - b(\theta)\right)$$
$$+\frac{1}{\sqrt{n}}\mathcal{C}^{-1}\left[\begin{array}{c}\left[\left(\mathcal{C}^{-1}\frac{\partial b'}{\partial \theta} W^*(\theta) k_{1_\beta}\right)' \frac{\partial^2 b'}{\partial \theta \partial \theta_j}\right]_{j=1,\ldots,p} W^* \\ +\frac{\partial b'}{\partial \theta} k_{1_{w^*}} + \frac{\partial b'}{\partial \theta}\left[\frac{\partial}{\partial \theta'} W^*(\theta)_{j,j'} k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q}\end{array}\right]\left(Id_q - \mathcal{C}^{-1}\frac{\partial b'}{\partial \theta} W^*(\theta)\right) k_{1_\beta}$$
$$-\frac{1}{\sqrt{n}}\mathcal{C}^{-1}\frac{\partial b'}{\partial \theta} W^*\left[\mathcal{I}_{\varphi_{V^*}}\left(k_{2_\beta}\right) + \frac{1}{2}\left[\sqrt{n}\left(\theta_n - \theta\right)' \frac{\partial b_j}{\partial \theta \partial \theta'}\sqrt{n}\left(\theta_n - \theta\right)\right]_{j=1,\ldots,q}\right]$$

Integrating with respect to $\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V^*(\theta)}(z)$, noting that $k_{1_\beta}(z,\theta)=z$, $k_{2_\beta}(z,\theta)=z\pi_1(z,\theta)$ we obtain that

$$\left\|E_\theta\sqrt{n}\left(\theta_n-\theta\right)-\frac{\xi_2(\theta)}{\sqrt{n}}\right\|=o\left(n^{-\frac{1}{2}}\right)$$

where

$$\begin{aligned}\xi_2(\theta) &= -\frac{1}{2}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_{\varphi_{V^*}}\left(\left[k_{1_\beta}'W^*\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\frac{\partial b_j}{\partial\theta\partial\theta'}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*k_{1_\beta}\right]_{j=1,\ldots,q}\right)\\ &\quad+\mathcal{C}^{-1}\mathcal{I}_{\varphi_{V^*}}\left(\left[\begin{array}{c}\left[k_{1_\beta}'W^*\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\frac{\partial^2 b'}{\partial\theta\partial\theta_j}\right]_{j=1,\ldots,p}W^*\\ +\frac{\partial b'}{\partial\theta}k_{1_{w^*}}+\frac{\partial b'}{\partial\theta}\left[\frac{\partial}{\partial\theta'}W_{j,j'}^*k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,q}\end{array}\right]\left(Id_q-\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\right)k_{1_\beta}\right),\end{aligned}$$

where the dependences of $W^*(\theta)$ and $b(\theta)$ on θ have been suppressed. Taking into account the expression of $\xi_1(\theta)$ in lemma 3.2 we get the result. ■

The following corollary is trivial and establishes general conditions under which the GMR2 estimator is second order unbiased.

Corollary 2 *When W^* is independent of x and θ and $b(\theta)$ is affine then $\xi_2(\theta)=\mathbf{0}_p$.*

GT **Estimator** Denoting with $\mathcal{D}=\frac{\partial b'}{\partial\theta}\frac{\partial c(\theta,b)}{\partial\beta'}W^{**}(\theta)\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}$, $\mathcal{E}=\frac{\partial b'}{\partial\theta}\frac{\partial c(\theta,b)}{\partial\beta'}W^{**}(\theta)$, $\mathcal{H}_j=\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\frac{\partial b}{\partial\theta'}-\left[\frac{\partial c_j(\theta,b)}{\partial\beta'}\frac{\partial^2 b}{\partial\theta'\partial\theta_r}\right]_{r=1,\ldots,p}$, $\mathcal{J}=k_{1_{w^{**}}}+\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,l}$, $\mathcal{J}^*=\left(\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\mathcal{D}^{-1}\mathcal{E}-\mathrm{Id}_l\right)\frac{\partial c(\theta,b)}{\partial\beta'}$ and $q_{1_\beta}=\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$ we obtain the following lemma.

Lemma 3.4 *Using A.12 suppose that $E_\theta c_n(\beta)=c(\theta,\beta)$. Furthermore let A.1, A.2, A.3, A.4, A.6, A.7, A.8, A.13 hold for $s^*\geq 3$, then uniformly on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$*

$$\left\|E_\theta\sqrt{n}\left(\theta_n-\theta\right)-\frac{\xi_3(\theta)}{\sqrt{n}}\right\|=o\left(n^{-\frac{1}{2}}\right)$$

where

$$\begin{aligned}
\xi_3(\theta) &= \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\mathcal{I}_V\left(k_{2_\beta}\right)+\frac{1}{2}\mathcal{D}^{-1}\mathcal{E}\left[\mathcal{I}_V\left(k_{1_\beta}'\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}k_{1_\beta}\right)\right]_{j=1,\ldots,l}\\
&-\mathcal{D}^{-1}\mathcal{E}\left[\mathcal{I}_V\left(q_{1_\beta}'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}k_{1_\beta}\right)\right]_{j=1,\ldots,l}+\frac{1}{2}\mathcal{D}^{-1}\mathcal{E}\left[\mathcal{I}_V\left(q_{1_\beta}'\mathcal{H}_j q_{1_\beta}\right)\right]_{j=1,\ldots,l}\\
&+\mathcal{D}^{-1}\mathcal{I}_V\left(\left[\mathcal{H}_j q_{1_\beta}-\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l}W^{**}(\theta)\mathcal{J}^* k_{1_\beta}\right)\\
&-\mathcal{D}^{-1}\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\mathcal{I}_V\left(\mathcal{J}\mathcal{J}^* k_{1_\beta}\right).
\end{aligned}$$

Proof: The assumptions and lemmas 2.2, 3.1 ensure the validity of the mean approximation. Then theorem 3.1 of Arvanitis and Demos implies that the relevant moment approximation can be obtained as follows. Due to the fact that $c(\theta, b(\theta)) = \mathbf{0}_l$ we obtain, by the implicit function theorem, that

$$\frac{\partial c(\theta,\beta)}{\partial\beta'}|_b\frac{\partial b}{\partial\theta'}=-\frac{\partial c(\theta,\beta)}{\partial\theta'}|_b$$

Moreover as $\frac{\partial c(\theta,b(\theta))}{\partial\theta'} = \mathbf{0}_{lxp}$ we have, by the same theorem, that for any j, we obtain that

$$\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\theta'}|_b=-\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'},$$

and

$$\begin{aligned}
&\frac{\partial}{\partial\theta'}\left(\frac{\partial c_j(\theta,\beta)}{\partial\theta}|_b+\frac{\partial b'}{\partial\theta}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right)=\frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\theta'}|_b\\
&+\frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}+\frac{\partial}{\partial\theta'}\left(\frac{\partial b'}{\partial\theta}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right)\\
=&\ \frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\theta'}|_b+\frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}+\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}\\
=&\ \frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\theta'}|_b-\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}+\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}
\end{aligned}$$

and therefore

$$\frac{\partial^2 c_j(\theta,\beta)}{\partial\theta\partial\theta'}|_b=\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}-\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}$$

Now

$$\begin{aligned}\sqrt{n}c\left(\theta_n,\beta_n\right) &= \frac{\partial c\left(\theta,b\right)}{\partial\beta'}\sqrt{n}\left(\beta_n-b\right)-\frac{\partial c\left(\theta,b\right)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\sqrt{n}\left(\theta_n-\theta\right)\\ &+\frac{1}{2\sqrt{n}}\left[tr\sqrt{n}\left(\beta_n-b\right)\sqrt{n}\left(\beta_n-b\right)^{/}\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}\right]_{j=1,\ldots,l}\\ &+\frac{1}{\sqrt{n}}\left[tr\sqrt{n}\left(\beta_n-b\right)\sqrt{n}\left(\theta_n-\theta\right)^{/}\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\beta\partial\theta'}\right]_{j=1,\ldots,l}\\ &+\frac{1}{2\sqrt{n}}\left[tr\sqrt{n}\left(\theta_n-\theta\right)\sqrt{n}\left(\theta_n-\theta\right)^{/}\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\theta\partial\theta'}\right]_{j=1,\ldots,l}\end{aligned}$$

and it follows that

$$\begin{aligned}&\sqrt{n}c\left(\theta_n,\beta_n\right)=\frac{\partial c\left(\theta,b\right)}{\partial\beta'}k_{1_\beta}-\frac{\partial c\left(\theta,b\right)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\sqrt{n}\left(\theta_n-\theta\right)\\ &+\frac{1}{2\sqrt{n}}\left[trk_{1_\beta}k_{1_\beta}'\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}\right]_{j=1,\ldots,l}\\ &-\frac{1}{\sqrt{n}}\left[trk_{1_\beta}\sqrt{n}\left(\theta_n-\theta\right)^{/}\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}\right]_{j=1,\ldots,l}\\ &+\frac{1}{2\sqrt{n}}\left[tr\sqrt{n}\left(\theta_n-\theta\right)\sqrt{n}\left(\theta_n-\theta\right)^{/}\left(\begin{array}{c}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'}\\ -\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}\end{array}\right)\right]_{j=1,\ldots,l}\end{aligned}$$

Moreover $W_n^{**}\left(\theta_n^*\right)$ is appropriately approximated by

$$W_n^{**}\left(\theta_n^*\right)=W_n^{**}\left(\theta\right)+\frac{1}{\sqrt{n}}\left[\frac{\partial}{\partial\theta^{/}}W_n^{**}\left(\theta\right)_{rj}\sqrt{n}\left(\theta_n^*-\theta\right)\right]_{r,j=1,\ldots,l}$$

that is by

$$W_n^{**}\left(\theta_n^*\right)=W^{**}\left(\theta\right)+\frac{1}{\sqrt{n}}k_{1_{w^{**}}}+\frac{1}{\sqrt{n}}\left[\frac{\partial}{\partial\theta^{/}}W^{**}\left(\theta\right)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,l}$$

and analogously $\frac{\partial c'(\theta_n,\beta_n)}{\partial\theta}$ is appropriately approximated by

$$\begin{aligned}&\frac{\partial c'\left(\theta,b\right)}{\partial\theta}+\frac{1}{\sqrt{n}}\left[\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\theta\partial\beta'}|_b\sqrt{n}\left(\beta_n-b\right)\right]_{j=1,\ldots,l}'\\ &+\frac{1}{\sqrt{n}}\left[\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\theta\partial\theta'}|_b\sqrt{n}\left(\theta_n-\theta\right)\right]_{j=1,\ldots,l}^{/}\end{aligned}$$

that is by

$$\frac{\partial c'(\theta_n,\beta_n)}{\partial\theta} = -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta} - \frac{1}{\sqrt{n}}\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\dots,l} + \frac{1}{\sqrt{n}}\left[\begin{pmatrix} \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,\beta)}{\partial\beta'}|_b\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\dots,p}\end{pmatrix}\sqrt{n}(\theta_n-\theta)\right]_{j=1,\dots,l}.$$

$$0 = -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta} + \frac{1}{\sqrt{n}}\left\{\begin{matrix} -\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\dots,l}W^{**}(\theta) \\ +\left[\begin{pmatrix} \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,\beta)}{\partial\beta'}|_b\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\dots,p}\end{pmatrix}\sqrt{n}(\theta_n-\theta)\right]_{j=1,\dots,l} W^{**}(\theta) \\ -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}k_{1_{w^{**}}} - \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\dots,l}\end{matrix}\right\}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$$
$$+\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\sqrt{n}(\theta_n-\theta)$$
$$-\frac{1}{\sqrt{n}}\left\{\begin{matrix} -\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\dots,l}W^{**}(\theta) \\ +\left[\begin{pmatrix} \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,b)}{\partial\beta'}\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\dots,p}\end{pmatrix}\sqrt{n}(\theta_n-\theta)\right]_{j=1,\dots,l} W^{**}(\theta) \\ -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}k_{1_{w^{**}}} - \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\dots,l}\end{matrix}\right\}$$
$$\times\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\sqrt{n}(\theta_n-\theta)$$
$$-\frac{1}{2\sqrt{n}}\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\left[tr k_{1_\beta}k'_{1_\beta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\right]_{j=1,\dots,l}$$
$$+\frac{1}{\sqrt{n}}\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\left[tr k_{1_\beta}\sqrt{n}(\theta_n-\theta)'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\right]_{j=1,\dots,l}$$
$$-\frac{1}{2\sqrt{n}}\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\left[tr\sqrt{n}(\theta_n-\theta)\sqrt{n}(\theta_n-\theta)'\begin{pmatrix}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\dots,p}\end{pmatrix}\right]_{j=1,\dots,l}$$

$\mathcal{E} = \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)$, It follows that $\sqrt{n}(\theta_n-\theta) = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$

$$-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left\{\begin{matrix} -\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\dots,l}W^{**}(\theta) \\ +\left[\begin{pmatrix} \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,\beta)}{\partial\beta'}|_b\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\dots,p}\end{pmatrix}\sqrt{n}(\theta_n-\theta)\right]_{j=1,\dots,l} W^{**}(\theta) \\ -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}k_{1_{w^{**}}} - \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\dots,l}\end{matrix}\right\}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$$

$$+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left\{\begin{array}{c} -\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l} W^{**}(\theta) \\ +\left[\left(\begin{array}{c}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,\beta)}{\partial\beta'}|_b\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\ldots,p}\end{array}\right)\sqrt{n}\left(\theta_n-\theta\right)\right]_{j=1,\ldots,l} W^{**}(\theta) \\ -\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}k_{1_{w^{**}}}-\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,l}\end{array}\right\}$$

$$\times\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\sqrt{n}\left(\theta_n-\theta\right)+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr k_{1_\beta}k'_{1_\beta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\right]_{j=1,\ldots,l}$$

$$-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}\right)'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l}$$

$$+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr\sqrt{n}\left(\theta_n-\theta\right)\sqrt{n}\left(\theta_n-\theta\right)'\left(\begin{array}{c}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}\end{array}\right)\right]_{j=1,\ldots,l},$$

where $\mathcal{D}=\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}$ and $\mathcal{E}=\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}(\theta)$. It follows that $\sqrt{n}\left(\theta_n-\theta\right)=\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$

$$+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left\{\begin{array}{c} -\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l} W^{**}(\theta) \\ +\left[\left(\begin{array}{c}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial c_j(\theta,\beta)}{\partial\beta'}|_b\frac{\partial^2 b}{\partial\theta'\partial\theta_{j'}}\right]_{j'=1,\ldots,p}\end{array}\right)\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}\right]_{j=1,\ldots,l} W^{**}(\theta)\end{array}\right\}\mathcal{J}^* k_{1_\beta}$$

$$-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}\mathcal{J}\mathcal{J}^* k_{1_\beta}+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr k_{1_\beta}k'_{1_\beta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\right]_{j=1,\ldots,l}$$

$$-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}\right)'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l}$$

$$+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[tr\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}\right)'\left(\begin{array}{c}\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,\beta)}{\partial\beta\partial\beta'}|_b\frac{\partial b}{\partial\theta'} \\ -\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,\beta)}{\partial\beta}|_b\right]_{j'=1,\ldots,p}\end{array}\right)\right]_{j=1,\ldots,l}$$

where $\mathcal{J}=k_{1_{w^{**}}}+\left[\frac{\partial}{\partial\theta'}W^{**}(\theta)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,l}$, $\mathcal{J}^*=\left(\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\mathcal{D}^{-1}\mathcal{E}-\mathrm{Id}_l\right)\frac{\partial c(\theta,b)}{\partial\beta'}$.

