TWO-WAY MODELS FOR GRAVITY

SUPPLEMENTARY MATERIAL

Koen Jochmans*

Draft: February 2, 2015 Revised: November 18, 2015

This version: April 11, 2016

I. Proofs of theorems in the main text

Proof of Theorem 1. We establish consistency by verifying Conditions (i)–(iv) of Theorem 2.1 in Newey and McFadden (1994). Assumptions 1 and 2 imply that Conditions (i)–(iii) hold. Condition (iv) states that $s(\psi)$ converges in probability to $\overline{s}(\psi)$ uniformly on $\mathcal{S}$ and remains to be shown. By definition we need to show that

$$\lim_{n\to\infty} \Pr\left(\sup_{\psi\in\mathcal{S}} \|s(\psi) - E[s(\psi)]\| > \epsilon\right) = 0$$

for any $\epsilon > 0$.

By symmetry,

$$s(\psi) - E[s(\psi)] = \frac{\varrho^{-1}}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i'\neq i} \sum_{j'\neq j} \overline{v}(\{i,i',j,j'\},\psi),$$

where $\overline{v}(\{i,i',j,j'\},\psi) = v(\{i,i',j,j'\},\psi) - E[v(\{i,i',j,j'\},\psi)]$ and we introduce the notational shorthand

$$v(\{i,i',j,j'\},\psi) = \phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)(u_{ij}(\psi)\, u_{i'j'}(\psi) - u_{ij'}(\psi)\, u_{i'j}(\psi)).$$

* Address: Sciences Po, Département d'économie, 28 rue des Saints Pères, 75007 Paris, France; koen.jochmans@sciencespo.fr.

By the Cauchy-Schwarz inequality,

$$E[\|\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi)\, u_{ij}(\psi)\, u_{i'j'}(\psi)\|^2]^2$$

is bounded by

$$E[\|\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi)\|^4] \times \sqrt{E[\|u_{ij}(\psi)\|^8]}\,\sqrt{E[\|u_{i'j'}(\psi)\|^8]}.$$

By Assumption 4, these terms are uniformly bounded on $\mathcal{S}$ for any i, i', j, j'. Therefore, there exists a constant C so that

$$E[\|v(\{i, i', j, j'\}, \psi)\|^2] < C.$$

This, in turn, implies that the variance of $s(\psi)$ is uniformly bounded. Chebychev's inequality yields

$$\Pr\left(\|s(\psi) - E[s(\psi)]\| > \epsilon\right) \leq \frac{E[\|s(\psi) - E[s(\psi)]\|^2]}{\epsilon}.$$

The numerator on the right-hand side is bounded by

$$\frac{\begin{matrix}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}\sum_{i_2\neq i_1}\sum_{j_2\neq j_1} \\ \sum_{i_3=1}^{n}\sum_{j_3=1}^{n}\sum_{i_4\neq i_3}\sum_{j_4\neq j_3}\end{matrix}\; E[\|\overline{v}(\{i_1, i_2, j_1, j_2\}, \psi)\|\,\|\overline{v}(\{i_3, i_4, j_3, j_4\}, \psi)\|]}{16\varrho^2}$$

The covariance between $v(\{i_1, i_2, j_1, j_2\}, \psi)$ and $v(\{i_3, i_4, j_3, j_4\}, \psi)$ depends on how many of the indices are common across the quadruples (i_1, i_2, j_1, j_2) and (i_3, i_4, j_3, j_4). The correlation is non-zero as soon as these sets overlap. When the sets are disjoint, the terms $v(\{i_1, i_2, j_1, j_2\}, \psi)$ and $v(\{i_3, i_4, j_3, j_4\}, \psi)$ are independent by virtue of Assumption 3. Of the $O(n^8)$ possible combinations of observations, $O(n^7)$ combinations have dyads that overlap. Hence, uniformly on $\mathcal{S}$,

