A. ONLINE SUPPLEMENT

to ***Simulation Based Estimation Using Extended Balanced Augmented Empirical Likelihood***

MINH KHOA NGYUEN, University of Essex
WING LON NG, University of Essex
STEVE PHELPS, University of Essex

Before we come to the main proofs we need to introduce some quantities, Lemmas and Remarks that will be needed later on.

LEMMA A.1. *Let Y_i be independent random variables with common distribution and suppose that $E\left(Y_i^2\right) < \infty$. Then*

$$\frac{1}{n}\sum_{i=1}^{n} Y_i^3 = o\left(n^{\frac{1}{2}}\right)$$

with probability 1 as $n \to \infty$.

PROOF. See [Owen 1990], p.98. □

REMARK A.2. *Let Y_i i.i.d. and suppose a measurable function f with $\Sigma_f = Var\left[f\left(Y\right)\right] < \infty$ and $E\left[f\right] = 0$. It follows that $f_i = f\left(Y_i\right)$ are also i.i.d. and $E\left[f^2\right] < \infty$. With Lemma A.1 we have*

$$\frac{1}{n}\sum_{i=1}^{n} f_i^3 = o_p\left(n^{\frac{1}{2}}\right).$$

We introduce the following quantities: a constant s and the function

$$g_{n+2}\left(\theta\right) = 2\overline{g}_n\left(\theta\right) + sc_u\left(\theta\right)u$$

where $\overline{g}_n\left(\theta\right) = \sum_{i=1}^{n} g_i\left(\theta\right)$ and $c_u\left(\theta\right) = \left(u'\hat{S}\left(\theta\right)^{-1}u\right)^{-1/2}$ with

$$\hat{S}\left(\theta\right) = \frac{1}{n}\sum_{i=1}^{n}\left(g_i\left(\theta\right) - \bar{g}_n\left(\theta\right)\right)\left(g_i\left(\theta\right) - \bar{g}_n\left(\theta\right)\right)'.$$

Moreover, $g_{n+1}\left(\theta\right) = -sc_u\left(\theta\right)u$ and

$$S\left(\theta\right) = \frac{1}{n}\sum_{i+1}^{n} g_i\left(\theta\right)g_i\left(\theta\right)',$$

$$\tilde{S}\left(\theta\right) = \frac{1}{n+2}\sum_{i=1}^{n+2} g_i\left(\theta\right)g_i\left(\theta\right)',$$

$$\tilde{\overline{g}}_n\left(\theta\right) = \frac{1}{n+2}\sum_{i=1}^{n+2} g_i\left(\theta\right).$$

Note that

$$\tilde{\overline{g}}_n\left(\theta\right) = \frac{1}{n+2}\left[n\overline{g}_n\left(\theta\right) + 2\overline{g}_n\left(\theta\right)\right] = \overline{g}_n\left(\theta\right).$$

LEMMA A.3. *For given θ suppose $E\left[g\left(Y,\theta\right)g\left(Y,\theta\right)'\right] < \infty$ and $E\left[g\left(Y,\theta\right)\right] < \infty$ then*

$$c_u\left(\theta\right) = O_p\left(1\right).$$

PROOF. Given the assumptions $\Sigma_g(\theta) = E\left[(g(Y,\theta) - E[g(Y,\theta)])(g(Y,\theta) - E[g(Y,\theta)])'\right] < \infty$ exists and $\hat{S}(\theta) \xrightarrow{p} \Sigma_g(\theta)$. Since the variance-covariance matrix $\Sigma_g(\theta)$ is positive-semidefinite (p.s.d.) and symmetric, it has positive eigenvalues. Let $\gamma_1(\theta) \geq \ldots \geq \gamma_d(\theta)$ be the eigenvalues of $\Sigma_g(\theta)$. As $\hat{S}(\theta) \xrightarrow{p} \Sigma_g(\theta)$, for any unit vector η we have $\gamma_1^{-1}(\theta) + o_p(1) \leq \eta' \hat{S}(\theta)^{-1} \eta \leq \gamma_d^{-1}(\theta) + o_p(1)$. With the latter it follows $c_u(\theta) = O_p(1)$. □

LEMMA A.4. *For given θ suppose $E\left[g(Y,\theta) g(Y,\theta)'\right] < \infty$ and $E[g(Y,\theta)] < \infty$ then*

$$\tilde{S}(\theta) \xrightarrow{p} S(\theta).$$

PROOF. With $u = \bar{g}_n(\theta) / \|\bar{g}_n(\theta)\|$ it is

$$\begin{aligned} \tilde{S}(\theta) &= \frac{1}{n+2}\left(\sum_{i=1}^{n} g_i(\theta) g_i(\theta)' + g_{n+1}(\theta) g_{n+1}(\theta)' + g_{n+2}(\theta) g_{n+2}(\theta)'\right) \\ &= \frac{n}{n+2} S(\theta) + \frac{s^2 c_u^2(\theta) + (2\|\bar{g}_n(\theta)\| + s c_u(\theta))^2}{n+2} uu'. \end{aligned} \tag{54}$$

As $\bar{g}_n \xrightarrow{p} E[g(Y,\theta)]$, $\bar{g}_n$ has order $O_p(1)$. With $s = O(1)$ and the given assumptions, $c_u(\theta) = O_p(1)$, (see Lemma A.3) the last term in (54) is of order

$$\begin{aligned} \left[O(1) O_p(1) + (O_p(1) + O(1) O_p(1))^2\right] O\left(n^{-1}\right) &= O_p(1) O\left(n^{-1}\right) \\ &= O_p\left(n^{-1}\right). \end{aligned}$$