Again as $\left[\frac{\partial^2 b'}{\partial\theta\partial\theta_{j'}}\frac{\partial c_j(\theta,b)}{\partial\beta}\right]_{j'=1,\ldots,p} = \left[\frac{\partial c_j(\theta,b)}{\partial\beta'}\frac{\partial^2 b}{\partial\theta'\partial\theta_r}\right]_{r=1,\ldots,p}$

$$
\begin{aligned}
&\sqrt{n}\left(\theta_n-\theta\right)=\mathcal{D}^{-1}\mathcal{E}\frac{\partial c\left(\theta,b\right)}{\partial\beta'}k_{1_\beta}\\
&+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[k_{1_\beta}'\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}k_{1_\beta}\right]_{j=1,\ldots,l}-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[q_{1_\beta}'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}k_{1_\beta}\right]_{j=1,\ldots,l}\\
&+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[q_{1_\beta}'\mathcal{H}_j q_{1_\beta}\right]_{j=1,\ldots,l}\\
&+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left[\mathcal{H}_j q_{1_\beta}-\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j\left(\theta,\beta\right)}{\partial\beta\partial\beta'}|_b k_{1_\beta}\right]_{j=1,\ldots,l}W^{**}\left(\theta\right)\mathcal{J}^* k_{1_\beta}\\
&-\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\frac{\partial b'}{\partial\theta}\frac{\partial c'\left(\theta,b\right)}{\partial\beta}\mathcal{J}\mathcal{J}^* k_{1_\beta}
\end{aligned}
$$

where $\mathcal{D} = \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}\left(\theta\right)\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}$, $\mathcal{E} = \frac{\partial b'}{\partial\theta}\frac{\partial c'(\theta,b)}{\partial\beta}W^{**}\left(\theta\right)$, $\mathcal{J} = k_{1_{w^{**}}} + \left[\frac{\partial}{\partial\theta'}W^{**}\left(\theta\right)_{j,j'}k_{1_{\theta^*}}\right]_{j,j'=1,\ldots,l}$, $\mathcal{J}^* = \left(\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\mathcal{D}^{-1}\mathcal{E}-\mathrm{Id}_l\right)\frac{\partial c(\theta,b)}{\partial\beta'}$, $q_{1_\beta} = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}$, $\mathcal{H}_j = \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\frac{\partial b}{\partial\theta'} - \left[\frac{\partial c_j(\theta,b)}{\partial\beta'}\frac{\partial^2 b}{\partial\theta'\partial\theta_r}\right]_{r=1,\ldots,p}$. Integrating the above w.r.t. $\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V(\theta)}\left(z\right)$ we get the result.■

Corollary 3 *When W^* is independent of x and θ and $b\left(\theta\right)$ is affine then*

$$
\begin{aligned}
\xi_3\left(\theta\right) \;=\;& \mathcal{D}^{-1}\mathcal{E}\frac{\partial c\left(\theta,b\right)}{\partial\beta'}\mathcal{I}_V\left(k_{2_\beta}\right)+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[\mathcal{I}_V\left(k_{1_\beta}'\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}k_{1_\beta}\right)\right]_{j=1,\ldots,l}\\
&+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[\mathcal{I}_V\left(q_{1_\beta}'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_\beta}-2k_{1_\beta}\right)\right)\right]_{j=1,\ldots,l}\\
&+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\mathcal{I}_V\left(\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j\left(\theta,b\right)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_\beta}-k_{1_\beta}\right)\right]_{j=1,\ldots,l}W^{**}\left(\theta\right)\mathcal{J}^* k_{1_\beta}\right).
\end{aligned}
$$

Moreover, even under the scope of stochastic weighting, when $p = q = l$ and $b\left(\theta\right)$ is affine, then $\xi_3\left(\theta\right) = \left(\frac{\partial b}{\partial\theta'}\right)^{-1}\mathcal{I}_V\left(k_{2_\beta}\right)$.

Proof. When W^{**} is independent of x and θ and $b\left(\theta\right)$ is affine then $\mathcal{J} = 0$ and $\mathcal{H}_j = \frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\frac{\partial b}{\partial\theta'}$. Hence by integrating w.r.t. $\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V(\theta)}\left(z\right)$ the following expression

$$
\begin{aligned}
&\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[k_{1_\beta}'\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}k_{1_\beta}\right]_{j=1,\ldots,l}\\
&+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[q_{1_\beta}'\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_\beta}-2k_{1_\beta}\right)\right]_{j=1,\ldots,l}\\
&+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_\beta}-k_{1_\beta}\right)\right]_{j=1,\ldots,l}W^{**}(\theta)\mathcal{J}^*k_{1_\beta}
\end{aligned}
$$

we get the result. On the other hand, when $p=q=l$ and $b(\theta)$ is affine then $\mathcal{D}^{-1}\mathcal{E}=\left(\frac{\partial c(\theta,b)}{\partial\beta'}\frac{\partial b}{\partial\theta'}\right)^{-1}$, $\mathcal{J}^*=0$, $q_{1_\beta}=\left(\frac{\partial b}{\partial\theta'}\right)^{-1}k_{1_\beta}$, $q_{1_\beta}=\left(\frac{\partial b}{\partial\theta'}\right)^{-1}k_{1_\beta}$, and $\mathcal{H}_j=\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\frac{\partial b}{\partial\theta'}$. Hence we the expression is

$$\left(\frac{\partial b}{\partial\theta'}\right)^{-1}k_{1_\beta},$$

and integrating the above w.r.t. $\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V(\theta)}(z)$ we get the result. ■

Lemma 3.5 *i). Under the assumptions in lemma 2.6.**A** and for $s^*\geq 3$ we have that $\xi_1(\theta)=\xi_3(\theta)$ uniformly over $\overline{\mathcal{O}}_\varepsilon(\theta_0)$. ii). Under the assumptions in lemma 2.6.**B** and for $s^*\geq 4$ we have that $\xi_2(\theta)=\xi_3(\theta)$ uniformly over $\overline{\mathcal{O}}_\varepsilon(\theta_0)$.*

Proof of Lemma 3.5. i). The result follows from lemmas 2.6.**A** and 3.2. Notice that as $p=q=l$, we have that $\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*=\left(\frac{\partial b}{\partial\theta'}\right)^{-1}$. ii). The result follows from lemmas 2.6.**B** and 3.3. ■

3.2 MSE 2^{nd} order Approximations for the Indirect Estimators

Lemma 3.6 *Let θ_n denote either the* GMR1*, or the* GMR2 *estimator. If $W^*(x,\theta)$ is independent of x and θ, b is affine and assumptions A.1, A.2, A.3, A.4, A.6 and A.7, A.8, A.9 hold for $s^*\geq 5$ then, for any $\varepsilon_*<\varepsilon$*

$$\left\|E_\theta\left(n(\theta_n-\theta)(\theta_n-\theta)'\right)-H_1(\theta)-\frac{H_2(\theta)}{\sqrt{n}}\right\|=o\left(n^{-1/2}\right)$$

where

$$
\begin{aligned}
H_1(\theta) &= \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*V(\theta)W^*\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\\
H_2(\theta) &= \mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^*\mathcal{I}_V\left(k_{2_\beta}k_{1_\beta}'\right)W^*\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}
\end{aligned}
$$

Proof. For both estimators we have that due to lemma 3.1, theorem 3.1 of Arvanitis and Demos [1] along with the approximations employed in lemmas 3.2, 3.3

$$E_{\theta}\left(n\left(\theta_{n}-\theta\right)\left(\theta_{n}-\theta\right)'\right) = \int_{\mathbb{R}^{q}}\mathcal{C}^{-1}\frac{\partial b'}{\partial\theta}W^{*}\left(k_{1_{\beta}}\left(z,\theta\right)+\frac{k_{2_{\beta}}\left(z,\theta\right)}{\sqrt{n}}\right)\left(k_{1_{\beta}}\left(z,\theta\right)+\frac{k_{2_{\beta}}\left(z,\theta\right)}{\sqrt{n}}\right)' W^{*}\frac{\partial b}{\partial\theta'}\mathcal{C}^{-1}\varphi_{V(\theta)}\left(z\right)dz,$$

where $k_{1_{\beta}}\left(z,\theta\right)=z$, $k_{2_{\beta}}\left(z,\theta\right)=z\pi_{1}\left(z,\theta\right)$. Keeping the relevant order terms, the result follows. ■

Lemma 3.7 *Let θ_n denote the* GT *estimator. If $W^{**}(x,\theta)$ is independent of x and θ, b is affine, $E_{\theta}c_{n}(\beta)=c(\theta,\beta)$ and assumptions A.1, A.2, A.3, A.4, A.6, A.7, A.8, and A.13 hold for $s^{*}\geq 4$ then, uniformly on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$*

$$\left\|E_{\theta}\left(n\left(\theta_{n}-\theta\right)\left(\theta_{n}-\theta\right)'\right)-H_{1}\left(\theta\right)-\frac{H_{2}\left(\theta\right)}{\sqrt{n}}\right\| = o\left(n^{-1/2}\right)$$

where

$$H_{1}\left(\theta\right) = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c\left(\theta,b\right)}{\partial\beta'}V\left(\theta\right)\frac{\partial c'\left(\theta,b\right)}{\partial\beta}\mathcal{E}'\mathcal{D}^{-1}$$

$$H_{2}\left(\theta\right) = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c\left(\theta,b\right)}{\partial\beta'}\mathcal{I}_{V}\left(k_{2_{\beta}}k_{1_{\beta}}'\right)\frac{\partial c'\left(\theta,b\right)}{\partial\beta}\mathcal{E}'\mathcal{D}^{-1}$$

Proof. Again we have that due to lemma 3.1, theorem 3.1 of Arvanitis and Demos [1] along with the approximations used in lemma 3 when W^{**} is independent of x and θ and $b(\theta)$ is affine, we get from the proof of lemma 3 that we have to integrate w.r.t. $\left(1+\frac{\pi_{1}(z,\theta)}{\sqrt{n}}\right)\varphi_{V(\theta)}(z)$ the following expression:

$$\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_{\beta}}k_{1_{\beta}}'\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[k_{1_{\beta}}'\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}k_{1_{\beta}}\right]_{j=1,\ldots,l}k_{1_{\beta}}'\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'$$
$$+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\left[q_{1_{\beta}}'\frac{\partial b'}{\partial\theta}\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_{\beta}}-2k_{1_{\beta}}\right)\right]_{j=1,\ldots,l}k_{1_{\beta}}'\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'$$
$$+\frac{1}{\sqrt{n}}\mathcal{D}^{-1}\left[\frac{\partial b'}{\partial\theta}\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_{\beta}}-k_{1_{\beta}}\right)\right]_{j=1,\ldots,l}W^{**}(\theta)\mathcal{J}^{*}k_{1_{\beta}}k_{1_{\beta}}'\frac{\partial c'(\theta,b)}{\partial\beta}\mathcal{E}'\mathcal{D}^{-1}$$
$$+\frac{1}{2\sqrt{n}}\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_{\beta}}\begin{bmatrix}\left(\left[k_{1_{\beta}}'\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}k_{1_{\beta}}\right]_{j=1,\ldots,l}\right)'\mathcal{E}'\\ \left(\left[q_{1_{\beta}}'\frac{\partial b'}{\partial\theta}\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_{\beta}}-2k_{1_{\beta}}\right)\right]_{j=1,\ldots,l}\right)'\mathcal{E}'\\ k_{1_{\beta}}'\left(W^{**}(\theta)\mathcal{J}^{*}\right)'\left(\left[\frac{\partial b'}{\partial\theta}\frac{\partial^{2}c_{j}(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_{\beta}}-k_{1_{\beta}}\right)\right]_{j=1,\ldots,l}\right)'\end{bmatrix}\mathcal{D}^{-1},$$

where $k_{1_\beta}(z,\theta) = z.$ and $q_{1_\beta} = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}z$. Now notice that
$\int_{\mathbb{R}^q} \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}k'_{1_\beta}\left(W^{**}(\theta)\mathcal{J}^*\right)'\left(\left[\frac{\partial b'}{\partial\theta}\frac{\partial^2 c_j(\theta,b)}{\partial\beta\partial\beta'}\left(\frac{\partial b}{\partial\theta'}q_{1_\beta} - k_{1_\beta}\right)\right]_{j=1,\ldots,l}\right)'\varphi_V(z)\,dz =$
0 as it involves the integral of $(z)^3$, which is zero-mean normally distributed. Hence by integrating the above expression w.r.t. $\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_{V(\theta)}(z)$ we get: $E_\theta\left(n(\theta_n-\theta)(\theta_n-\theta)'\right) = \int_{\mathbb{R}^q}\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}k'_{1_\beta}\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'\left(1+\frac{\pi_1(z,\theta)}{\sqrt{n}}\right)\varphi_V(z)\,dz$ and taking into account that $k_{1_\beta}(z,\theta) = z$, $k_{2_\beta}(z,\theta) = z\pi_1(z,\theta)$ and that
$\int_{\mathbb{R}^p}\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}k_{1_\beta}k'_{1_\beta}\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'\varphi_V(z)\,dz = \mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}V(\theta)\left(\mathcal{D}^{-1}\mathcal{E}\frac{\partial c(\theta,b)}{\partial\beta'}\right)'$
we get the result. ■

4 Recursive GMR2

Let $\theta_n^{(0}$ denote any estimator of θ.