$$\Pr\left(\|s(\psi) - E[s(\psi)]\| > \epsilon\right) = O(n^{-1}),$$

which converges to zero as $n \to \infty$ for any $\epsilon > 0$. Therefore, uniform convergence of the empirical moment condition $s(\psi)$ to $\overline{s}(\psi)$ has been established. With all conditions of Theorem 2.1 in Newey and McFadden (1994) fullfilled we have established that

$$\psi_n - \psi_0 \xrightarrow{p} 0$$

as $n \to \infty$. The proof is complete. □

Proof of Theorem 2. The proof proceeds in three steps. We first show (2.2). We next establish the uniform convergence of the Jacobian matrix of the moment conditions. We then combine these results with a Taylor expansion to establish that

$$\text{(a)} \qquad n\,(\psi_n - \psi_0) = -(\Sigma' \Omega \Sigma)^{-1}\, \Sigma' \Omega\, \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\,(\varepsilon_{ij} - 1) + o_p(1),$$

and apply a suitable central limit theorem to the right-hand side of this equation to validate Theorem 2.

(i) Asymptotic approximation of the moment conditions. At ψ_0 the empirical moment conditions are

$$\text{(b)} \qquad s(\psi_0) = \frac{\varrho^{-1}}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i' \neq i} \sum_{j' \neq j} \phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)(u_{ij}\, u_{i'j'} - u_{ij'}\, u_{i'j}),$$

where we have exploited the symmetry of $s(\psi_0)$ in (i, i') and (j, j'). A small calculation shows that the Hájek projection (van der Vaart, 2000, Section 11.3) of $s(\psi_0)$, conditional on the covariates, equals

$$p_n = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\,(\varepsilon_{ij} - 1).$$

Note that $E[\,p_n\,|\,x_{11},\ldots,x_{nn}] = 0$ and that

$$V_* = n^2\,\mathrm{var}(p_n) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E[w_{ij}w_{ij}'\,\sigma_{ij}^2].$$

To show that $s(\psi_0)$ is asymptotically equivalent to p_n in the sense of (2.2) it suffices to show that

$$\text{(c)}\qquad n^2\,E[(p_n - s(\psi_0))(p_n - s(\psi_0))'] \to 0$$

as $n \to \infty$ (see, e.g., van der Vaart 2000, Chapter 12).

The main step needed to establish (c) is the calculation of the variance of the moment conditions $s(\psi_0)$. Use (b) to see that $\mathrm{var}(s(\psi_0)) = E[s(\psi_0)s(\psi_0)']$ equals the expectation of the matrix

$$\text{(d)}\qquad \begin{aligned} &\frac{\varrho^{-1}}{4}\sum_{i_1=1}^{n}\sum_{i_2\neq i_1}\sum_{j_1=1}^{n}\sum_{j_2\neq j_1}\phi(x_{i_1j_1},x_{i_1j_2},x_{i_2j_1},x_{i_2j_2};\psi_0)\,(u_{i_1j_1}\,u_{i_2j_2} - u_{i_1j_2}\,u_{i_2j_1}) \\ \times &\frac{\varrho^{-1}}{4}\sum_{i_3=1}^{n}\sum_{i_4\neq i_3}\sum_{j_3=1}^{n}\sum_{j_4\neq j_3}\phi(x_{i_3j_3},x_{i_3j_4},x_{i_4j_3},x_{i_4j_4};\psi_0)'(u_{i_3j_3}\,u_{i_4j_4} - u_{i_3j_4}\,u_{i_4j_3}). \end{aligned}$$