Hence, $\tilde{S}(\theta) - S(\theta) \to 0$ in probability as $n \to \infty$. □

A.1. Proof of Theorem 1

The proof in this section is similar to the one in Owen [1990] and Emerson and Owen [2009]. However, to the best of our knowledge there has been no proof published to demonstrate the distributional convergence of the BAEL for unbiased estimation equations. Throughout this proof we assume $\theta = \theta_0$ for which we have $E[g(Y,\theta_0)] = E_{F_0}[g] = 0$. For the rest of this section we will write the argument θ_0 only for emphasis, otherwise we will drop the argument for convenience, i.e. $\bar{g}_n = \sum_{i=1}^{n} g_i(\theta_0)$. Moreover we define $g^* = \max_{i=1:n} \|g_i\|$, $\tilde{g}^* = \max_{i=1:n+2} \|g_i\|$ and the following magnitudes hold: i)[1] $g^* = o_p\left(n^{\frac{1}{2}}\right)$, ii)[2] $\bar{g}_n = O_p\left(n^{-1/2}\right)$ iii) $g_{n+1} = O_p(1)$, iv) $g_{n+2} = O_p(1)$ and v) $\tilde{g}^* = o_p\left(n^{\frac{1}{2}}\right)$.

Note that by assumption, iii) follows from Lemma A.3; that is $c_u = c_u(\theta_0) = O_p(1)$ since Theorem 1 assumes

$$\Sigma_g = E\left[g(Y,\theta_0) g(Y,\theta_0)'\right] < \infty.$$

The latter gives $g_{n+1} = O_p(1)$. Moreover using $c_u = O_p(1)$ and $\bar{g}_n = O_p\left(n^{-1/2}\right)$ with the definition of g_{n+2}, we get

$$\begin{aligned} g_{n+2} &= O_p\left(n^{-1/2}\right) + O_p(1) \\ &= O_p(1). \end{aligned}$$

Finally, as $g_{n+2} = g_{n+1} = O_p(1)$, $\tilde{g}^*$ has the same order as g^*, i.e. $\tilde{g}^* = o_p\left(n^{\frac{1}{2}}\right)$. Before we come to the main proof we need the following Lemma and Remark.

LEMMA A.5. *Let $\theta = \theta_0$ and $E\left[g(Y,\theta_0) g(Y,\theta_0)'\right] < \infty$ and $E[g(Y,\theta_0)] < \infty$, then*

$$\tilde{S} - S \xrightarrow{p} 0$$

[1] [Chen 2008] p.22.
[2] [Qin and Lawless 1994] p.318.

as $n \to \infty$.

PROOF. Similarly to Lemma A.4, $\tilde{S} = \tilde{S}(\theta_0)$ can be written as

$$\tilde{S} = \frac{n}{n+2}S + \frac{s^2c_u^2 + (b_n \|\bar{g}_n\| + sc_u)^2}{n+2}uu'.$$

As c_u is of order $O_p(1)$, $s = O(1)$ and from above $\bar{g}_n$ is order of $O_p\left(n^{-\frac{1}{2}}\right)$, the order of the last term is

$$\begin{aligned}&\left[O(1)O_p(1) + \left(O_p\left(n^{-\frac{1}{2}}\right) + O(1)O_p(1)\right)^2\right]O\left(n^{-1}\right)\\ &= \left[O_p(1) + (o_p(1) + O_p(1))^2\right]O\left(n^{-1}\right)\\ &= [O_p(1) + O_p(1)]\,O\left(n^{-1}\right)\\ &= O\left(n^{-1}\right)\end{aligned}$$

Hence $\widetilde{S} - S \to 0$ in probability as $n \to \infty$. □

REMARK A.6. *As above for $\theta = \theta_0$ we have $E_{F_0}[g] = 0$ and $\Sigma_g = E\left[g(Y,\theta_0)\,g(Y,\theta_0)'\right] < \infty$, then $S \to \Sigma_g$ in probability as $n \to \infty$. Furthermore $\bar{g}_n \to E_{F_0}[g] = 0$ in probability as $n \to \infty$, it follows $\frac{1}{n}\sum_{i+1}^{n}\|g_i\|^2 \to E_{F_0}\left[\|g\|^2\right] < \infty$ in probability as $n \to \infty$. According to Lemma A.5 we have $\widetilde{S} \to \Sigma_g < \infty$ and for $b_n = 2$, it is $\tilde{\bar{g}}_n = \bar{g}_n \to E_{F_0}[g] = 0$ in probability as $n \to \infty$. From the latter two statements it follows $\frac{1}{n+2}\sum_{i+1}^{n+2}\|g_i\|^2 \to E_{F_0}\left[\|g\|^2\right] < \infty$ in probability as $n \to \infty$.*

The remainder of this Section proves Theorem 1. The proof is outlined as follows. First we derive that $\|\lambda\| = O_p\left(n^{-\frac{1}{2}}\right)$. Knowing that, we show $\lambda = \tilde{S}^{-1}\tilde{\bar{g}}_n + o_p\left(n^{-\frac{1}{2}}\right)$, for the sample covariance matrix $\tilde{S}$. We complete the proof by substituting this expression for λ into the profile (balance adjusted) empirical log likelihood ratio statistic $-\tilde{2W}(\theta_0)$, verifying that some other terms are negligible and using Lemma A.5. Accordingly the proof of Theorem 1 is divided into three parts.