Definition D.5 *Let $\zeta \in \mathbb{N}$, the recursive $\zeta -$ GMR2 estimator (denoted by $\theta_n^{(\zeta}$) is defined in the following steps:*

1. $\theta_n^{(1} = \arg\min_\theta \left\|\theta_n^{(0} - E_\theta\theta_n^{(0}\right\|$,
2. *for* $\zeta > 1$ $\theta_n^{(\zeta} = \arg\min_\theta \left\|\theta_n^{(\zeta-1} - E_\theta\theta_n^{(\zeta-1}\right\|$.

Using the results of the previous section, we are now able to prove the following lemma.

Lemma 4.1 *Suppose that assumptions A.6, A.8, A.11 hold for $\theta_n^{(0}$ for $s^* \geq 2\zeta+4$. Moreover suppose that $E_\theta\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\|n\overline{s}_n\|^2 < +\infty$ and $E_\theta\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|n\overline{H}_n\right\| < +\infty$ for all $\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)$ and $\sqrt{n}\overline{s}_n(\theta)$ admits a locally uniform Edgeworth expansion of order 6. Then the $\zeta -$ GMR2 estimator is of order $s = 2\zeta+1$ unbiased and has the same MSE with the $(\zeta-1) -$ GMR2, up to 2ζ order, uniformly over $\overline{\mathcal{O}}_{\varepsilon_*}(\theta_0)$ for any $\varepsilon_* < \varepsilon$.*

Proof. First notice that in any step of the procedure the binding function is the identity. Next the $o\left(n^{-a^*}\right)$ uniform consistency of $\theta_n^{(0}$ ensures the analogous for any step of the recursion. Then validity of the Edgeworth expansion for $\sqrt{n}\overline{s}_n(\theta)$ along with lemma 3.1 and remark R.1 imply that

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} E\left\|n\overline{s}_n(\theta)\overline{s}'_n(\theta) + E\overline{H}_n(\theta)\right\|^2 = O(1)$$

and since by the same lemma $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} E\left\|\left(\theta_n^{(0}-\theta\right)\right\|^2=O\left(\frac{1}{n}\right)$ and $E_\theta \sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}\left\|n\overline{s}_n\right\|^2<+\infty$ and $E_\theta \sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}\left\|n\overline{H}_n\right\|<+\infty$ we have that $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}\left\|D^2 E_\theta \theta_n^{(1}\right\|<M$. Hence lemma 2.5 applies and accordingly $\theta_n^{(1}$ admits a locally uniform Edgeworth expansion of order s^*. Given this the exact same reasoning implies the same result for $\theta_n^{(h}$ for any h. Moreover assumption A.11 follows for the expansions in every step of the procedure due to the previous. The proof for the moment approximations for the case $h=1$ follows easily. Using induction, let us assume that the result holds for some h, i.e. assume that the appropriate expression for $\sqrt{n}\left(\theta_n^{(h}-\theta\right)$ is given by:

$$E_\theta\sqrt{n}\left(\theta_n^{(h}-\theta\right)=\frac{1}{n^{\frac{2h+1}{2}}}\mathcal{I}_V\left(k_{2h+2}\right)+\frac{1}{n^{\frac{2h+2}{2}}}\mathcal{I}_V\left(k_{2h+3}\right)+o\left(n^{-\frac{2h+2}{2}}\right).$$

uniformly over $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$. Hence for $\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, by lemma 2.4 it follows that

$$\sqrt{n}\left(E_{\theta_n^{(h+1}}\theta_n^{(h}-E_\theta\theta_n^{(h}\right)-\left(\mathrm{Id}_p+\frac{1}{n^{\frac{2h+2}{2}}}\frac{\partial\mathcal{I}_V\left(k^{/}_{2h+2}\right)}{\partial\theta}\right)\sqrt{n}\left(\theta_n^{(h+1}-\theta\right)$$

is bounded by a real sequence of order $o\left(n^{-\frac{2h+3}{2}}\right)$ that is independent of θ, with P_θ-probability $1-o\left(n^{-\frac{2h+3}{2}}\right)$ independent of θ. The $h+1^{st}$-step GMR2 estimator satisfies with P_θ-probability $1-o\left(n^{-\frac{2h+3}{2}}\right)$ independent of θ, $\theta_n^{(h}=E_{\theta_n^{(h+1}}\theta_n^{(h}$. Hence lemma 3.1 and Theorem 3.1 of Arvanitis and Demos [1] imply that the required approximation would be given by the integration of the Edgeworth density in the h^{th} step of the following approximation

$$\sqrt{n}\left(\theta_n^{(h}-\theta\right)-\left(\frac{1}{n^{\frac{2h+1}{2}}}\mathcal{I}_V\left(k_{2h+2}\right)+\frac{1}{n^{\frac{2h+2}{2}}}\mathcal{I}_V\left(k_{2h+3}\right)\right)$$

This integration gives

$$\begin{aligned}&\frac{1}{n^{\frac{2h+1}{2}}}\mathcal{I}_V\left(k_{2h+2}\right)+\frac{1}{n^{\frac{2h+2}{2}}}\mathcal{I}_V\left(k_{2h+3}\right)\\&-\left(\frac{1}{n^{\frac{2h+1}{2}}}\mathcal{I}_V\left(k_{2h+2}\right)+\frac{1}{n^{\frac{2h+2}{2}}}\mathcal{I}_V\left(k_{2h+3}\right)\right)+o\left(n^{-\frac{2h+2}{2}}\right)\end{aligned}$$

as $\int_{\mathbb{R}^p}\left(1+\sum_{i=1}^{2h+2}\frac{\pi_i(z,\theta)}{n^{i/2}}\right)\varphi_{V(\theta)}\left(z\right)dz=1+o\left(n^{-\frac{2h+2}{2}}\right)$ due to the validity of the Edgeworth approximation of the distribution of $\sqrt{n}\left(\theta_n^{(h}-\theta\right)$ and the result follows. For the MSE approximation the result follows analogously, by simply noticing that $\left(\mathrm{Id}_p+\frac{1}{n^{\frac{2h+2}{2}}}\frac{\partial\mathcal{I}_V\left(k^{/}_{2h+2}\right)}{\partial\theta}\right)^{-1}=\mathrm{Id}_p+o\left(1\right)$. ■

5 Examples and Monte Carlo Experiments

In this Appendix we present an analytic proof the GARCH$(1,1)$ example only.

5.1 The GARCH(1,1) Case

Consider the set of stationary ergodic and covariance stationary processes defined by the recursion

$$\begin{aligned} y_j^2 &= \varepsilon_j^2 h_j \\ h_j &= \theta_1 (1-\theta_2-\theta_3) + \left(\theta_2 \varepsilon_{j-1}^2 + \theta_3\right) h_{j-1} \end{aligned}$$

where the (ε_j) are iid, $E\varepsilon_0 = 0$, $E\varepsilon_0^2 = 1$, $E\varepsilon_0^{28} < +\infty$ the distribution of ε_0 admits a positive continuous density and $\theta = (\theta_1, \theta_2, \theta_3)' \in \Theta = \left[\underline{\eta}_\omega, \overline{\eta}_\omega\right] \times \left[\underline{\eta}_\alpha, \overline{\eta}_\alpha\right] \times \left[\underline{\eta}_\beta, \overline{\eta}_\beta\right]$ where $\underline{\eta}_\omega, \underline{\eta}_\alpha, \underline{\eta}_\beta > 0$ and for any $\theta \in \Theta$, $E\left(\theta_2 \varepsilon_0^2 + \theta_3\right)^{14} < 1$.

Let

$$b(\theta) = \left(\theta_1, \frac{\theta_2 \left(1 - (\theta_2+\theta_3)\theta_3\right)}{1 - 2\theta_2\theta_3 - \theta_3^2}, \theta_2 + \theta_3\right)'$$

and for some compact $B \supseteq b(\Theta)$ and $c_n(\beta) = \left(\left(\overline{y^2}, \widehat{\rho}_1, \widehat{\frac{\rho_2}{\rho_1}}\right) - \beta\right)'$ define

$$\beta_n = \arg\min_{\beta \in B} \frac{1}{2} \left\| c_n(\beta) \right\|^2$$

where $\overline{y^2} = \frac{1}{n}\sum_{j=1}^n y_j^2$, $\widehat{\rho}_i = \frac{\frac{1}{n}\sum_{j=1}^n \left(y_t^2 y_{t-i}^2\right) - \left(\overline{y^2}\right)^2}{\frac{1}{n}\sum_{j=1}^n \left(y_t^4\right) - \left(\overline{y^2}\right)^2}$. Furthermore define

$$\mathrm{GMR1} = \arg\min_{\theta \in \Theta} \frac{1}{2} \left\| \beta_n - b(\theta) \right\|^2.$$

Now employing the GMR2 estimator, treating the GMR1 as an auxiliary one, we get the $1-\mathrm{GMR2}$ estimator. Again, the $E_\theta(\mathrm{GMR1})$ needs to be evaluated.

Proposition 4 *If the distribution of ε_0 admits a positive and continuous density then β_n and* GMR1 *admit 4^{th} order valid Edgeworth expansions, uniformly over Θ. Furthermore if the distribution of ε_0 is standard normal, then* GMR2, $1-$GMR2 *and* GT *admit 4^{th} order valid Edgeworth expansions, uniformly over any compact subset of Θ.*

Proof: For any $\theta \in \times$ let $X_j(\theta) = \begin{pmatrix} y_j^2 & y_j^4 & y_j^2 y_{j-1}^2 & y_j^2 y_{j-2}^2 \end{pmatrix}'$, and $S_n(\theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i(\theta) - EX_0(\theta))$. Then as $E(\theta_2 \varepsilon_0^2 + \theta_3)^{14} < 1$, the monotonicity of h w.r.t. θ and a dominated convergence argument imply that $E\left(y_j^m(\theta)\right)$ exists and is continuous on Θ for any $m = 1, \ldots, 24$. Therefore $\sup_{\theta \in \times} E \|X_0(\theta)\|^7 < +\infty$ establishing A.2-M in Arvanitis and Demos [1]. This also implies that if the formal Edgeworth expansion is valid, the polynomials of its density are equicontinuous functions of these moments and the covariance matrix is continuous on Θ and positive definite. The validity of the the formal Edgeworth expansion follows from the verification of conditions A.2-WD, A.3-CPD and A.3-NDD in Arvanitis and Demos [1] (for details see proposition 1 in Arvanitis and Demos [1]).

Let us define the function fact that $f(x) = \left(x_1, \frac{x_3 - x_1^2}{x_2 - x_1^2}, \frac{x_4 - x_1^2}{x_2 - x_1^2}\right)$ which is continuous. A 4^{th} order Taylor expansion of f-which is independent of θ- around $E(X_0(\theta))$ of gives

$$\sqrt{n}\left(\left(\overline{y^2}, \widehat{\rho}_1, \frac{\widehat{\rho}_2}{\widehat{\rho}_1}\right)' - b'(\theta)\right) = \sum_{i=0}^3 \frac{1}{n^{i/2}} D^{(i+1)} f(E(X_0(\theta)))(S_n(\theta))^{i+1} + R_n(\theta)$$

where

$$R_n(\theta) = \frac{1}{n^{3/2}}\left(D^4 f\left(R_n^+(\theta)\right)(S_n(\theta))^4 - D^4 f(E(X_0(\theta)))(S_n(\theta))^4\right)$$

$R_n^+(\theta)$ lies between $\frac{1}{n}\sum_{j=1}^n X_j(\theta)$ and $E(X_0(\theta))$ with probability $1 - o\left(n^{-\frac{3}{2}}\right)$ that does not depend on θ. Due to the continuity of $D^4 f$ on some compact neighborhood of $E(X_0(\theta))$ we have that

$$\|R_n(\theta)\| \leq \frac{\|R_n^+(\theta)\| \, \|S_n(\theta)\|^4}{n^{3/2}}$$

Hence the definition of $R_n^+(\theta)$, along with the fact that $S_n(\theta)$ has a valid Edgeworth expansion uniformly on Θ proposition, and lemmas AL.2 and 3.3 in Arvanitis and Demos [1] imply that the result will hold if

$$\sum_{i=0}^3 \frac{1}{n^{i/2}} D^{(i+1)} f(E(X_0(\theta)))(S_n(\theta))^{i+1}$$

admits the relevant Edgeworth expansion. But this holds due to the fact that $Df(E(X_0(\theta)))$ has rank 3 for any θ. Hence by theorem 3.1 in Arvanitis and Demos [1] it follow that $\sqrt{n}\left(\left(\overline{y^2}, \widehat{\rho}_1, \frac{\widehat{\rho_2}}{\widehat{\rho_1}}\right) - b(\theta)\right) - b'(\theta)$ admits a *locally uniform Edgeworth expansion of order* 4. As now $\beta_n = \left(\overline{y^2}, \widehat{\rho}_1, \frac{\widehat{\rho_2}}{\widehat{\rho_1}}\right)$ with probability $1-o\left(n^{-\frac{s-2}{2}}\right)$ that does not locally depend on θ, by lemma 3.3 in Arvanitis and Demos [1] we get $\sqrt{n}\left(\beta_n - b\left(\theta\right)\right)$ admits a *locally uniform Edgeworth expansion of order* 4 with Edgeworth polynomials that are, locally on Θ, equicontinuous functions.