Because $u_{ij} = \alpha_i\gamma_j\,\varepsilon_{ij}$ and the ε_{ij} are independent across both i and j, we have that

$$\text{(e)}\qquad E[(u_{i_1j_1}\,u_{i_2j_2} - u_{i_1j_2}\,u_{i_2j_1})(u_{i_3j_3}\,u_{i_4j_4} - u_{i_3j_4}\,u_{i_4j_3})|\,x_{11},\ldots,x_{nn}]$$

equals zero unless any of the dyads in $\{(i_1,j_1),(i_2,j_2),(i_1,j_2),(i_2,j_1)\}$ co-incides with any of the dyads in $\{(i_3,j_3),(i_4,j_4),(i_3,j_4),(i_4,j_3)\}$. Of the $O(n^8)$ terms in $\mathrm{var}(s(\psi_0))$, $O(n^6)$ have at least dyad in common. Moreover, the number of terms with two or more dyads in common is $O(n^4)$. Because

$$\mathrm{var}(s(\psi_0)) = \frac{O(n^6)}{\varrho^2} = \frac{O(n^6)}{O(n^8)} = O(n^{-2}),$$

only terms with at least one dyad in common provide a non-zero contribution to the asymptotic variance. By symmetry of (d), all the expressions are permutation invariant and so we are free to choose a dyad that is common across terms in our calculations and multiply through the resulting expression by 4^2, thereby accounting for all possible choices. With $(i_3, j_3) = (i_1, j_1)$, the expectation in (e) equals

$$\alpha_{i_1}^2 \gamma_{j_1}^2 \alpha_{i_2} \gamma_{j_2} \alpha_{i_4} \gamma_{j_4} \sigma_{i_1 j_1}^2, \tag{f}$$

Setting $(i_3, j_3) = (i_1, j_1)$ in (d), and using (f) and the definition of w_{ij} given in the text we find

$$n^2 \operatorname{var}(s(\psi_0)) = V_* + o(1).$$

The same argument can be used to show that $n^2 E[s(\psi_0) \, p_n'] = V_* + o(1)$. Hence,

$$E[(p_n - s(\psi_0))(p_n - s(\psi_0))'] = o(n^{-2})$$

and (c) has been shown.

(ii) Uniform convergence of the Jacobian matrix. Differentiating $s(\psi)$ gives the Jacobian as

$$\begin{aligned} \varrho \, S(\psi) = \sum_{i=1}^{n} \sum_{i<i'} \sum_{j=1}^{n} \sum_{j<j'} \phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi) \, \frac{\partial (u_{ij}(\psi) u_{i'j'}(\psi) - u_{i'j}(\psi) u_{ij'}(\psi))}{\partial \psi'} \\ + \frac{\partial \phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi)}{\partial \psi} \, (u_{ij}(\psi) u_{i'j'}(\psi) - u_{i'j}(\psi) u_{ij'}(\psi)). \end{aligned}$$

Convergence of the second term to its expectation follows as in the proof of Theorem 1, with ϕ' replacing ϕ, by Assumptions 4 and 5. For the first term, observe that

$$\frac{\partial (u_{ij}(\psi) u_{i'j'}(\psi) - u_{i'j}(\psi) u_{ij'}(\psi))}{\partial \psi'}$$

equals

$$u_{i'j}(\psi)u_{ij'}(\psi)(\tau(x_{i'j};\psi)'+\tau(x_{ij'};\psi)') - u_{ij}(\psi)u_{i'j'}(\psi)(\tau(x_{ij};\psi)'+\tau(x_{i'j'};\psi)').$$

Assumptions 4 and 5 imply that

$$E[\|\phi(x_{ij},x_{ij'},x_{i'j},x_{i'j'};\psi)\, u_{ij}(\psi)u_{i'j'}(\psi)\, \tau(x_{ij};\psi)'\|^2] < C$$

for some finite constant C. Therefore, again by the same argument as in the proof of Theorem 1, this term converges uniformly to its expectation which, as will be verified below, equals Σ. By Theorem 1, $\|\psi_n - \psi_0\| = o_p(1)$ as $n \to \infty$. Therefore,

$$\|S(\psi) - \Sigma\| \xrightarrow{p} 0$$

for any ψ that lies in between ψ_n and ψ_0. This conclusion, together with the asymptotic equivalence of $s(\psi_0)$ and p_n, can be combined with a mean-value expansion of $s(\psi_n)$ around ψ_0 to obtain the sampling-error representation for $n\,(\psi_n - \psi_0)$ in (a).