Part 1:

PROOF. Without loss of generality let $\sigma_1^2 \leq \ldots \leq \sigma_m^2$ be the eigenvalues of $\Sigma_g = E_{F_0}[gg']$ with $\sigma_1^2 = 1$. For $\theta = \theta_0$ using $\frac{1}{1+x} = 1 - \frac{x}{1+x}$ and $\hat{\lambda} = \lambda/\rho$, $\rho = \|\lambda\|$ in (23) it follows

$$\begin{aligned}0 = \frac{\hat{\lambda}'}{n+2}\sum_{i=1}^{n+2}\frac{g_i}{1+\lambda' g_i} &= \frac{\hat{\lambda}'}{n+2}\sum_{i=1}^{n+2}g_i - \frac{\hat{\lambda}'}{n+2}\sum_{i=1}^{n+2}\frac{g_i\lambda' g_i}{1+\lambda' g_i}\\ &= \hat{\lambda}'\tilde{\bar{g}}_n - \frac{1}{n+2}\sum_{i=1}^{n+2}\frac{\hat{\lambda}' g_i\rho\hat{\lambda}' g_i}{1+\rho\hat{\lambda}' g_i}\\ &= \hat{\lambda}'\tilde{\bar{g}}_n - \frac{\rho}{n+2}\sum_{i=1}^{n+2}\frac{\hat{\lambda}' g_i g_i'\hat{\lambda}}{1+\rho\hat{\lambda}' g_i}\\ &\leq \hat{\lambda}'\tilde{\bar{g}}_n - \frac{\rho}{1+\rho\tilde{g}^*}\hat{\lambda}'\tilde{S}\hat{\lambda}\\ &\leq \hat{\lambda}'\tilde{\bar{g}}_n - \frac{\rho(1-\varepsilon)}{1+\rho\tilde{g}^*}.\end{aligned} \tag{55}$$

The last inequality follows from the fact that $\tilde{S} \xrightarrow{p} \Sigma_g$ (,using $S \xrightarrow{p} \Sigma_g$ and Lemma A.5). Therefore in probabilty for some some $\varepsilon > 0$ we have

$$\hat{\lambda}'\tilde{S}\hat{\lambda} \geq (1-\varepsilon)\,\sigma_1^2 = (1-\varepsilon).$$

Using (55) gives

$$\frac{\rho}{(1+\rho\tilde{g}^*)} \leq \frac{\hat{\lambda}'\tilde{\bar{g}}_n}{1-\varepsilon}. \tag{56}$$

Since $\tilde{\overline{g}}_n = \overline{g}_n$ and $\frac{\hat{\lambda}' \tilde{\overline{g}}_n}{1-\varepsilon}$ is of order $O_p\left(n^{-\frac{1}{2}}\right)$ with equation (56) it follows

$$\rho = \|\lambda\| = O_p\left(n^{-\frac{1}{2}}\right). \tag{57}$$

□

Part 2:

PROOF.
First define $\vartheta_i = \lambda^{'} g_i$. Having established an order bound for $\|\lambda\|$ and with $\tilde{g}^* = o_p\left(n^{\frac{1}{2}}\right)$ it is

$$\max_{i=1:n+2} |\vartheta_i| = O_p\left(n^{-\frac{1}{2}}\right) o_p\left(n^{\frac{1}{2}}\right) = o_p(1). \tag{58}$$

Using $\frac{1}{1+x} = 1 - x - \frac{x^2}{1+x}$ in (23) we get

$$\begin{aligned} 0 = \frac{1}{n+2} \ \textstyle\sum_{i=1}^{n+2} \frac{g_i}{1+\lambda^{'} g_i} &= \textstyle\frac{1}{n+2} \sum_{i=1}^{n+2} g_i \left(1 - \lambda^{'} g_i + \frac{\left(\lambda^{'} g_i\right)^2}{1+\lambda^{'} g_i}\right) \\ = & \quad \tilde{\overline{g}}_n - \tilde{S}\lambda + \textstyle\frac{1}{n+2} \sum_{i=1}^{n+2} \frac{g_i \left(\lambda^{'} g_i\right)^2}{1+\lambda^{'} g_i}. \end{aligned} \tag{59}$$

The last term is bounded above by norm

$$\begin{aligned} \frac{1}{n+2} \sum_{i=1}^{n+2} \frac{g_i\left(\lambda^{'} g_i\right)^2}{1+\lambda^{'} g_i} &\leq \max_{i=1:n+2} \|g_i\| \frac{1}{n+2} \sum_{i=1}^{n+2} \|\lambda\|^2 \|g_i\|^2 \left|1+\lambda^{'} g_i\right|^{-1} \\ &= \tilde{g}^* \|\lambda\|^2 \frac{1}{n+2} \sum_{i=1}^{n+2} \|g_i\|^2 \left|1+\lambda^{'} g_i\right|^{-1}. \end{aligned} \tag{60}$$

With the given order of $\tilde{g}^*$ and λ, Remark A.6 and (58), the order of equation (60) becomes

$$o_p\left(n^{\frac{1}{2}}\right)\left(O_p\left(n^{-\frac{1}{2}}\right)\right)^2 O_p(1)\, O_p(1) = o_p\left(n^{-\frac{1}{2}}\right).$$