Let us call GMR1 by θ_n. Initially observe that due to the first part, for some $\Theta^* = \left[\underline{\eta}^*_\omega, \overline{\eta}^*_\omega\right] \times \left[\underline{\eta}^*_\alpha, \overline{\eta}^*_\alpha\right] \times \left[\underline{\eta}^*_\beta, \overline{\eta}^*_\beta\right]$ where $0 < \underline{\eta}^*_m < \underline{\eta}_m, \overline{\eta}^*_m > \overline{\eta}_m$ for $m = \omega, \alpha, \beta$, such that $\mathrm{Int}\left(\Theta\right) \supset \Theta^* \supset \Theta'$

$$\sup_{\theta \in \overline{\mathcal{O}}(\theta_0, \delta)} P\left(\beta_n\left(\theta\right) \in \overline{\mathcal{O}}\left(\theta_0, \delta^*\right)\right) = 1 - o\left(n^{-\frac{s-2}{2}}\right)$$

and it is easy to see that $\frac{\partial b}{\partial \theta'}$ has full rank for any θ in $\overline{\mathcal{O}}\left(\theta_0, \delta^*\right)$, hence with probability $1-o\left(n^{-\frac{3}{2}}\right)$ that does not locally depend on θ, θ_n satisfies $\beta_n = b\left(\theta_n\right)$. The mean value theorem along with the constant full rank and continuity of $\frac{\partial b}{\partial \theta'}$ on Θ' imply that for some $c > 0$ independent of θ

$$\sup_{\theta \in \overline{\mathcal{O}}(\theta_0, \delta)} P\left(\sqrt{n}\left\|\theta_n - \theta\right\| \leq c\sqrt{n}\left\|\beta_n - b\left(\theta\right)\right\|\right) = 1 - o\left(n^{-\frac{3}{2}}\right)$$

which along with lemma AL.2 in Arvanitis and Demos [1] imply that for some $C^* > 0$ independent of θ

$$\sup_{\theta \in \overline{\mathcal{O}}(\theta_0, \delta)} P\left(\sqrt{n}\left\|\theta_n - \theta\right\| > C^* \ln^{1/2} n\right) = o\left(n^{-\frac{s-2}{2}}\right)$$

A Taylor expansion of $b\left(\theta_n\right)$ around $b\left(\theta\right)$ of order 4 implies that

$$0_{3\times 1} = \sqrt{n}\left(\beta_n - b\left(\theta\right)\right) + \sqrt{n}\sum\nolimits_{i=0}^{3} \frac{1}{n^{i/2}} D^{(i+1)} b\left(\theta\right)\left(\sqrt{n}\left(\theta_n - \theta\right)\right)^{i+1} + R_n\left(\theta\right)$$

where

$$R_n\left(\theta\right) = \frac{1}{n^{3/2}}\left(D^4 b\left(\theta_n^+\right)\left(\sqrt{n}\left(\theta_n - \theta\right)\right)^4 - D^4 b\left(\theta\right)\left(\sqrt{n}\left(\theta_n - \theta\right)\right)^4\right)$$

θ_n^+ lies between θ_n and θ with probability $1-o\left(n^{-\frac{3}{2}}\right)$ that does not depend on θ. Due to the continuity of $D^4 b\left(\theta\right)$ on some compact neighborhood of θ we have that

$$\left\|R_n\left(\theta\right)\right\| \leq \frac{\left\|\theta_n^+ - \theta\right\| \left\|\sqrt{n}\left(\theta_n - \theta\right)\right\|^4}{n^{3/2}}$$

Hence due to the definition of θ_n^+, the fact that θ_n is uniformly tight, the uniform expansion of β_n and the constant full rank of the Jacobian of b and application of theorem 3.2 in Arvanitis and Demos [1] delivers the result for θ_n.

Let us now call GMR2 as θ_n^*. Notice first that uniform consistency of β_n to $b(\theta)$ along with the boundeness of Θ imply by uniform integrability that

$$\sup_{\theta\in\Theta} |E_\theta \beta_n - b(\theta)| = o(1) \tag{7}$$

hence for any $\varepsilon > 0$

$$\begin{aligned} &\sup_{\theta^*\in\Theta} P\left(\sup_{\theta\in\Theta} ||\beta_n - E\beta_n(\theta)| - |b(\theta^*) - b(\theta)|| > \varepsilon\right) \\ \leq\ &\sup_{\theta^*\in\Theta} P\left(|\beta_n - b(\theta^*)| + o(1) > \varepsilon\right) = o\left(n^{-\frac{3}{2}}\right) \end{aligned}$$

due to the analogous consistency of β_n. Hence

$$\sup_{\theta^*\in\Theta} P\left(\theta_n^* \in \mathcal{O}(\theta^*, \varepsilon) \cap \Theta\right) = 1 - o\left(n^{-\frac{3}{2}}\right)$$

for any $\varepsilon > 0$. Then from lemma AL.9 and lemma 2.5 we obtain that

$$\sup_{\theta^*\in\Theta''} P\left(\sqrt{n}\,|\theta_n^* - \theta| > C \ln^{1/2} n\right) = o\left(n^{-\frac{3}{2}}\right) \tag{8}$$

for some appropriate $C > 0$. Now by recursive examination it is easy to see that $Eh_0^m(\theta)$ is 4 times continuously differentiable for any θ in Θ'' for all $m = 1, \ldots, 5$. This along the analogous differentiability of f imply that the π_i there are also 4 times continuously differentiable for any θ in Θ'' for any $z \in \mathbb{R}$. Then dominated convergence implies the same for $\mathcal{I}_V(k_i(z, \theta))$ for all $i = 1, \ldots, 3$. Then lemma 2.4 along with lemma AL.9 imply that $\frac{\partial E_{\theta_n^*}(\beta_n)}{\partial\theta}$ converges to $\frac{\partial b(\theta)}{\partial\theta'}$ for any θ in Θ'' with probability $1 - o\left(n^{-\frac{3}{2}}\right)$ independent of θ, hence with the same probability θ_n^* satisfies $\beta_n = E_{\theta_n^*}\theta_n$. Hence with probability $1 - o\left(n^{-\frac{3}{2}}\right)$ independent of θ, θ_n^* satisfies

$$0 = \sqrt{n}\left(\beta_n - E_\theta \theta_n^*\right) + A_n(\theta) + R_n(\theta)$$

where $\sup_{\theta\in\Theta''} P\left(\|R_n(\theta)\| > o\left(n^{-1}\right)\right) = o\left(n^{-3/2}\right)$. The result follows from 8, proposition AL.8, lemma AL.2 and theorem 3.2 in Arvanitis and Demos [1]. Notice that by the definition of $c_n(\beta)$ we have that $E_\theta(c_n(\beta_n)) = E_\theta\beta_n - \beta_n$, i.e. GT = GMR2.

Finally, the case of $1-$GMR2 follows in complete analogy to the previous by simply replacing in the previous proof any invocation to f with $b^{-1}(\varphi) = \left(\varphi_1, \frac{1-\varphi_3^2-\sqrt{\left(1-(2\varphi_2-\varphi_3)^2\right)\left(1-\varphi_3^2\right)}}{2(\varphi_2-\varphi_3)}, \frac{-\left(1-2\varphi_2\varphi_3+\varphi_3^2\right)+\sqrt{\left(1-(2\varphi_2-\varphi_3)^2\right)\left(1-\varphi_3^2\right)}}{2(\varphi_2-\varphi_3)}\right)$ and of b with the identity. ■

References

[1] Arvanitis S. and A. Demos (2012) "Valid Locally Uniform Edgeworth Expansions Under Weak Dependence and Sequences of Smooth Transformations", working paper No. 1214, Dept. of IEES, AUEB (http://wpa.deos.aueb.gr/docs/Uniform-Edg.pdf).

[2] Corradi V. and E.M. Iglesias (2008), "Bootstrap refinements for QML estimators of the GARCH(1,1) parameters", *Journal of Econometrics*, 144, pp. 500–510.

[3] Gotze F., and C. Hipp (1994), "Asymptotic distribution of statistics in time series", *The Annals of Statistics*, 22, pp. 2062-2088.

[4] Gradshteyn I. S., and I.M. Ryzhik (1994), "Table of Integrals, Series, and Products" A. Jeffrey Editor, Academic Press.

[5] Magdalinos, M.A. (1992), "Stochastic Expansions and Asymptotic Approximations", *Econometric Theory*, Vol. 8, No. 3, pp. 343-367.

[6] Skovgaard I.M. (1981), "Transformation of an Edgeworth expansion by a sequence of smooth functions", *Scandinavian Journal* of Statistics 8, 207-217.

[7] van der Vaart, A.W. and J.A. Wellner, (2000) "*Weak convergence and empirical processes*" Springer.

Appendix-General Proofs

The following are a collection of helpful results that are frequently used in the proofs of the main results.

Lemma AL.1 *Suppose that:*
-UUC

$$\sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left|c_n(\beta)-c(\theta,\beta)\right|>\varepsilon\right)=o\left(n^{-a}\right), \quad \forall\varepsilon>0$$

-AB $c(\theta,\beta)$ *is jointly continuous and* $\gamma(\theta)=\arg\min_{\beta\in B} c(\theta,\beta)$, *then*

$$\sup_{\theta\in\Theta} P_\theta\left(\left\|\beta_n-\gamma(\theta)\right\|>\varepsilon\right)=o\left(n^{-a}\right), \quad \forall\varepsilon>0$$

where $\beta_n\in\arg\min_{\beta\in B} c_n(\beta)$.

Proof. For $\varepsilon>0$ independent of θ, and for any β for which

$$\left\|\beta-\gamma(\theta)\right\|>\varepsilon$$

there must exist a $\delta>0$ such that

$$c(\theta,\beta)-c(\theta,\gamma(\theta))>\delta$$

due to the compactness of B the continuity of $c(\theta,\cdot)$ and the uniqueness of $b(\theta)$ as a minimizer of $c(\theta,\beta)$ for any θ. The compactness of $\Theta\times B$ and the *joint continuity* of c implies that it can be chosen independent of θ. Suppose that this is *not* the case which implies that $\inf_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\delta=0$. Then there exists a sequence θ_m in Θ for which, for any $\varepsilon>0$ there exists an $m(\varepsilon)$ such that $c(\theta_m,\beta)-c(\theta_m,\gamma(\theta))<\varepsilon$ for all $m\geq m(\varepsilon)$. Due to compactness θ_m can be chosen convergent, say to θ_*. Then due to the joint continuity of c and the continuity of b we have that $c(\theta_*,\beta_n)-c(\theta_*,\gamma(\theta_*))=0$ which is impossible if $\beta\neq\gamma(\theta_*)$ due to the property of γ. Hence

$$\begin{aligned}
&\sup_{\theta\in\Theta} P_\theta\left(\left\|\beta_n-\gamma(\theta)\right\|>\varepsilon\right)\\
\leq\;&\sup_{\theta\in\Theta} P_\theta\left(\left|c(\theta,\beta_n)-c(\theta,\gamma(\theta))\right|>\delta\right)\\
\leq\;&\sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left|c_n(\beta)-c(\theta,\beta)\right|>\frac{\delta}{2}\right)=o\left(n^{-a^*}\right)
\end{aligned}$$

which implies the result. ■

Lemma AL.2 *Let assumptions A.7.a) and A.6 hold. Then for* $j = *, **$

$$\sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta\right)\right\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0$$

Furthermore, there exists $K > 0$ *for which*

$$\sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right)\right\| > K\right) = o\left(n^{-a^*}\right)$$

Proof. Assumptions A.6, A.7a) and the triangle inequality imply that for any $\varepsilon > 0$

$$\begin{aligned}
& \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta\right)\right\| > \varepsilon\right) \qquad (9)\\
\leq\ & \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta_n^*\right)\right\| > \frac{\varepsilon}{2}\right) + \sup_{\theta\in\Theta} P_\theta\left(\left\|E_\theta W^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta\right)\right\| > \frac{\varepsilon}{2}\right)\\
\leq\ & o\left(n^{-a^*}\right) + \sup_{\theta\in\Theta} P_\theta\left(\left\|E_\theta W^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta\right)\right\| > \frac{\varepsilon}{2}\right) \quad \text{by assumption A.7.a)}\\
\leq\ & o\left(n^{-a^*}\right) + \sup_{\theta\in\Theta} P_\theta\left(\kappa^*\left(\theta\right)\left\|\theta_n^* - \theta\right\| > \frac{\varepsilon}{2}\right) \quad \text{by assumption A.7.a)}\\
=\ & o\left(n^{-a^*}\right) \quad \text{by assumption A.6}
\end{aligned}$$

due to the fact that $\sup_{\theta\in\Theta}\kappa^j\left(\theta\right) < +\infty$. Now for $K > \sup_{\theta\in\Theta}\left\|E_\theta W^j\left(\theta\right)\right\| > 0$ which exists due to assumption A.7.a) and $\varepsilon = K - \sup_{\theta\in\Theta}\left\|E_\theta W^j\left(\theta\right)\right\|$we have that

$$\begin{aligned}
\sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right)\right\| > K\right) &= \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right)\right\| > \varepsilon + \left\|E_\theta W^j\left(\theta\right)\right\|\right)\\
&= \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right)\right\| - \left\|E_\theta W^j\left(\theta\right)\right\| > \varepsilon\right)\\
&\leq \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j\left(\theta_n^*\right) - E_\theta W^j\left(\theta\right)\right\| > \varepsilon\right)\\
&= o\left(n^{-a^*}\right).
\end{aligned}$$