To see that the limit of the expectation of $S(\psi_0)$ equals Σ, first note that the term involving ϕ' drops out because

$$E[\phi'(x_{ij},x_{ij'},x_{i'j},x_{i'j'};\psi_0)\,(u_{ij}u_{i'j'} - u_{i'j}u_{ij'})] = 0.$$

Therefore, up to $o_p(1)$, $S(\psi_0)$ equals

$$\varrho^{-1}\sum_{i=1}^{n}\sum_{i<i'}\sum_{j=1}^{n}\sum_{j<j'} E[\phi(x_{ij},x_{ij'},x_{i'j},x_{i'j'};\psi_0)\,((\tau_{i'j}+\tau_{ij'})-(\tau_{ij}+\tau_{i'j'}))'\,\alpha_i\alpha_{i'}\gamma_j\gamma_{j'}],$$

where we let $\tau_{ij} = \tau(x_{ij};\psi_0)$. Exploit symmetry and expand the sum on the right-hand side to see that

$$\begin{aligned} S(\psi_0) = \frac{\varrho^{-1}}{4} \sum_{i=1}^n \sum_{i' \neq i} \sum_{j=1}^n \sum_{j' \neq j} E[\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)\, (\alpha_i \alpha_{i'} \gamma_j \gamma_{j'})\, \tau'_{i'j}] \\ + \frac{\varrho^{-1}}{4} \sum_{i=1}^n \sum_{i' \neq i} \sum_{j=1}^n \sum_{j' \neq j} E[\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)\, (\alpha_i \alpha_{i'} \gamma_j \gamma_{j'}) \tau'_{ij'}\,] \\ - \frac{\varrho^{-1}}{4} \sum_{i=1}^n \sum_{i' \neq i} \sum_{j=1}^n \sum_{j' \neq j} E[\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)\, (\alpha_i \alpha_{i'} \gamma_j \gamma_{j'}) \tau'_{ij}\;\,] \\ - \frac{\varrho^{-1}}{4} \sum_{i=1}^n \sum_{i' \neq i} \sum_{j=1}^n \sum_{j' \neq j} E[\phi(x_{ij}, x_{ij'}, x_{i'j}, x_{i'j'}; \psi_0)\, (\alpha_i \alpha_{i'} \gamma_j \gamma_{j'}) \tau'_{i'j'}] + o_p(1). \end{aligned}$$

By permutation invariance, the fourth right-hand side term is identical to the third while the first and second right-hand side terms are identical to the third up to sign. Collapsing the four expressions on the right-hand side and using the definition of w_{ij} we therefore find that

$$S(\psi_0) = -\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E[w_{ij}\, \tau'_{ij}] + o_p(1) \to \Sigma$$

as $n \to \infty$.

(iii) Central limit theorem. Steps (i) and (ii) validate the linear approximation stated in (a). Theorem 2 will then follow by showing that

$$n\, V^{-1/2}\, p_n \xrightarrow{d} N(0, I), \qquad \text{(g)}$$

where I denotes the identity matrix of conformable dimension. To do so note that, conditional on $x_{11}, \ldots, x_{nn}$, the Hájek projection p_n is an average of independent heterogeneously-distributed zero-mean random variables with variance $n^{-2}\, W_*$, where

$$W_* = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} w'_{ij}\, \sigma^2_{ij}.$$

By virtue of Assumption 4, p_n satisfies Lyapunov's condition and so, applying a conditional version of the central limit theorem (see, e.g., Prakasa Rao 2009), we have

$$n\, W_*^{-1/2}\, p_n \xrightarrow{d} N(0, I),$$

conditional on $x_{11}, \ldots, x_{nn}$. Now, $\|W_* - V\| \leq \|W_* - V_*\| + \|V_* - V\|$ by the triangle inequality, and each of these right-hand side terms converges to zero in probability as $n \to \infty$. Therefore, the conditional limit distribution of $n\, p_n$ is normal with zero mean and constant covariance V. Because this distribution is independent of the covariate values $x_{11}, \ldots, x_{nn}$ it equals the unconditional limit distribution. This yields (g). Therefore, the theorem has been shown. □

II. Instrument calculations

We calculate instruments as in Chamberlain (1987) for 2×2 (one-quad) data (so that there is only one conditional moment condition) for exponential-regression models.