Using the latter in equation (59) gives

$$\lambda \ = \ \tilde{S}^{-1}\tilde{\overline{g}}_n + o_p\left(n^{-\frac{1}{2}}\right). \tag{61}$$

□

Part 3:

PROOF. By (58) we may expand

$$\log\left(1+\vartheta_i\right) = \vartheta_i - \frac{1}{2}\vartheta_i^2 + \eta_i, \tag{62}$$

where for some finite $B > 0$,

$$P\left(|\eta_i| \leq B\,|\vartheta_i|^3, 1 \leq i \leq n+2\right) \to 1 \tag{63}$$

as $n \to \infty$. Substituting (62) in (22) we get

$$-2\tilde{W}(\theta_0) = 2\sum_{i=1}^{n+2}\log(1+\vartheta_i) = 2\sum_{i=1}^{n+2}\vartheta_i - \sum_{i=1}^{n+2}\vartheta_i^2 + 2\sum_{i=1}^{n+2}\eta_i.$$

Remark[3] A.2 and (63) give an order bound for the last term

$$\begin{aligned} 2\left|\sum_{i=1}^{n+2}\eta_i\right| &\leq 2B\|\lambda\|^3\sum_{i=1}^{n+2}\|g_i\|^3 \\ &= 2B\|\lambda\|^3\left[\sum_{i=1}^{n}\|g_i\|^3 + \|g_{n+1}\|^3 + \|g_{n+2}\|^3\right] \\ &= 2BO_p\left(n^{-\frac{1}{2}}\right)^3\left[o_p\left(n^{\frac{3}{2}}\right) + O_p(1) + O_p(1)\right] \\ &= 2BO_p\left(n^{-\frac{3}{2}}\right)\left[o_p\left(n^{\frac{3}{2}}\right)\right] = o_p(1). \end{aligned} \tag{64}$$

Let us rewrite (64) by

$$\lambda = \tilde{S}^{-1}\tilde{\bar{g}}_n + \beta,$$

with $\|\beta\| = o_p\left(n^{-\frac{1}{2}}\right)$. Using the latter and re-substituting $\vartheta_i = \lambda' g_i$ in (22) gives

$$\begin{aligned} -2\tilde{W}(\theta_0) &= 2\sum_{i=1}^{n+2}\lambda' g_i - \sum_{i=1}^{n+2}(\lambda' g_i)^2 + o_p(1) \\ &= 2(n+2)\lambda'\tilde{\bar{g}}_n - (n+2)\lambda'\tilde{S}\lambda + o_p(1) \\ &= 2(n+2)\left(\tilde{S}^{-1}\tilde{\bar{g}}_n + \beta\right)'\tilde{\bar{g}}_n - (n+2)\left(\tilde{S}^{-1}\tilde{\bar{g}}_n + \beta\right)'\tilde{S}\left(\tilde{S}^{-1}\tilde{\bar{g}}_n + \beta\right) + o_p(1) \\ &= 2(n+2)\left[\tilde{\bar{g}}_n'\tilde{S}^{-1}\tilde{\bar{g}}_n + \beta'\tilde{\bar{g}}_n\right] - (n+2)\left[\tilde{\bar{g}}_n'\tilde{S}^{-1}\tilde{\bar{g}}_n + 2\beta'\tilde{\bar{g}}_n + \beta'\tilde{S}\beta\right] + o_p(1) \\ &= (n+2)\left[\tilde{\bar{g}}_n'\tilde{S}^{-1}\tilde{\bar{g}}_n\right] + (n+2)\beta'\tilde{S}\beta + o_p(1) \\ &= (n+2)\left[\bar{g}_n'\tilde{S}^{-1}\bar{g}_n\right] + o_p(1). \end{aligned} \tag{65}$$

As $\tilde{S} = O_p(1)$ (using Lemma A.5 and $S \xrightarrow{p} \Sigma_g$), the last equality holds because $\tilde{\bar{g}}_n = \bar{g}_n'$ and

$$(n+2)\beta'\tilde{S}\beta = O(n)\,o_p\left(n^{-1/2}\right)O_p(1)\,o_p\left(n^{-1/2}\right) = o_p(1).$$

Moreover, as $n\bar{g}_n' S^{-1}\bar{g}_n$ converges to a χ^2 distribution with q degrees of freedom, $\tilde{S} \xrightarrow{p} S$ and $\frac{n}{n+2} \to 1$, it follows $-2\tilde{W}(\theta_0) \to \chi_q^2$ in probability as $n \to \infty$. □

A.2. Proof of Theorem 2

PROOF. Suppose $\theta \neq \theta_0$. As before we drop the argument θ, e.g. $\bar{g}_n = \frac{1}{n}\sum_{i=1}^{n} g(y_i,\theta) = \frac{1}{n}\sum_{i=1}^{n} g_i$, $g_{n+1} = -sc_u(\theta)\,u$ and $g_{n+2} = 2\bar{g}_n(\theta) + sc_u(\theta)\,u$. Note, due to the law of large numbers, $\|\bar{g}_n'\bar{g}_n\| \to \delta^2$ and $\bar{g}_n \to \boldsymbol{\mu}(\theta) := E[g(Y,\theta)]$ in probability as $n \to \infty$. By assumption $\Sigma_g(\theta) < \infty$, with Lemma A.3 we have $c_u = O_p(1)$. As $E\left[g(Y,\theta)\,g(Y,\theta)'\right] = \Sigma_g(\theta) + \boldsymbol{\mu}(\theta)\,\boldsymbol{\mu}(\theta)' < \infty$ and $S \xrightarrow{p} E\left[g(Y,\theta)\,g(Y,\theta)'\right]$ with Lemma A.3 ($\tilde{S} \xrightarrow{p} S$) it follows $\tilde{S} = O_p(1)$.