■

Lemma AL.3 *Suppose that*

$$c_n\left(\beta\right) = \sqrt{q_n'\left(\beta\right) W_n^j\left(\theta_n^*\right) q_n\left(\beta\right)}$$

for some appropriate random element q_n *where* W_n^j*,* θ_n^* *satisfy assumptions A.7.a), A.6 and for* q *an appropriate jointly continuous function on* $\Theta\times B$

$$\sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left\|q_n\left(\beta\right) - q\left(\theta,\beta\right)\right\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0$$

*Then AL.1.*UUC *holds for* $c\left(\theta,\beta\right) = \sqrt{q'\left(\theta,\beta\right) E_\theta W^j\left(\theta\right) q\left(\theta,\beta\right)}$ *which is jointly continuous.*

Proof. Due to the triangle inequality the submultiplicativity and the monotonicity of the square root, we have pointwise that

$$\begin{aligned}
& \left|c_n(\beta) - c(\theta,\beta)\right| \\
\leq\ & \left|c_n(\beta) \pm \sqrt{q'(\theta,\beta) W_n^j(\theta_n^*) q(\theta,\beta)} - c(\theta,\beta)\right| \\
\leq\ & \left\|q_n'(\beta) - q(\theta,\beta)\right\|_{W_n^j(\theta_n^*)} + \sqrt{\left|q'(\theta,\beta)\left(W_n^j(\theta_n^*) - E_\theta W^j(\theta)\right) q(\theta,\beta)\right|} \\
\leq\ & \left\|q_n'(\beta) - q(\theta,\beta)\right\|_{W_n^j(\theta_n^*)} + \left\|q'(\theta,\beta)\right\| \sqrt{\left\|W_n^j(\theta_n^*) - E_\theta W^j(\theta)\right\|} \\
\leq\ & \left\|q_n'(\beta) - q(\theta,\beta)\right\| \sqrt{\left\|W_n^j(\theta_n^*)\right\|} + \left\|q(\theta,\beta)\right\| \sqrt{\left\|W_n^j(\theta_n^*) - E_\theta W^j(\theta)\right\|}
\end{aligned}$$

therefore

$$\begin{aligned}
& \sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left\|c_n(\beta) - c(\theta,\beta)\right\| > \varepsilon\right) \\
\leq\ & \sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left\|q_n'(\beta) - q(\theta,\beta)\right\| \sqrt{\left\|W_n^j(\theta_n^*)\right\|} > \frac{\varepsilon}{2}\right) \\
& + \sup_{\theta\in\Theta} P_\theta\left(\sup_{(\theta,\beta)\in\Theta\times B}\left\|q(\theta,\beta)\right\| \sqrt{\left\|W_n^j(\theta_n^*) - E_\theta W^j(\theta)\right\|} > \frac{\varepsilon}{2}\right)
\end{aligned}$$

Now continuity of q and compactness of $\Theta\times B$ imply that $\sup_{(\theta,\beta)\in\Theta\times B}\left\|q(\theta,\beta)\right\| < M$. Furthermore, for $c = \sqrt{K}$ and K as in lemma AL.2, that applies due to assumptions A.7.a), A.6 we have that the right hand side of the previous inequality is bounded by

$$\begin{aligned}
& \sup_{\theta\in\Theta} P_\theta\left(\sup_{\beta\in B}\left\|q_n'(\beta) - q(\theta,\beta)\right\| > \frac{\varepsilon}{2c}\right) \\
& + \sup_{\theta\in\Theta} P_\theta\left(\left\|W_n^j(\theta_n^*) - E_\theta W^j(\theta)\right\| > \sqrt{\frac{\varepsilon}{2M}}\right)
\end{aligned}$$

and AL.1.UUC follows due to the hypotheses and lemma AL.2. The joint continuity follows from the hypothesis for q and the the fact that $E_\theta W^j(\theta)$ is continuous due to A.7.a). ■

Lemma AL.4 *Suppose* W_n^j, θ_n^* *satisfy assumptions A.7, A.6,* β_n *and* $\gamma(\theta)$ *are as in lemma AL.1,* γ *is continuous on* $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ *and that:*
-FOC β_n *satisfies*

$$\frac{\partial q_n'(\beta_n)}{\partial\beta} W_n^j(\theta_n^*) q_n(\beta_n) = \mathbf{0}$$

with P_θ*-probability* $1 - o\left(n^{-a^*}\right)$ *that is independent of* θ*,*
-HUB *for some* $\delta, M > 0$ *independent of* θ *such that* $\gamma\left(\overline{\mathcal{O}}_\varepsilon(\theta_0)\right) \subset \overline{\mathcal{O}}_\delta(\gamma(\theta_0))$

and for all i, $\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sup_{\beta\in\overline{\mathcal{O}}_\delta(\gamma(\theta_0))}\left\|\frac{\partial^2 q_n'(\beta_n)}{\partial\beta\partial\beta_i}\right\| > M\right) = o\left(n^{-a^*}\right)$,
-RANK *for any* $\beta \in \overline{\mathcal{O}}_\delta(\gamma(\theta_0))$, $\frac{\partial q_n(\beta)}{\partial\beta'}$ *is of full rank with* P_θ*-probability* $1-o\left(n^{-a^*}\right)$ *that is independent of* θ *and,*
-TIGHT *for some* $C>0$ *independent of* θ, $\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|q_n(\gamma(\theta))\right\| > C\ln^{1/2} n\right) = o\left(n^{-a^*}\right)$, *then*

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\beta_n-\gamma(\theta)\right\| > C^+\ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

for some $C^+>0$ *independent of* θ.

Proof. Due to AL.4.HUB-RANK, A.7 and the mean value theorem we have that with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ

$$\frac{\partial q_n'(b(\theta))}{\partial\beta}W_n^j(\theta_n^*)\sqrt{n}q_n(\gamma(\theta)) + A_n\sqrt{n}\left\|\beta_n-\gamma(\theta)\right\| = \mathbf{0}$$

with

$$A_n = \left[\frac{\partial^2 q_n'\left(\beta_n^+\right)}{\partial\beta\partial\beta_i}W_n^j(\theta_n^*)\,q_n\left(\beta_n^+\right)\right]_i + \frac{\partial q_n'\left(\beta_n^+\right)}{\partial\beta}W_n^j(\theta_n^*)\frac{\partial q_n\left(\beta_n^+\right)}{\partial\beta'}$$

where β_n^+ lies between β_n and $\gamma(\theta)$. We have that due to submultiplicativity

$$\begin{aligned}&\left\|\frac{\partial q_n'(\gamma(\theta))}{\partial\beta}W_n^j(\theta_n^*)\sqrt{n}q_n(\gamma(\theta))\right\|\\ \leq\ &\left\|\frac{\partial q_n'(\gamma(\theta))}{\partial\beta}\right\|\left\|W_n^j(\theta_n^*)\right\|\sqrt{n}\left\|q_n(\gamma(\theta))\right\|\end{aligned}$$

and due to AL.4.HUB we have that $\frac{\partial q_n'(\gamma(\theta))}{\partial\beta}$ is asymptotically equi-Lipschitz and therefore there exists some constant $m^*>0$, independent of θ for which

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)}\left\|\frac{\partial q_n'(\gamma(\theta))}{\partial\beta}\right\| > m^*\right) = o\left(n^{-a^*}\right)$$

furthermore assumptions A.7, A.6 along with lemma AL.2 imply that there exists $K>0$ independent of θ for which

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|W_n^j(\theta_n^*)\right\| > K\right) = o\left(n^{-a^*}\right)$$

hence due to AL.4.TIGHT

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|\frac{\partial q_n'(b(\theta))}{\partial\beta}W_n^j(\theta_n^*)\sqrt{n}q_n(\gamma(\theta))\right\| > C^*\ln^{1/2}n\right) = o\left(n^{-a^*}\right)$$

for any $C^* \geq \frac{C}{m^*K}$ which is obviously independent of θ. Furthermore, due to AL.4.HUB and the mean value theorem with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ

$$q_n\left(\beta_n^+\right) = q_n(\gamma(\theta)) + \frac{\partial q_n'\left(\beta_n^{++}\right)}{\partial\beta}\left(\beta_n^+ - b(\theta)\right)$$

where β_n^{++} lies between β_n^+ and $\gamma(\theta)$. As before due to the definitions of β_n^+, β_n^{++} and due to AL.4.TIGHT

$$\begin{aligned}\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|\frac{\partial q_n'\left(\beta_n^{++}\right)}{\partial\beta}\right\| > m^*\right) &= o\left(n^{-a^*}\right), \\ \sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|\beta_n^+ - \gamma(\theta)\right\| > \varepsilon\right) &= o\left(n^{-a^*}\right) \text{ for any } \varepsilon > 0\end{aligned}$$

and due to AL.4.TIGHT

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|q_n\left(\beta_n^+\right)\right\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0$$

which furthermore along with AL.4.HUB and lemma AL.2 imply that

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|\left[\frac{\partial^2 q_n'\left(\beta_n^+\right)}{\partial\beta\partial\beta_i}W_n^j(\theta_n^*)\,q_n\left(\beta_n^+\right)\right]_i\right\| > \varepsilon\right) = o\left(n^{-a^*}\right), \forall\varepsilon > 0$$

Also, AL.4.RANK via the Weierstrass theorem which implies that with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ, $\inf_{\beta\in\overline{\mathcal{O}}_{\delta}(b(\theta))}\operatorname{rank}\frac{\partial q_n'(\beta)}{\partial\beta}$ is full. A.7 imply that, with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\lambda_n^{\min} < k\right) = o\left(n^{-a^*}\right)$ for some $k > 0$ independent of θ, where $\lambda_n^{\min}$ denotes the smallest absolute eigenvalue of $W_n^j(\theta_n^*)$. These imply that with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\mu_n^{\min} < k^*\right) = o\left(n^{-a^*}\right)$ for some $k^* > 0$ independent of θ, where $\mu_n^{\min}$ denotes the smallest absolute eigenvalue of $\left(\frac{\partial q_n'\left(\beta_n^+\right)}{\partial\beta}W_n^j(\theta_n^*)\frac{\partial q_n\left(\beta_n^+\right)}{\partial\beta'}\right)$. Hence with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ, A_n^{-1} exists and is of the form $\left(\frac{\partial q_n'\left(\beta_n^+\right)}{\partial\beta}W_n^j(\theta_n^*)\frac{\partial q_n\left(\beta_n^+\right)}{\partial\beta'}\right)^{-1} + B_n$, with $\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\|B_n\| > \varepsilon\right) =$

$o\left(n^{-a^*}\right)$ for any $\varepsilon > 0$. Furthermore due to the fact that $\left(\frac{\partial q_n'\left(\beta_n^+\right)}{\partial \beta} W_n^j\left(\theta_n^*\right) \frac{\partial q_n\left(\beta_n^+\right)}{\partial \beta'}\right)^{-1}$ is symmetric we have that

$$\left\| \left(\frac{\partial q_n'\left(\beta_n^+\right)}{\partial \beta} W_n^j\left(\theta_n^*\right) \frac{\partial q_n\left(\beta_n^+\right)}{\partial \beta'}\right)^{-1} \right\| \leq \frac{r}{\left(\mu_n^{\min}\right)^2}$$

where r is the rank of the matrix. Hence for an $\varepsilon > 0$

$$\begin{aligned} & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|A_n^{-1}\right\| > c\right) \\ \leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{r}{\left(\mu_n^{\min}\right)^2} + \varepsilon > c\right) + o\left(n^{-a^*}\right) \\ \leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{r}{\left(k^*\right)^2} + \varepsilon > c\right) + o\left(n^{-a^*}\right) = o\left(n^{-a^*}\right) \end{aligned}$$

for any $c \geq \frac{r}{(k^*)^2} + \varepsilon$. These imply that

$$\begin{aligned} & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\beta_n - \gamma(\theta)\right\| > C^+ \ln^{1/2} n\right) \\ \leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|A_n^{-1}\right\| \left\|\frac{\partial q_n'\left(b(\theta)\right)}{\partial \beta} W_n^j\left(\theta_n^*\right) \sqrt{n} q_n\left(\gamma(\theta)\right)\right\| > C^+ \ln^{1/2} n\right) + o\left(n^{-a^*}\right) \\ \leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|\frac{\partial q_n'\left(b(\theta)\right)}{\partial \beta} W_n^j\left(\theta_n^*\right) \sqrt{n} q_n\left(\gamma(\theta)\right)\right\| > \frac{C^+}{c} \ln^{1/2} n\right) + o\left(n^{-a^*}\right) \end{aligned}$$

which is $o\left(n^{-a^*}\right)$ for any $C^+ \geq cC^*$. ■

Lemma AL.5 *Suppose that:*
-FOC β_n *satisfies*

$$Q_n\left(\beta_n\right) W_n^j\left(\theta_n^*\right) q_n\left(\beta_n\right) = \mathbf{0}$$

with P_θ-probability $1 - o\left(n^{-a^*}\right)$ that is independent of θ,
-UTIGHT *There exists a $C^+ > 0$ independent of θ for which*

$$\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\sqrt{n}\left\|\beta_n - \gamma(\theta)\right\| > C^+ \ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