Write $\varphi_{ij} = \varphi(x'_{ij}\psi)$, let φ'_{ij} denote the first derivative, and let σ^2_{ij} denote the conditional variance of ε_{ij}. A small calculation reveals that Chamberlain's instrument here equals

$$A = \frac{1}{\alpha_1\alpha_2\gamma_1\gamma_2}\, \frac{\left(\frac{\varphi'_{11}}{\varphi_{11}} x_{11} + \frac{\varphi'_{22}}{\varphi_{22}} x_{22}\right) - \left(\frac{\varphi'_{12}}{\varphi_{12}} x_{12} + \frac{\varphi'_{21}}{\varphi_{21}} x_{21}\right)}{\sigma^2_{11} + \sigma^2_{22} + \sigma^2_{11}\sigma^2_{22} + \sigma^2_{12} + \sigma^2_{21} + \sigma^2_{12}\sigma^2_{21}},$$

which is a function of the fixed effects.

In an exponential-regression model, $\varphi'_{ij}/\varphi_{ij} = 1$ and the numerator of A simplifies to

$$(x_{11} - x_{12}) - (x_{21} - x_{22}). \qquad \text{(h)}$$

When errors are homoskedastic with variance σ^2, the denominator in the second term of A simplifies to $2\sigma^2(2 + \sigma^2)$. When data is Poisson distributed, this term has the more complicated form

$$\frac{1+\varphi_{11}\alpha_1\gamma_1+\varphi_{22}\alpha_2\gamma_2}{\varphi_{11}\varphi_{22}}+\frac{1+\varphi_{12}\alpha_1\gamma_2+\varphi_{21}\alpha_2\gamma_1}{\varphi_{12}\varphi_{21}}.$$

The full expression for A in the Poisson case then becomes

(i) $$A=\frac{(x_{11}-x_{12})-(x_{21}-x_{22})\varphi_{11}\varphi_{22}\varphi_{12}\varphi_{21}}{Q},$$

where

$$\begin{aligned}Q=&\ \varphi_{12}\varphi_{21}+\varphi_{11}\varphi_{12}\varphi_{21}\alpha_1\gamma_1+\varphi_{12}\varphi_{21}\varphi_{22}\alpha_2\gamma_2\\&+\varphi_{11}\varphi_{22}+\varphi_{11}\varphi_{12}\varphi_{22}\alpha_1\gamma_2+\varphi_{11}\varphi_{21}\varphi_{22}\alpha_2\gamma_1.\end{aligned}$$

This depends on the fixed effects.

In the one-way Poisson model for short panel data (see, e.g., Wooldridge 1999), where outcomes y_{ij} (given x_{ij} and α_i) are Poisson distributed with conditional mean

$$e^{x'_{ij}\psi}\alpha_i,$$

the maximum-likelihood estimator uses the optimal instrument, which does not depend on the α_i (Hahn 1997a,b). This is so because the α_i cancel out in the numerator and denominator of (the corresponding version of) A. This is no longer true in the two-way model considered here. The cause is that, while in the one-way model the moment condition is based on differencing of first moments, in the two-way model here, we difference second-order moments. There might be a connection between this finding and the fact that the Poisson first-order conditions are not unbiased (see Charbonneau 2013 for an example of this later fact).