Now, for $i = 1, ..., n$ the terms $g_i - \bar{g}_n$ have expected value zero

$$E[g_i - \bar{g}_n] = 0$$

[3]Under the mild condition of g being a measurable function, it follows with Lemma A.1 that $\sum_{i=1}^{n}\|g_i\|^3 = o\left(n^{\frac{3}{2}}\right)$ as $\frac{1}{n}\sum_{i=1}^{n}\|g_i\|^3 = o\left(n^{\frac{1}{2}}\right)$.

and satisfying all moment conditions such that with Lemma 3 in [Owen 1990, p. 98] it follows that

$$\max_{i=1,\dots,n} \{\|g_i - \bar{g}_n\|\} = o_p\left(n^{1/2}\right). \tag{66}$$

Let $\tilde{\lambda} = n^{-2/3}\bar{g}_n M$ for a positive constant $M > 0$. For $i = 1, \dots, n$

$$\tilde{\lambda}' g_i = \tilde{\lambda}' (g_i - \bar{g}_n) + \tilde{\lambda}' \bar{g}_n. \tag{67}$$

From the above $\bar{g}_n$ is of order $O_p(1)$ therefore the maximum of the first term on the right hand side in (67) is with (66) of order $o_p\left(n^{-2/3}n^{1/2}\right) = o_p(1)$. The last term in (67) has the order $n^{-2/3}O_p(1) = o_p(1)$ hence

$$\max_{i=1,\dots,n} \left\{\left\|\tilde{\lambda}' g_i\right\|\right\} = o_p(1). \tag{68}$$

Since s and u are of $O(1)$ and $c_u = O_p(1)$ it follows that $g_{n+1} = O_p(1)$ and $g_{n+2} = O_p(1)$. Hence $\tilde{\lambda}' g_{n+1} = o_p(1)$ and $\tilde{\lambda}' g_{n+2} = o_p(1)$ therefore

$$\max_{i=1:n+2} \left\{\left\|\tilde{\lambda}' g_i\right\|\right\} = o_p(1). \tag{69}$$

With (69) for $i = 1, \dots, n+2$ we have $1 + \tilde{\lambda}' g_i > 0$ with probability going to 1. Hence using the Taylor expansion:

$$\log(1+x) = x - \frac{x^2}{2(1+\xi)^2} \tag{70}$$

for some ξ between 0 and x and the duality of the maximization problem it is

$$\begin{aligned}
\tilde{W}(\theta) &= -\sup_{\lambda} \left\{\sum_{i=1}^{n+2} \log\left(1 + \lambda' g_i\right)\right\} \\
&\le -\sum_{i=1}^{n+2} \log\left(1 + \tilde{\lambda}' g_i\right) \\
&= -\left[\sum_{i=1}^{n+2} \tilde{\lambda}' g_i - \frac{1}{2}\sum_{i=1}^{n+2} \frac{\left(\tilde{\lambda}' g_i\right)^2}{(1+\xi_i)^2}\right].
\end{aligned} \tag{71}$$

Note, from (69) all ξ_i are within $o_p(1)$ neighborhood of 0 uniformly. Therefore the second term in the last line of (71) is no larger than

$$\sum_{i=1}^{n+2} \left(\tilde{\lambda}' g_i\right)^2 = (n+2)\,\tilde{\lambda}' \tilde{S} \tilde{\lambda} = O(n)\, O_p\left(n^{-2/3}\right) O_p(1)\, O_p\left(n^{-2/3}\right) = o_p(1).$$

The first term is

$$\sum_{i=1}^{n+2} \tilde{\lambda}' g_i = \tilde{\lambda}' n \bar{g}_n + 2\tilde{\lambda}' \bar{g}_n = n^{1/3}\delta^2 M + o_p(1).$$

Therefore (71) gives

$$\tilde{W}(\theta) \le -n^{1/3}\delta^2 M + o_p(1). \tag{72}$$

Since M can be arbitrarily large, we have $-2n^{-1/3}\tilde{W}(\theta) \to \infty$ for any $\theta \neq \theta_0$. □

A.3. Proof of Theorem 3

Before we come to the proof the following Remark.

REMARK A.7. *The last two elements in Assumption 4.1 satisfy the conditions of Lemma A.3, hence*

$$c_u(\theta) = O_p(1)$$

for all $\theta \in \Theta$. The considered assumptions and the WLN give

$$\bar{g}_n(\theta) \xrightarrow{p} E[g(Y,\theta)] < \infty$$

for all $\theta \in \Theta$, that is $\bar{g}_n(\theta) = O_p(1)$. Altogether this result in $\|g_{n+1}(\theta)\| = \|g_{n+2}(\theta)\| = O_p(1)$ for all $\theta \in \Theta$.

The proof of Theorem 3 is almost the same as that of Newey and Smith [2004] and is divided into four parts (three Lemmas and the main proof).