-UEDGE ***a.*** *There exists a random element $M_n(\theta)$ with values in an Euclidean space, containing the elements of $\sqrt{n}\left(\theta_n^* - \theta\right)$, the distribution of which admits a uniform over $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ Edgeworth expansion $\Psi_{n,s}(\theta)$. The i^{th} polynomial, say, $\pi_i(z,\theta)$ of $\Psi_{n,s}(\theta)$ is equicontinuous on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ $\forall z \in \mathbb{R}^q$, for*

$i=1,\ldots,s-2$, and if $\Sigma(\theta)$ denotes the variance matrix in the density of $\Psi_{n,s}(\theta)$ then it is continuous on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ and positive definite.
-EXPAND *The following hold with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ*

$$Q_n(\beta_n)=\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}\sum\nolimits_{j=0}^{i}C^*_{ij_n}(\theta)\left(M_n(\theta)^j,S_n(\theta)^{i-j}\right)+R^*_n(\beta_n,\theta)$$

$$W_n^j(\theta_n^*)=\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}C^{**}_{i_n}(\theta)\left(M_n(\theta)^i\right)+R^{**}_n(\theta^*_n,\theta)$$

$$\sqrt{n}q_n(\beta_n)=\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}\sum\nolimits_{j=0}^{i+1}C^{\#}_{ij_n}(\theta)\left(M_n(\theta)^j,S_n(\theta)^{i+1-j}\right)+R^{\#}_n\left(\widetilde{\beta}_n,\theta\right)$$

where $S_n(\theta)=\sqrt{n}(\beta_n-\gamma(\theta))$, $C^*_{ij_n}:\overline{\mathcal{O}}_{\varepsilon}(\theta_0)\times\mathbb{R}^{q^i}\to\mathbb{R}^p$, $C^{**}_{i_n}:\overline{\mathcal{O}}_{\varepsilon}(\theta_0)\times\mathbb{R}^{q^i}\to\mathbb{R}^p$ are i-linear, $C_{ij_n}:\Theta\times\mathbb{R}^{q^{i+1}}\to\mathbb{R}^p$ is $(i+1)$-linear $\forall\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, $C^*_{00_n}$, $C^{**}_{0_n}$, $C^{\#}_{00_n}(\theta),C^{\#}_{01_n}(\theta)$ are independent of n and have full rank $\forall\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, $C^*_{i_n}$, $C^{**}_{i_n}$, $C^{\#}_{ij_n}$ are equicontinuous on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$, and

$$\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon}(\theta_0)}P_\theta\left(\left\|R^l_n\right\|>\gamma^l_n\right)=o\left(n^{-a^*}\right),\ l=*,**,\#$$

*for real sequence $\gamma^l_n=o\left(n^{-a^*}\right)$ independent of θ, for , $l=*,**,\#$.
Then $\sqrt{n}(\beta_n-\gamma(\theta))$ admits a locally uniform Edgeworth expansion, $\Psi^*_{n,s}(\theta)$, over $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$. The i^{th} polynomial, say, $\pi^*_i(z,\theta)$ of the density of $\Psi^*_{n,s}(\theta)$ is equicontinuous on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ $\forall z\in\mathbb{R}^q$, for $i=1,\ldots,s-2$, and if $\Sigma^*(\theta)$ denotes the variance matrix in the density of $\Psi^*_{n,s}(\theta)$ then it is continuous on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ and positive definite.*

Proof. Due to conditions UTIGHT and EXPAND condition FOC implies that with P_θ-probability $1-o\left(n^{-a^*}\right)$ that is independent of θ

$$\begin{aligned}&\left(\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}\sum\nolimits_{j=0}^{i}C^*_{ij_n}(\theta)\left(M_n(\theta)^j,\left(\sqrt{n}(\beta_n-\gamma(\theta))\right)^{i-j}\right)+R^*_n(\beta_n,\theta)\right)\times\\&\left(\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}C^{**}_{i_n}(\theta)\left(M_n(\theta)^i\right)+R^{**}_n(\theta^*_n,\theta)\right)\times\\&\left(\sum_{i=0}^{s^*-1}\frac{1}{n^{i/2}}\sum\nolimits_{j=0}^{i+1}C^{\#}_{ij_n}(\theta)\left(M_n(\theta)^j,\left(\sqrt{n}(\beta_n-\gamma(\theta))\right)^{i+1-j}\right)+R^{\#}_n(\beta_n,\theta)\right)\\=&\ \mathbf{0}\end{aligned}$$

Gathering terms of the same order we obtain that with P_θ-probability $1 - o\left(n^{-a^*}\right)$ that is independent of θ

$$\sum_{i=0}^{s^*-1} \frac{1}{n^{i/2}} \sum\nolimits_{j=0}^{i+1} C_{ij_n}(\theta)\left(M_n(\theta)^j, \left(\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right)^{i+1-j}\right) + R_n\left(\beta_n, \theta_n^*, \theta\right) = \mathbf{0}$$

with $C_{00_n}(\theta) = C^*_{00_n}(\theta)\, C^{**}_{0_n}(\theta)\, C^{\#}_{00_n}(\theta)$, $C_{01_n}(\theta) = C^*_{00_n}(\theta)\, C^{**}_{00_n}(\theta)\, C^{\#}_{01_n}(\theta)$ which are obviously independent of n and of full rank,

$$C_{ij_n}(\theta) = \sum_{j_0+j_1+j_2=j, i_0+i_1+j_1=i} C^*_{i_0j_{0n}}(\theta)\, C^{**}_{j_{1n}}(\theta)\, C^{\#}_{i_1j_{2n}}(\theta)$$

which is obviously equicontinuous on $\overline{\mathcal{O}}_\varepsilon(\theta_0)$. Moreover $R_n\left(\beta_n, \theta_n^*, \theta\right)$ *is a sum* containing terms of the form

$$\begin{aligned} A_n \;=\; & \frac{1}{n^{(i_0+i_1+j_1)/2}} C^*_{i_0j_{0n}}(\theta)\left(M_n(\theta)^{j_0}, \left(\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right)^{i_0-j_0}\right) C^{**}_{j_{1n}}(\theta)\left(M_n(\theta)^{j_1}\right) \times \\ & C^{\#}_{i_1j_{2n}}(\theta)\left(M_n(\theta)^{j_2}, \left(\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right)^{i_1+1-j_2}\right) \end{aligned}$$

for which $i_1 + i_2 + j_1 > s^* - 1$. Due to equicontinuity which along with the compactness of $\overline{\mathcal{O}}_\varepsilon(\theta_0)$ imply that the C functions and thereby their products have uniformly bounded coefficients, and the submultiplicativity we have that for some $M > 0$ independent of θ

$$\|A_n\| \leq \frac{M}{n^{(i_0+i_1+j_1)/2}} \left\|\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right\|^{i_0+i_1+1-j_1-j_2} \|M_n(\theta)\|^{j_1+j_2}$$

hence due to UTIGHT, UEDGE which along with lemma AL.2 of Arvanitis and Demos [1] imply the existence of a constant $C > 0$ independent of θ for which

$$\begin{aligned} & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\|A_n\| > \gamma_n\right) \\ \leq\; & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{n^{(i_0+i_1+j_1)/2}} \left\|\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right\|^{i_0+i_1+1-j_1-j_2} \|M_n(\theta)\|^{j_1+j_2} > \gamma_n\right) \\ \leq\; & o\left(n^{-a^*}\right) + \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M\left(C^+\right)^{i_0+i_1+1-j_1-j_2} C^{j_1+j_2}}{n^{(i_0+i_1+j_1)/2}} \ln^{\frac{i_1+i_2+1}{2}} n > \gamma_n\right) \end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \frac{M\left(C^+\right)^{i_0+i_1+1-j_1-j_2} C^{j_1+j_2}}{n^{(i_0+i_1+j_1)/2}} \ln^{\frac{i_1+i_2+1}{2}} n = o\left(n^{-a^*}\right)$ and is obviously independent of θ. Furthermore $R_n\left(\beta_n, \theta\right)$ contains terms of the form

$$B_n = \frac{R_n^*\left(\beta_n, \theta\right)}{n^{(i_1+i_2)/2}} C^{**}_{i_{1n}}(\theta)\left(M_n(\theta)^{i_1}\right) C^{\#}_{i_2j_n}(\theta)\left(M_n(\theta)^j, \left(\sqrt{n}\left(\beta_n - \gamma(\theta)\right)\right)^{i_2+1-j}\right)$$

and of the form

$$\Gamma_n = \frac{R_n^*(\beta_n,\theta)}{n^{i_1/2}} C_{i_{1n}}^{**}(\theta)\left(M_n(\theta)^{i_1}\right) R_n^{\#}(\beta_n,\theta)$$

and of the form

$$\begin{aligned}\Delta_n &= \frac{1}{n^{(i_1+i_2)/2}} C_{i_1 j_0 n}^{*}(\theta)\left(M_n(\theta)^{j_0}, \left(\sqrt{n}\left(\beta_n-\gamma(\theta)\right)\right)^{i_1-j_0}\right) R_n^{**}(\theta_n^*,\theta) \\ &\quad \times C_{i_2 j_2 n}^{\#}(\theta)\left(M_n(\theta)^{j_2}, \left(\sqrt{n}\left(\beta_n-\gamma(\theta)\right)\right)^{i_2+1-j_2}\right)\end{aligned}$$

and of the form

$$E_n = \frac{1}{n^{i_1/2}} C_{i_1 j_0 n}^{*}(\theta)\left(M_n(\theta)^{j_0}, \left(\sqrt{n}\left(\beta_n-\gamma(\theta)\right)\right)^{i_1-j_0}\right) R_n^{**}(\theta_n^*,\theta) R_n^{\#}(\beta_n,\theta)$$

for any compatible i_0, i_1, i_2, j, j_1, j_2 and finally the term

$$Z_n = R_n^*(\beta_n,\theta) R_n^{**}(\theta_n^*,\theta) R_n^{\#}(\beta_n,\theta)$$

and using the same arguments as before along with condition EXPAND, we have that

$$\begin{aligned}&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\|B_n\|>\gamma_n\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{n^{(i_1+i_2)/2}}\left\|R_n^*(\beta_n,\theta)\right\|\left\|M_n(\theta)\right\|^{i_1}\left\|\sqrt{n}\left(\beta_n-\gamma(\theta)\right)\right\|^{i_2+1-j}>\gamma_n\right) \\ \leq\ & o\left(n^{-a^*}\right) + \sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M\left(C^+\right)^{i_2+1-j} C^{i_1}}{n^{(i_1+i_2)/2}} \ln^{\frac{i_1+i_2-j+1}{2}} n\gamma_n^* > \gamma_n\right)\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \frac{M\left(C^+\right)^{i_2+1-j} C^{i_1}}{n^{(i_1+i_2)/2}} \ln^{\frac{i_1+i_2-j+1}{2}} n\gamma_n^* = o\left(n^{-a^*}\right)$ and is obviously independent of θ, and

$$\begin{aligned}&\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\|\Gamma_n\|>\gamma_n\right) \\ \leq\ &\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{n^{i_1/2}}\left\|R_n^*(\beta_n,\theta)\right\|\left\|M_n(\theta)\right\|^{i_1}\left\|R_n^{\#}(\beta_n,\theta)\right\|>\gamma_n\right) \\ \leq\ & o\left(n^{-a^*}\right) + \sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{MC^{i_1}}{n^{i_1/2}} \ln^{\frac{i_1}{2}} n\gamma_n^*\gamma_n^{\#} > \gamma_n\right)\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \frac{MC^{i_1}}{n^{i_1/2}} \ln^{\frac{i_1}{2}} n\gamma_n^* \gamma_n^\# = o\left(n^{-a^*}\right)$ and is obviously independent of θ, and

$$\begin{aligned}
& \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|\Delta_n\right\| > \gamma_n\right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{n^{(i_1+i_2)/2}} \left\|\sqrt{n}\left(\beta_n - \gamma\left(\theta\right)\right)\right\|^{i_1+i_2+1-j_0-j_2} \left\|M_n\left(\theta\right)\right\|^{j_0+j_2} \left\|R_n^{**}\left(\theta_n^*, \theta\right)\right\| > \gamma_n\right) \\
\leq & \ o\left(n^{-a^*}\right) + \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{MC^{j_0+j_2}\left(C^+\right)^{i_1+i_2+1-j_0-j_2}}{n^{(i_1+i_2)/2}} \ln^{\frac{i_1+i_2+1}{2}} n\gamma_n^{**} > \gamma_n\right)
\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \frac{MC^{j_0+j_2}\left(C^+\right)^{i_1+i_2+1-j_0-j_2}}{n^{(i_1+i_2)/2}} \ln^{\frac{i_1+i_2+1}{2}} n\gamma_n^{**} = o\left(n^{-a^*}\right)$ and is obviously independent of θ and

$$\begin{aligned}
& \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|E_n\right\| > \gamma_n\right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M}{n^{i_1/2}} \left\|M_n\left(\theta\right)\right\|^{j_0} \left\|\sqrt{n}\left(\beta_n - \gamma\left(\theta\right)\right)\right\|^{i_1-j_0} \left\|R_n^{**}\left(\beta_n, \theta\right)\right\| \left\|R_n^\#\left(\beta_n, \theta\right)\right\| > \gamma_n\right) \\
\leq & \ o\left(n^{-a^*}\right) + \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{M\left(C^+\right)^{i_1-j_0} C^{j_0}}{n^{i_1/2}} \ln^{\frac{i_1}{2}} n\gamma_n^{**}\gamma_n^\# > \gamma_n\right)
\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \frac{M\left(C^+\right)^{i_1-j_0} C^{j_0}}{n^{i_1/2}} \ln^{\frac{i_1}{2}} n\gamma_n^{**}\gamma_n^\# = o\left(n^{-a^*}\right)$ and is obviously independent of θ and finally

$$\begin{aligned}
& \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|Z_n\right\| > \gamma_n\right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|R_n^{**}\left(\beta_n, \theta\right)\right\| \left\|R_n^{**}\left(\theta_n^*, \theta\right)\right\| \left\|R_n^\#\left(\beta_n, \theta\right)\right\| > \gamma_n\right) \\
\leq & \sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\gamma_n^*\gamma_n^{**}\gamma_n^\# > \gamma_n\right)
\end{aligned}$$

which is of order $o\left(n^{-a^*}\right)$ for $\gamma_n = \gamma_n^*\gamma_n^{**}\gamma_n^\# = o\left(n^{-a^*}\right)$ and is obviously independent of θ. Hence there exists a real sequence $\gamma_n = o\left(n^{-a^*}\right)$ independent of θ for which

$$\sup_{\theta \in \overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\left\|R_n\left(\beta_n, \theta_n^*, \theta\right)\right\| > \gamma_n\right) = o\left(n^{-a^*}\right)$$

The result follows then from theorem 3.2 of Arvanitis and Demos [1]. ■

Lemma AL.6 *Under assumptions A.7 and A.8 condition* EXPAND *hold for* $W_n^j\left(\theta_n^*\right)$ *where* $M_n\left(\theta\right) = \sqrt{n}m_n\left(\theta\right)$.