III. Computationally-efficient moment evaluation

The empirical moment vector of the GMM estimator is a fourth-order U-statistic. Brute-force evaluation of $s(\psi)$ requires $O(n^4)$ operations. Consequently, for large n, the computational cost of such an approach may be prohibitive. Fortunately, careful inspection of $s(\psi)$ shows that it can be computed much more efficiently; in some cases even without looping over any of the indices, making calculation extremely fast. Here, we illustrate this for the estimators GMM1 and GMM2 (as defined in the main text).

Recall that the empirical moment for GMM1 is

$$s(\psi) = \varrho^{-1} \sum_{i=1}^{n} \sum_{i<i'} \sum_{j=1}^{n} \sum_{j<j'} (x_{ij} - x_{ij'}) - (x_{i'j} - x_{i'j'})(u_{ij}(\psi)\, u_{i'j'}(\psi) - u_{ij'}(\psi)\, u_{i'j}(\psi)).$$

First exploit symmetry across both i and i' and across j and j', and note that terms for which either $i = i'$ or $j = j'$ are zero for any ψ, to see that

$$\text{(j)} \qquad s(\psi) = \varrho^{-1} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{n} \sum_{j'=1}^{n} x_{ij}\,(u_{ij}(\psi)\, u_{i'j'}(\psi) - u_{ij'}(\psi)\, u_{i'j}(\psi)).$$

By evaluating each of the terms of the summand in (j), we obtain

$$\text{(k)} \qquad s(\psi) = \varrho^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}\ \left(u_{ij}(\psi)\, \overline{u}(\psi) - \overline{u}_{i\cdot}(\psi)\, \overline{u}_{\cdot j}(\psi)\right),$$

where we define

$$\overline{u}_{i\cdot}(\psi) = \sum_{j=1}^{n} u_{ij}(\psi), \qquad \overline{u}_{\cdot j}(\psi) = \sum_{i=1}^{n} u_{ij}(\psi), \qquad \overline{u}(\psi) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij}(\psi).$$

Computation of $s(\psi)$ using (k) is immediate in any matrix-based language. A computationally-efficient representation of the Jacobian follows readily in the same way.

Proceeding similarly for GMM2 yields

$$
\begin{aligned}
s(\psi) &= \varrho^{-1} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{n} \sum_{j'=1}^{n} x_{ij} \left(y_{ij} y_{i'j'} \, \varphi_{ij'}(\psi) \varphi_{i'j}(\psi) - y_{i'j} y_{ij'} \, \varphi_{ij}(\psi) \varphi_{i'j'}(\psi) \right) \\
&= \varrho^{-1} \sum_{i=1}^{n} \sum_{j'=1}^{n} x_{ij} \left(y_{ij} \, a_{ij}(\psi) - \varphi_{ij}(\psi) \, b_{ij}(\psi) \right),
\end{aligned}
$$

for

$$
a_{ij}(\psi) = \sum_{i'=1}^{n} \sum_{j'=1}^{n} y_{i'j'} \, \varphi_{i'j}(\psi) \, \varphi_{ij'}(\psi), \qquad b_{ij}(\psi) = \sum_{i'=1}^{n} \sum_{j'=1}^{n} \varphi_{i'j'}(\psi) \, y_{i'j} \, y_{ij'}.
$$

Matlab routines that perform rapid calculation of both GMM1 and GMM2 (as well as the associated standard errors, as defined in the main text) are available as supplementary material.

References

Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Econometrica 34*, 305–334.

Charbonneau, K. B. (2013). *Multiple fixed effects in theoretical and applied econometrics*. Ph. D. thesis, Princeton University.

Hahn, J. (1997a). Efficient estimation of panel data models with sequential moment restrictions. *Journal of Econometrics 79*, 1–21.

Hahn, J. (1997b). A note on the efficient semiparametric estimation of some exponential panel models. *Econometric Theory 13*, 583–588.

Newey, W. K. and D. L. McFadden (1994). Large sample estimation and hypothesis testing. In R. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, Chapter 36, pp. 2111–2245. Elsevier.

Prakasa Rao, B. L. S. (2009). Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics 61*, 441–460.

van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics 90*, 77–97.