LEMMA A.8. *If Assumption 4.1 is satisfied, then for any ζ with $1/\alpha < \zeta < 1/2$ and $\Lambda_n = \left\{\lambda : \|\lambda\| \le n^{-\zeta}\right\}$, we have $\sup_{\theta\in\Theta,\lambda\in\Lambda_n,i=1,\ldots,n+2} |\lambda' g_i(\theta)| \xrightarrow{p} 0$ and with probability approaching (w.p.a.) 1, $\Lambda_n \subseteq \hat{\Lambda}_n(\theta)$ for all $\theta \in \Theta$.*

PROOF.

$$\begin{aligned}
\sup_{\theta\in\Theta,\lambda\in\Lambda_n,i=1,\ldots,n+2} \left|\lambda' g_i(\theta)\right| &\le \sup_{\lambda\in\Lambda_n} \|\lambda\| \max_{i=1,\ldots,n+2} \sup_{\theta\in\Theta} \|g_i(\theta)\| \\
&= \sup_{\lambda\in\Lambda_n} \|\lambda\| \left(\max_{i=1,\ldots,n+2} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha\right)^{1/\alpha} \\
&\le \sup_{\lambda\in\Lambda_n} \|\lambda\| \left(\sum_{i=1}^{n+2} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha\right)^{1/\alpha} \\
&= n^{1/\alpha} \sup_{\lambda\in\Lambda_n} \|\lambda\| \left(\frac{1}{n}\sum_{i=1}^{n+2} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha\right)^{1/\alpha} \\
&= n^{1/\alpha} \sup_{\lambda\in\Lambda_n} \|\lambda\| \left(\frac{1}{n}\sum_{i=1}^{n} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha + \frac{1}{n}\sum_{i=n+1}^{n+2} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha\right)^{1/\alpha} \\
&= n^{1/\alpha} O\left(n^{-\zeta}\right)\left(O_p(1) + O_p\left(n^{-1}\right)\right) \\
&= o_p(1)
\end{aligned}$$

The second to last line holds due to Remark A.7 and due the assumption $E\left[\sup_{\theta\in\Theta} \|g(y,\theta)\|^\alpha\right] < \infty$ for some $\alpha > 2$ that gives $\frac{1}{n}\sum_{i=1}^{n} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha \xrightarrow{p} E\left[\sup_{\theta\in\Theta} \|g(y,\theta)\|^\alpha\right] < \infty$, i.e. $\frac{1}{n}\sum_{i=1}^{n} \sup_{\theta\in\Theta} \|g_i(\theta)\|^\alpha = O_p(1)$. Overall it follows w.p.a. 1 $\lambda' g_i(\theta) \in \Im$ for all $\theta \in \Theta$ and $\|\lambda\| \le n^{-\zeta}$. $\square$

LEMMA A.9. *If Assumption 4.1 is satisfied, $\bar{\theta} \in \Theta$, with $\bar{\theta} \xrightarrow{p} \theta_0$ and $\tilde{\bar{g}}_n(\bar{\theta}) = O_p\left(n^{-1/2}\right)$, then $\grave{\lambda} = argmax_{\lambda\in\hat{\Lambda}_n(\bar{\theta})} \hat{P}(\bar{\theta},\lambda) = argmax_{\lambda\in\hat{\Lambda}_n(\bar{\theta})} \sum_{i=1}^{n+2} \rho\left(\lambda' g_i(\bar{\theta})\right)/(n+2)$ exists with w.p.a. 1, $\grave{\lambda} = O_p\left(n^{-1/2}\right)$ and $\sup_{\lambda\in\hat{\Lambda}_n(\bar{\theta})} \hat{P}(\bar{\theta},\lambda) \le \rho_0 + O_p\left(n^{-1}\right)$.*

PROOF. Since Assumption 4.1 satisfy the conditions of Lemma A.4 we have $\tilde{S}(\bar{\theta}) \xrightarrow{p} S(\bar{\theta})$. By assumptions and the UWL (Uniform Weak Law of Large Numbers) we have $S(\bar{\theta}) \xrightarrow{p} \sum_g(\theta_0)$, hence $\tilde{S}(\bar{\theta}) \xrightarrow{p} \sum_g(\theta_0)$. By $\sum_g(\theta_0) < \infty$ the smallest eigenvalue $\tilde{S}(\bar{\theta})$ is bounded away from 0 w.p.a. 1. Since $\rho(v)$ is twice continuously differentiable in the neighborhood of 0 with Lemma A.8 it follows $\hat{P}(\bar{\theta},\lambda)$ is twice continuously differentiable on Λ_n with w.p.a. 1. Hence, $\check{\lambda} = argmax_{\lambda\in\Lambda_n}\hat{P}(\bar{\theta},\lambda)$ exists with w.p.a 1. Furthermore, for $\bar{g}_i = g_i(\bar{\theta})$ and any $\dot{\lambda}$ on the line joining $\check{\lambda}$ and 0 it follows from Lemma A.8 and $\rho_2 = -1$ that $max_{1\le i\le n+2}\rho_2\left(\dot{\lambda}g_i(\bar{\theta})\right) < -1/2$ with w.p.a. 1. Then using the Taylor Expansion of $\hat{P}(\bar{\theta},\lambda)$ around $\lambda = 0$ and $\dot{\lambda}$ on the line joining $\check{\lambda}$ and 0 we get