Proof. Due to assumption A.7.b) for any $\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)$ and any θ_* sufficiently close to θ, $W_n^j(\theta_*)$ admits a Taylor expansion of order s^*-1 around θ of the form

$$\begin{aligned} W_n^j(\theta_*) &= \sum_{i=0}^{s^*-1} \frac{1}{i!} D^i W_n^j(\theta)\left((\theta_* - \theta)^i\right) \\ &\quad + \frac{1}{(s^*-1)!}\left(D^{s^*-1}W_n^j(\theta^+) - D^{s^*-1}W_n^j(\theta)\right)\left((\theta_* - \theta)^{s^*-1}\right) \end{aligned}$$

where θ^+ lies between θ_* and θ. Due to the assumption A.8 the elements of $\sqrt{n}(\theta^* - \theta)$ are in $M_n(\theta)$. Furthermore there exist $K^i(\theta)$ i-linear functions such that the coefficients of $\sqrt{n}(D^i W_n^j(\theta) - K^i(\theta))$ are also in $M_n(\theta)$. Due to assumption A.7.b) the elements of $K^i(\theta)$ can be identified as the uniform probability limits of the corresponding elements of $D^i W_n^j(\theta)$ and thereby are continuous on $\overline{\mathcal{O}}_{\varepsilon}(\theta_0)$. Obviously $K^0(\theta) = E_\theta W_n^j(\theta)$ due to A.7.a). The previous along with lemma AL.2 of Arvanitis and Demos [1] imply the existence of a constant $C > 0$ independent of θ for which

$$\sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\sqrt{n}\left\|\theta_n^* - \theta\right\| > C \ln^{1/2} n\right) = o\left(n^{-a^*}\right)$$

hence we obtain that with probability $1 - o\left(n^{-a^*}\right)$ that is independent of θ

$$\begin{aligned} W_n^j(\theta_n^*) &= E_\theta W_n^j(\theta) + \sum_{i=1}^{s^*-1} \frac{1}{i!}\frac{1}{n^{i/2}} K^i(\theta)\left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^i\right) \\ &\quad + \sum_{i=1}^{s^*-1} \frac{1}{i!}\frac{1}{n^{i/2}}\sqrt{n}\left(D^{i-1}W_n^j(\theta) - K^{i-1}(\theta)\right)\left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^i\right) \\ &\quad + R_n^{**}(\theta_n^*, \theta) \end{aligned}$$

with

$$\begin{aligned} R_n^{**}(\theta_n^*, \theta) &= \frac{1}{(s^*-1)!}\frac{1}{n^{\frac{s^*-1}{2}}}\left(D^{s^*-1}W_n^j(\theta^+) - D^{s^*-1}W_n^j(\theta)\right)\left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^{s^*-1}\right) \\ &\quad + \frac{1}{(s^*-1)!}\frac{1}{n^{\frac{s^*}{2}}}\sqrt{n}\left(D^{s^*-1}W_n^j(\theta) - K^{s^*-1}(\theta)\right)\left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^{s^*-1}\right) \end{aligned}$$

Furthermore due to submultiplicativity, A.7.b), A.8 and lemma AL.2 of Arvanitis and Demos [1] there exist $M, C > 0$ and independent of θ such that

$$\begin{aligned} &\sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\|R_n^{**}(\theta_n^*, \theta)\right\| > \gamma_n^{**}\right) \\ \leq &\sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\| \begin{array}{c} \frac{1}{(s^*-1)!}\frac{1}{n^{\frac{s^*-1}{2}}}\left(D^{s^*-1}W_n^j(\theta^+) - D^{s^*-1}W_n^j(\theta)\right) \times \\ \left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^{s^*-1}\right) \end{array} \right\| > \frac{\gamma_n^{**}}{2}\right) \\ &+ \sup_{\theta \in \overline{\mathcal{O}}_{\varepsilon}(\theta_0)} P_\theta\left(\left\| \begin{array}{c} \frac{1}{(s^*-1)!}\frac{1}{n^{\frac{s^*}{2}}}\sqrt{n}\left(D^{s^*-1}W_n^j(\theta) - K^{s^*-1}(\theta)\right) \times \\ \left(\left(\sqrt{n}(\theta_n^* - \theta)\right)^{s^*-1}\right) \end{array} \right\| > \frac{\gamma_n^{**}}{2}\right) \end{aligned}$$

which is less than or equal to

$$\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\begin{array}{c}\frac{1}{n^{\frac{s^*-1}{2}}}\sup_{\theta\in\overline{\mathcal{O}}_{\varepsilon_0}(\theta_0)}\left\|D^{s^*}W_n^j(\theta)\right\|\left\|\theta_n^*-\theta\right\|\times\\ \left\|\sqrt{n}\left(\theta_n^*-\theta\right)\right\|^{s^*-1}>\frac{\gamma_n^{**}}{2}\end{array}\right)$$
$$+\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\begin{array}{c}\frac{1}{n^{\frac{s^*}{2}}}\left\|\sqrt{n}\left(D^{s^*-1}W_n^j(\theta)-K^{s^*-1}(\theta)\right)\right\|\times\\ \left\|\sqrt{n}\left(\theta_n^*-\theta\right)\right\|^{s^*-1}>\frac{\gamma_n^{**}}{2}\end{array}\right)$$
$$\leq o\left(n^{-a^*}\right)+\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{1}{(s^*-1)!}\frac{M}{n^{\frac{s^*}{2}}}\ln^{s^*/2}n>\frac{\gamma_n^{**}}{2}\right)$$
$$+\sup_{\theta\in\overline{\mathcal{O}}_\varepsilon(\theta_0)} P_\theta\left(\frac{1}{(s^*-1)!}\frac{C^{s^*}}{n^{\frac{s^*}{2}}}\ln^{s^*/2}n>\frac{\gamma_n^{**}}{2}\right)$$

which is of order $o\left(n^{-a^*}\right)$ when $\gamma_n^{**}=\frac{1}{(s^*-1)!}\frac{2\max\left(M,C^{s^*}\right)}{n^{\frac{s^*}{2}}}\ln^{s^*/2}n=o\left(n^{-a^*}\right)$ independent of θ. Hence due to the rank condition on $E_\theta W_n^j(\theta)$ from assumption A.7.a) the result follows. ■

Lemma AL.7 *For real valued functions f_n, f defined on $\Theta'\supseteq\Theta$, suppose that:* $\sup_{\theta\in\Theta}|f_n-f|=o(1)$, *and* $\sup_{\theta\in\Theta}\left\|D^2f_n\right\|,\sup_{\theta\in\Theta}\left\|D^2f\right\|<M$. *Then* $\sup_{\theta\in\Theta}\left\|Df_n-Df\right\|=o(1)$.

Proof. For any with $\theta_m\neq\theta$ and $D_i=\frac{\partial}{\partial\theta_i}$ for any i

$$\sup_{\theta\in\Theta}\left|D_if_n(\theta)-D_if(\theta)\right|$$
$$\leq\sup_{\theta\in\Theta}\left|D_if_n(\theta)-\frac{f_n(\theta_m)-f_n(\theta)}{\left|\theta_{i_m}-\theta\right|}\right|+\sup_{\theta\in\Theta}\left|\frac{f(\theta_m)-f(\theta)}{\left|\theta_{i_m}-\theta\right|}-D_if(\theta)\right|$$
$$+\sup_{\theta\in\Theta}\left|\frac{f_n(\theta_m)-f(\theta_m)}{\left|\theta_{i_m}-\theta\right|}\right|+\sup_{\theta\in\Theta}\left|\frac{f_n(\theta)-f(\theta)}{\left|\theta_{i_m}-\theta\right|}\right|$$

which is less than or equal

$$2M\left\|\theta_m-\theta\right\|+\frac{1}{c_m}\left(\sup_{\theta\in\Theta}\left|f_n(\theta_m)-f(\theta_m)\right|+\sup_{\theta\in\Theta}\left|f_n(\theta)-f(\theta)\right|\right)$$

where $c_m=\min_{\theta\in\Theta}\left|\theta_{i_m}-\theta\right|$ which exists due to the compactness of Θ and continuity and it is different from zero due to the definition of θ_m, which converges as $n\to\infty$ to

$$2M\left\|\theta_m-\theta\right\|$$

letting *then* $\theta_m\to\theta$ we obtain the needed result. ■

Lemma AL.8 *Suppose that* $\sqrt{n}m_n(\theta)$ *admits a locally uniform Edgeworth expansion, say* $\Psi_{n,s}(\theta)$*, of order* s *over* Θ'*, the polynomials of the density of which, say,* $\pi_i(z,\theta)$ *of* $\Psi_{n,s}(\theta)$ *are equicontinuous on* Θ $\forall z \in \mathbb{R}^q$*, for* $i = 1, \ldots, s-1$*, and* $V(\theta)$ *denotes the variance matrix in the density of* $\Psi_{n,s}(\theta)$ *then it is continuous on* Θ *and positive definite. Let the random element* $\sqrt{n}\gamma_n(\theta)$ *be comprised by elements of* $\sqrt{n}m_n(\theta)$ *such that its support is bounded by* $\sqrt{n}\Gamma$ *for* Γ *a bounded set of some Euclidean space. Then* $\sqrt{n}m_n^*(\theta) = \sqrt{n}\begin{pmatrix} m_n(\theta) \\ \gamma_n(\theta) - E\gamma_n(\theta) \end{pmatrix}$ *admits a locally uniform Edgeworth expansion of order* $s-1$ *over* $\overline{\mathcal{O}}(\theta_0, \delta)$*, the polynomials of the density of which are equicontinuous, as well.*

Proof. As $\sqrt{n}\gamma_n(\theta)$ is part of $\sqrt{n}m_n(\theta)$ (a projection) we have that $\sqrt{n}\gamma_n(\theta)$ admits *a locally uniform Edgeworth expansion of order s* over Θ' (see lemma AL.1 in Arvanitis and Demos [1]), the polynomials of the density of which are equicontinuous on Θ. Due to lemma 3.1, above, we have that

$$\begin{aligned} &\sup_{\theta\in\Theta'}\left|\sqrt{n}E_\theta\gamma_n - \int_{\mathbb{R}} z\left(1+\sum_{i=1}^{s-2}\frac{\pi_i(z,\theta)}{n^{i/2}}\right)\varphi_{V(\theta)}(z)\,dz\right| \\ = &\sup_{\theta\in\Theta'}\left|\sqrt{n}E_\theta\gamma_n - \sum\nolimits_{i=1}^{s-2}\frac{\mathcal{I}_V(k_i(z,\theta))}{n^{i/2}}\right| = o\left(n^{-\frac{s-2}{2}}\right) \end{aligned}$$

where $\left(1+\sum_{i=1}^{s-2}\frac{\pi_i(z,\theta)}{n^{i/2}}\right)\varphi_{V(\theta)}(z)$ denotes the density of the Edgeworth distribution truncated up to the $O\left(n^{-\frac{s-2}{2}}\right)$ order, i.e. of the (obviously) valid locally uniform Edgeworth expansion of order $s-1$, $k_i(z,\theta) = z\pi_i(z,\theta)$ and $\mathcal{I}_V(k_i(z,\theta)) = \int_{\mathbb{R}} k_i(z,\theta)\varphi_{V(\theta)}(z)\,dz$. Using the fact that the π_i's are equicontinuous on Θ it is easy to see that so do the $\mathcal{I}_V(k_i(z,\theta))$. It is also obvious that the random vector $\sqrt{n}l_n(\theta) = \sqrt{n}\begin{pmatrix} m_n(\theta) \\ \gamma_n(\theta)\end{pmatrix}$ admits *a locally uniform Edgeworth expansion of order* $s-1$ over Θ', the polynomials of the density (say π_i^*) of which are equicontinuous on Θ. Consider the vector

$v_n = \begin{pmatrix} 0_{\dim(m_n)} \\ \sum_{i=1}^{s-2} \frac{\mathcal{I}_V(k_i(z,\theta))}{n^{i/2}} \end{pmatrix}$. For an arbitrary Borel set A due to the previous

$$\begin{aligned} & P\left(\sqrt{n}m_n^*(\theta) \in A\right) \\ = \; & P\left(\sqrt{n}l_n(\theta) \in A + v_n + o\left(n^{-\frac{s-2}{2}}\right)\right) \\ = \; & \int_{A\cap\mathcal{H}_n^c(C)} \left(1 + \sum_{i=1}^{s-3} \frac{\pi_i^*\left(z + \begin{pmatrix} 0_{\dim(m_n)} \\ \sum_{i=1}^{s-2} \frac{\mathcal{I}_V(k_i(z,\theta))}{n^{i/2}} \end{pmatrix} + o\left(n^{-\frac{s-2}{2}}\right), \theta\right)}{n^{i/2}}\right) \\ & \times\varphi_{V(\theta)}\left(z + \begin{pmatrix} 0_{\dim(m_n)} \\ \sum_{i=1}^{s-2} \frac{\mathcal{I}_V(k_i(z,\theta))}{n^{i/2}} \end{pmatrix} + o\left(n^{-\frac{s-2}{2}}\right)\right) dz + o\left(n^{-\frac{s-2}{2}}\right) \end{aligned}$$

where $\mathcal{H}_n^c(C)$ analogously to the relevant term in the proof of theorem 3.1 in Arvanitis Demos [1]. Expanding and holding terms of relevant order, by noticing that the π_i are polynomial in z, and that the $o\left(n^{-\frac{s-2}{2}}\right)$ are independent of θ we obtain the needed result. ■

The second auxiliary result is the only one employing the assumption of normality.