$$\begin{aligned}\rho_0 = \hat{P}(\bar{\theta},0) \le \hat{P}(\bar{\theta},\check{\lambda}) &= \rho_0 - \check{\lambda}'\tilde{\bar{g}}_n(\bar{\theta}) + \frac{1}{2}\check{\lambda}'\left[\sum_{i=1}^{n+2}\rho_2\left(\dot{\lambda}'\bar{g}_i\right)\bar{g}_i\bar{g}_i'/(n+2)\right]\check{\lambda} \\ &\le \rho_0 - \check{\lambda}\tilde{\bar{g}}_n(\bar{\theta}) - \frac{1}{4}\check{\lambda}'\tilde{S}(\bar{\theta})\check{\lambda} \\ &\le \rho_0 + \|\check{\lambda}\|\,\|\tilde{\bar{g}}_n(\bar{\theta})\| - C_1\|\check{\lambda}\|^2, \end{aligned} \tag{73}$$

where C_1 a positive constant. Subtracting $\rho_0 - C_1\|\check{\lambda}\|^2$ from both sides and dividing $\|\check{\lambda}\|^2$ we get $C_1\|\check{\lambda}\| \le \|\tilde{\bar{g}}_n\|$w.p.a. 1. By assumption we have $\tilde{\bar{g}}_n(\bar{\theta}) = O_p\left(n^{-1/2}\right)$, therefore $\|\check{\lambda}\| = O_p\left(n^{-1/2}\right) = o_p\left(n^{-\zeta}\right)$. From the latter it follows that $\check{\lambda}\in int(\Lambda_n)$ w.p.a.1 and with Lemma A.8 $\check{\lambda}\in\hat{\Lambda}_n(\bar{\theta})$ w.p.a. 1. By concavity of $\hat{P}(\bar{\theta},\lambda)$ and convexity of $\hat{\Lambda}_n(\bar{\theta})$ it follows $\hat{P}(\bar{\theta},\check{\lambda}) = \sup_{\lambda\in\hat{\Lambda}_n(\bar{\theta})}\hat{P}(\bar{\theta},\lambda)$ and therefore $\grave{\lambda} = \check{\lambda}$. Using $\tilde{\bar{g}}_n(\bar{\theta}) = O_p\left(n^{-1/2}\right)$, $\left\|\grave{\lambda}\right\| = O_p\left(n^{-1/2}\right)$ and in (73) we get

$$\hat{P}\left(\bar{\theta},\grave{\lambda}\right) \le \rho_0 + \left\|\grave{\lambda}\right\|\|\tilde{\bar{g}}_n(\bar{\theta})\| - C_1\left\|\grave{\lambda}\right\|^2 = \rho_0 + O_p\left(n^{-1}\right).$$

□

LEMMA A.10. *If Assumption 4.1, then*

$$\left\|\tilde{\bar{g}}_n\left(\hat{\theta}\right)\right\| = O_p\left(n^{-1/2}\right).$$

PROOF. Let $\hat{g}_i = g_i\left(\hat{\theta}\right)$, $\hat{g} = \tilde{\bar{g}}_n\left(\hat{\theta}\right)$ and for ζ in Lemma A.8, $\check{\lambda} = -n^{-\zeta}\hat{g}/\|\hat{g}\|$. With Lemma A.8 it follows $\max_{1\le i\le n+2}\left|\check{\lambda}'\hat{g}_i\right| \xrightarrow{p} 0$ and $\check{\lambda}\in\Lambda_n\left(\hat{\theta}\right)$w.p.a. 1. Then for any $\dot{\lambda}$ on the line joining $\check{\lambda}$ and 0 w.p.a 1 we have $\rho_2\left(\dot{\lambda}'\hat{g}_i\right) \ge -C_2$ for all $i = 1,...,n+2$, where C_2 is a positive constant. Given Assumption 4.1, Lemma (A.4) gives $\frac{1}{n+2}\sum_{i=1}^{n+2}\hat{g}_i\hat{g}_i' \xrightarrow{p} \frac{1}{n}\sum_i\hat{g}_i\hat{g}_i'$ and by CS (Cauchy-Schwarz inequality) and UWL it is $\frac{1}{n}\sum_{i=1}^{n}\hat{g}_i\hat{g}_i' \le \left(\frac{1}{n}\sum_{i=1}^{n}\left(\sup_{\theta\in\Theta}\|g_i(\theta)\|\right)^2\right)I \xrightarrow{p} C_3I$, where C_3 is a positive constant. From the latter it follows that the largest eigenvalue of $\frac{1}{n+2}\sum_i\hat{g}_i\hat{g}_i'$ is bounded above w.p.a. 1. Using Taylor Expansion as before

$$\begin{aligned}\hat{P}\left(\hat{\theta},\check{\lambda}\right) &= \rho_0 - \check{\lambda}'\hat{g} + \frac{1}{2}\check{\lambda}'\left[\sum_{i=1}^{n+2}\rho_2\left(\dot{\lambda}'\hat{g}_i\right)\hat{g}_i\hat{g}_i'/(n+2)\right]\check{\lambda} \\ &\ge \rho_0 + n^{-\zeta}\|\hat{g}\| - C_2\frac{1}{2}\check{\lambda}'\left[\sum_{i=1}^{n+2}\hat{g}_i\hat{g}_i'/(n+2)\right]\check{\lambda} \\ &\ge \rho_0 + n^{-\zeta}\|\hat{g}\| - Cn^{-2\zeta} \end{aligned} \tag{74}$$

w.p.a. 1, where $C = C_2C_3$. By the Lindeberg-Levy central limit theorem the hypothesis of Lemma A.9 are satisfied by[4] $\bar{\theta} = \theta_0$. As $\hat{\theta}$ and $\hat{\lambda}$ being saddle point solutions, (74) and Lemma A.9 gives:

[4]Note, $\tilde{\bar{g}}_n(\theta_0) = \bar{g}_n(\theta_0)$.

$$\rho_0 + n^{-\zeta} \|\hat{g}\| - Cn^{-2\zeta} \leq \hat{P}\left(\hat{\theta}, \check{\lambda}\right) \leq \hat{P}\left(\hat{\theta}, \hat{\lambda}\right) \leq \sup_{\lambda \in \hat{\Lambda}_n(\theta_0)} \hat{P}\left(\theta_0, \lambda\right) \leq \rho_0 + O_p\left(n^{-1}\right). \tag{75}$$

Solving the latter for $\|\hat{g}\|$ gives

$$\|\hat{g}\| \leq O_p\left(n^{\zeta-1}\right) + Cn^{-\zeta} = O_p\left(n^{-\zeta}\right). \tag{76}$$

The last equality holds because by assumption $\zeta < 1/2$, thus $\zeta - 1 < -1/2 < -\zeta$. Now consider $\varepsilon_n \to 0$ and let $\grave{\lambda} = -\varepsilon_n \hat{g}$, with (76) $\grave{\lambda} = o_p\left(n^{-\zeta}\right)$, so that $\grave{\lambda} \in \Lambda_n$ w.p.a. 1. Then as in (75)

$$\rho_0 - \grave{\lambda}'\hat{g} - C\left\|\grave{\lambda}\right\|^2 = \rho_0 + \varepsilon_n \|\hat{g}\|^2 - C\varepsilon_n^2 \|\hat{g}\|^2 = \rho_0 + (1 - C\varepsilon_n)\,\varepsilon_n \|\hat{g}\|^2 \leq \rho_0 + O_p\left(n^{-1}\right).$$

Since for large enough n, $1 - C_1\varepsilon_n$ is bounded away from 0 w.p.a. 1 and it follows from the latter equation, $\varepsilon_n \|\hat{g}\|^2 = O_p\left(n^{-1}\right)$. The final conclusion follows by standard result from probability theory, that if $\varepsilon_n Y_n = O_p(n^{-1})$ for all $\varepsilon_n \to 0$ then $Y_n = O_p\left(n^{-1}\right)$. $\square$

Provided with the given Lemma A.8-A.10 the following proofs Theorem 3.

PROOF. First note, $\tilde{\bar{g}}_n(\theta) = \bar{g}_n(\theta)$ then

$$\left\|\tilde{\bar{g}}_n\left(\hat{\theta}\right) - E\left[g\left(Y, \hat{\theta}\right)\right]\right\| = \left\|\bar{g}_n\left(\hat{\theta}\right) - E\left[g\left(Y, \hat{\theta}\right)\right]\right\| \leq \sup_{\theta \in \Theta} \|\bar{g}_n(\theta) - E\left[g\left(Y, \theta\right)\right]\| \xrightarrow{p} 0,$$

where the latter follows from the assumptions and the UWL. As Lemma A.10 gives $\tilde{\bar{g}}_n\left(\hat{\theta}\right) \xrightarrow{p} 0$ it follows from above $E\left[g\left(Y, \hat{\theta}\right)\right] \xrightarrow{p} 0$. By assumption $E\left[g\left(Y, \theta\right)\right] = 0$ has a unique solution at θ_0, hence $\|E\left[g\left(Y, \theta\right)\right]\|$ must be bounded away from 0 outside any neighborhood of θ_0. Therefore $\hat{\theta}$ must be inside any neighborhood of θ_0 w.p.a. 1, i.e. $\hat{\theta} \xrightarrow{p} \theta_0$. With Lemma A.10 ($\left\|\tilde{\bar{g}}_n\left(\hat{\theta}\right)\right\| = O_p\left(n^{-1/2}\right)$) and $\bar{\theta} = \hat{\theta}$ the hypotheses in Lemma A.9 are satisfied, hence $\hat{\lambda} = argmax_{\lambda \in \hat{\Lambda}_n(\hat{\theta})} \hat{P}\left(\hat{\theta}, \lambda\right) = argmax_{\lambda \in \hat{\Lambda}_n(\hat{\theta})} \sum_{i=1}^{n+2} \rho\left(\lambda' g_i\left(\hat{\theta}\right)\right) / (n+2)$ exists with w.p.a. 1, $\hat{\lambda} = O_p\left(n^{-1/2}\right)$. $\square$

References

CHEN, J. 2008. Adjusted Empirical Likelihood and its Properties. *Journal of Computational and Graphical Statistics 17,* 2, 426–443.

CHEN, J., VARIYATH, A., AND ABRAHAM, B. 2008. Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics 17,* 2, 426–443.

EMERSON, S. AND OWEN, A. 2009. Calibration of the empirical likelihood method for a vector mean. *Electronic Journal of Statistics 3*, 1161–1192.

NEWEY, W. AND SMITH, R. 2004. Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica 72,* 1, 219–255.

OWEN, A. 1990. Empirical Likelihood Ratio Confidence Region. *The Annals of Statistics 18,* 1, 90–120.

QIN, J. AND LAWLESS, J. 1994. Empirical Likelihood and General Estimating Equations. *The Annals of Statistics 22,* 1, 300–325.