Lemma AL.9 *Suppose that $\sqrt{n}(\varphi_n - b(\theta))$ and $\sqrt{n}(\theta_n - \theta)$ admit locally uniform Edgeworth expansions of order s over Θ' the polynomials of the densities of which, say, $\pi_i(z,\theta)$ are equicontinuous on Θ' $\forall z \in \mathbb{R}^3$, for $i = 1,\ldots,s-1$, and the distribution of ε_0 is standard normal. Then $E(\varphi_n(\theta))$ and $E(\theta_n(\theta))$ are two times differentiable on Θ' and for any $\theta \in \Theta'$ and any sequence $\theta_n \neq \theta$ with values in Θ' such that $\|\theta_n - \theta\| \leq C\frac{\ln^{1/2} n}{n^{1/2}}$ for $C > 0$, $i = 1,2$ $\left\|\frac{\partial M_{in}(\theta_n)}{\partial\theta'} - K_i(\theta)\right\| = o(1)$ where $M_{1_n}(\theta) = E(\varphi_n(\theta))$, $M_{2_n}(\theta) = E(\theta_n(\theta))$, $K_1 = \frac{\partial b}{\partial\theta'}$, $K_2 = \mathrm{id}_{\mathbb{R}^3}$.*

Proof. Consider first the case of $E(\varphi_n(\theta))$. Let $\sigma(\varepsilon_0)$ the smallest sub σ-algebra of $\mathcal{F}$ w.r.t. the $\varepsilon_0, \varepsilon_{-1}, \ldots$ are measurable. We have that

$$E(\varphi_n(\theta)) = E(E(\varphi_n(\theta)/\sigma(\varepsilon_0)))$$

Now notice that

$$E(\varphi_n(\theta)/\sigma(\varepsilon_0)) = \int_{\mathbb{R}^n} \varphi_n \frac{1}{\sqrt{(2\pi)^n \prod_{j=1}^n h_i(\theta)}} \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{y_j^2(\theta)}{h_j(\theta)}\right) dz$$

and the differentiability result would follow via the dominated convergence theorem if

$$E\left(\sup_{\theta\in\Theta'}\|s_n(\theta)\|\right) \text{ and } E\left(\sup_{\theta\in\Theta'}\|H_n(\theta)\|\right)$$

are finite where $s_n(\theta) \doteqdot \sum_{j=1}^{n}\left(\varepsilon_j^2-1\right)\frac{1}{h_j(\theta)}\frac{\partial h_j(\theta)}{\partial\theta}$, $H_n(\theta) \doteqdot \sum_{j=1}^{n}\left(\varepsilon_j^2-1\right)\frac{1}{h_j(\theta)}\frac{\partial^2 h_j(\theta)}{\partial\theta\partial\theta'} - \sum_{j=1}^{n}\left(2\varepsilon_j^2-1\right)\frac{1}{h_j^2(\theta)}\frac{\partial h_j(\theta)}{\partial\theta}\frac{\partial h_j(\theta)}{\partial\theta'}$, $\overline{s}_n(\theta) = \frac{1}{n}s_n(\theta)$, $\overline{H}_n(\theta) = \frac{1}{n}H_n(\theta)$. First notice that $h_j(\theta) \geq \underline{\eta}_\omega\left(1-\overline{\eta}_\alpha-\overline{\eta}_\beta\right) \doteqdot c_*$ and due to the fact that

$$\begin{aligned}
\frac{\partial h_j(\theta)}{\partial\theta_1} &= (1-\theta_2-\theta_3)+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial h_{j-1}(\theta)}{\partial\theta_1} \\
\frac{\partial h_j(\theta)}{\partial\theta_2} &= -\theta_1+\varepsilon_{j-1}^2 h_{j-1}(\theta)+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial h_{j-1}(\theta)}{\partial\theta_2} \\
\frac{\partial h_j(\theta)}{\partial\theta_3} &= -\theta_1+h_{j-1}(\theta)+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial h_{j-1}(\theta)}{\partial\theta_3}
\end{aligned}$$

hence

$$\begin{aligned}
& E\left(\sup_{\theta\in\Theta'}\left\|\sum_{j=1}^{n}\left(\varepsilon_j^2-1\right)\frac{1}{h_j(\theta)}\frac{\partial h_j(\theta)}{\partial\theta}\right\|\right) \\
\leq\ & \frac{1}{c_*}\sum_{j=1}^{n}E^{1/2}\left|\varepsilon_j^2-1\right|^2 E^{1/2}\sup_{\theta\in\Theta'}\left\|\frac{\partial h_j(\theta)}{\partial\theta}\right\|^2
\end{aligned}$$

and for $\theta^* = \left(\overline{\eta}_\omega^*, \underline{\eta}_\alpha^*, \underline{\eta}_\beta^*\right)'$ it is easy to see that

$$E\sup_{\theta\in\Theta'}\left\|\frac{\partial h_j(\theta)}{\partial\theta}\right\|^2 \leq E\left\|\frac{\partial h_j(\theta^*)}{\partial\theta}\right\|^2 < +\infty$$

Furthermore, since

$$\begin{aligned}
\frac{\partial^2 h_j(\theta)}{\partial\theta_1^2} &= 0 \\
\frac{\partial^2 h_j(\theta)}{\partial\theta_2^2} &= -\theta_1+\varepsilon_{j-1}^2\frac{\partial h_j(\theta)}{\partial\theta_2}+\varepsilon_{j-1}^2\frac{\partial h_{j-1}(\theta)}{\partial\theta_2}+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial h_{j-1}(\theta)}{\partial\theta_2^2} \\
\frac{\partial^2 h_j(\theta)}{\partial\theta_3^2} &= -\theta_1+2\frac{\partial h_{j-1}(\theta)}{\partial\theta_3}+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial^2 h_{j-1}(\theta)}{\partial\theta_3^2} \\
\frac{\partial^2 h_j(\theta)}{\partial\theta_1\partial\theta_2} &= -1+\varepsilon_{j-1}^2\frac{\partial h_{j-1}(\theta)}{\partial\theta_1}+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial h_{j-1}(\theta)}{\partial\theta_1\partial\theta_2} \\
\frac{\partial^2 h_j(\theta)}{\partial\theta_1\partial\theta_3} &= -1+\frac{\partial h_{j-1}(\theta)}{\partial\theta_1}+\left(\theta_2\varepsilon_{j-1}^2+\theta_3\right)\frac{\partial^2 h_j(\theta)}{\partial\theta_1\partial\theta_3}
\end{aligned}$$

we have that

$$
\begin{aligned}
& E\left(\sup_{\theta\in\Theta'}\left\|\sum\nolimits_{j=1}^{n}\left(\varepsilon_j^2-1\right)\frac{1}{h_j\left(\theta\right)}\frac{\partial^2 h_j\left(\theta\right)}{\partial\theta\partial\theta'}\right\|\right)\\
\leq\; & \frac{1}{c_*}\sum\nolimits_{j=1}^{n}E^{1/2}\left|\varepsilon_j^2-1\right|^2 E^{1/2}\left\|\frac{\partial^2 h_j\left(\theta^*\right)}{\partial\theta\partial\theta'}\right\|^2<+\infty
\end{aligned}
$$

and

$$
\begin{aligned}
& E\left(\sup_{\theta\in\Theta'}\left\|\sum\nolimits_{j=1}^{n}\left(2\varepsilon_j^2-1\right)\frac{1}{h_j^2\left(\theta\right)}\frac{\partial h_j\left(\theta\right)}{\partial\theta}\frac{\partial h_j\left(\theta\right)}{\partial\theta'}\right\|\right)\\
\leq\; & \frac{1}{c_*^2}\sum\nolimits_{j=1}^{n}E^{1/2}\left|2\varepsilon_j^2-1\right|^2 E^{1/2}\left\|\frac{\partial h_j\left(\theta^*\right)}{\partial\theta}\right\|^4<+\infty
\end{aligned}
$$

Next notice that for any θ in Θ' any $i=1,\ldots,3$, and any sequence θ_n as described above we have that

$$
\begin{aligned}
& \left\|\frac{\partial E\left(\varphi_n\left(\theta_n\right)\right)}{\partial\theta_i}-\frac{\partial b\left(\theta\right)}{\partial\theta_i}\right\|\\
\leq\; & 2\sup_{\theta^*\in\Theta'}\left\|\frac{\partial^2 E\left(\varphi_n\left(\theta^*\right)\right)}{\partial\theta_i\partial\theta'}\right\|\left\|\theta_n-\theta\right\|+\left\|\frac{E\left(\varphi_n\left(\theta_n\right)\right)-E\left(\varphi_n\left(\theta\right)\right)}{\theta_{i_n}-\theta_i}-\frac{\partial b\left(\theta\right)}{\partial\theta_i}\right\|
\end{aligned}
$$

Then lemma 2.4, above, implies that due to the behavior of θ_n the last term on the right hand side of the last display is $o\left(1\right)$. Hence the result would follow if $\sup_{\theta^*\in\Theta''}\left\|\frac{\partial^2 E\left(\varphi_n\left(\theta^*\right)\right)}{\partial\theta_i\partial\theta'}\right\|=o\left(\frac{\sqrt{n}}{\ln^{1/2}n}\right)$. The previous along with an application of the Cauchy-Schwarz and the triangle inequalities imply that for any i

$$
\begin{aligned}
& \sup_{\theta\in\Theta'}\left\|\frac{\partial^2 E\left(\varphi_n\left(\theta\right)\right)}{\partial\theta_i\partial\theta'}\right\|\\
\leq\; & \sup_{\theta\in\Theta'}E^{1/2}\left\|\varphi_n\left(\theta\right)-\theta\right\|^2\\
& \times\left(\sup_{\theta\in\Theta'}E^{1/2}\left\|s_n\left(\theta\right)s_n'\left(\theta\right)-EH_n\left(\theta\right)\right\|^2+\sup_{\theta\in\Theta'}E^{1/2}\left\|H_n\left(\theta\right)-EH_n\left(\theta\right)\right\|^2\right)
\end{aligned}
$$

Furthermore, due to assumed Edgeworth approximation for $\sqrt{n}\left(\varphi_n\left(\theta\right)-\theta\right)$, and the fact that $s\geq 5$ lemma 3.1 along with theorem 3.1 in Arvanitis Demos [1] imply that $\sup_{\theta\in\Theta'}E^{1/2}\left\|\left(\varphi_n-b\left(\theta\right)\right)\right\|^2=O\left(\frac{1}{\sqrt{n}}\right)$. Hence the result would follow if

$$
\begin{aligned}
\sup_{\theta\in\Theta'}E\left\|n\overline{s}_n\left(\theta\right)\overline{s}_n'\left(\theta\right)+E\overline{H}_n\left(\theta\right)\right\|^2 &= o\left(\frac{n}{\ln n}\right)\\
\sup_{\theta\in\Theta'}E\left\|\overline{H}_n\left(\theta\right)-E\overline{H}_n\left(\theta\right)\right\|^2 &= o\left(\frac{n}{\ln n}\right)
\end{aligned}
$$

From the proof of Lemma A.1 of Corradi and Inglesias [2], we can prove that $\sqrt{n}\left(S_n^*(\theta) - E\left(S_n^*(\theta)\right)\right)$, where S_n^* contains stacked the elements of $\overline{s}_n$ and $\overline{H}_n$ admits *a locally uniform Edgeworth expansion of order* $s-3$ over Θ' by establishing the conditions A2.M-WD and A3.EL-CPD in Arvanitis Demos [1] through the provision of bounds being independent of θ using the compactness of Θ' and condition A3.NDD in Arvanitis Demos [1] using the result of the referenced proof, the P almost everywhere continuity of the elements of $S_n^*(\theta)$ on Θ', the continuity of determinant and the compactness of Θ'. Then remark R.3 implies that

$$\begin{aligned}
\sup_{\theta\in\Theta'} E\left\|n\overline{s}_n(\theta)\,\overline{s}_n'(\theta) + E\overline{H}_n(\theta)\right\|^2 &= O(1) \\
\sup_{\theta\in\Theta'} E\left\|\overline{H}_n(\theta) - E\overline{H}_n(\theta)\right\|^2 &= O\left(\frac{1}{n}\right)
\end{aligned}$$

which establish the needed bounds. The result about $E\left(\theta_n(\theta)\right)$ is derived analogously. ■