ECONOMIC RESEARCH REPORTS

A UNIFORM LAW OF LARGE NUMBERS FOR DEPENDENT AND HETEROGENEOUS DATA PROCESSES

by

Benedikt M. Potscher

and

Ingmar R. Prucha

R.R. #87-26

July 1987

C. V. STARR CENTER
FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

A UNIFORM LAW OF LARGE NUMBERS FOR DEPENDENT AND HETEROGENEOUS DATA PROCESSES

Benedikt M. Pötscher* and Ingmar R. Prucha**

April 1987

---

* Department of Econometrics, University of Technology Vienna, 1040 Vienna, Austria; and Department of Statistics, Yale University, New Haven, CT 06520.

** Department of Economics, University of Maryland, College Park, MD 20742; National Bureau of Economic Research, Cambridge, MA 02138; and C.V. Starr Center for Applied Economics, New York University, New York, NY 10003.

Abstract

Uniform laws of large numbers constitute important tools in proving consistency of estimators in nonlinear econometric models. A uniform law of large numbers used widely in the recent econometric literature is that of Hoadly (1971). It turns out that the assumptions maintained by this theorem (or some modified versions of it) are severe in that they preclude the analysis of many estimators and models of interest in economics. This paper introduces an alternative uniform law of large numbers for dependent and heterogeneous data processes that is based on a quite general and easy-to-verify catalog of assumptions.

1. Introduction[1]

Uniform laws of large numbers (ULLNs) constitute important tools in proving consistency of estimators in nonlinear econometric models. ULLNs consider sums of the form: $n^{-1}\Sigma_{t=1}^{n}[q_t(z_t,\theta)-Eq_t(z_t,\theta)]$, where (z_t) with $z_t \in Z$ denotes a stochastic data generating process, θ is an element of the parameter space Θ, and $q_t: Z \times \Theta \longrightarrow R$. ULLNs provide conditions under which the above sum converges to zero uniformly over the parameter space.

Hoadley (1971) introduced a ULLN that allows for non-i.i.d. data processes. This ULLN (or some modified version of it) has been used widely in the recent econometric literature, see e.g. Bates and White (1985), Domowitz and White (1982), Levine (1983), White (1980) and White and Domowitz (1984). However, Andrews (1986) and Pötscher and Prucha (1986a,b) point out that the uniform continuity assumption maintained by this ULLN is severe. It precludes the analysis of many estimators and models of interest in economics.[2] This suggests the need for alternative ULLNs.

The purpose of the present note is to introduce an alternative generic ULLN. The ULLN is generic in the sense that it transforms pointwise laws of large numbers into corresponding uniform ones. It maintains a set of assumptions that is relatively easy to verify and allows at the same time the

---

[1] We would like to thank Charles Bates, Manfred Deistler, Ian Domowitz, Ronald Gallant, Lars Peter Hansen, Nanhua Hu, Harry Kelejian, Whitney Newey, Ching-Zong Wei, Halbert White, and Ernest Zampelli for helpful comments. We assume responsibility, however, for any errors. The research was supported by the Max Kade Foundation and by the C.V. Starr Center's Focus Program for Capital Formation, Technological Change, Financial Structure and Tax Policy.

[2] We note that in the above papers by Bates, Domowitz, Levine and White the proofs of theorems regarding consistency are such that Hoadley's ULLN can be replaced by some alternative ULLN. I.e., the theorems can be readily rectified and/or restored to their intended generality by use of an alternative ULLN accompanied by a corresponding change in the catalogs of assumptions. For a detailed discussion of this issue see, e.g., Pötscher and Prucha (1986b).

analysis of a wide variety of estimators and models of interest in economics. The ULLN allows for temporal dependence and heterogeneity of the data process (z_t) as well as for heterogeneity in the functions q_t.[3] In addition to giving conditions under which $n^{-1}\Sigma_{t=1}^{n}[q_t(z_t,\theta)-Eq_t(z_t,\theta)]$ converges to zero uniformly, we also give additional conditions which guarantee that $n^{-1}\Sigma_{t=1}^{n}q_t(z_t,\theta)$ converges uniformly to a finite limit. This latter result is essentially a generic generalization of ULLNs given in Bierens (1981,1984) and Pötscher and Prucha (1986a).

In a recent paper, Andrews (1986) introduces an alternative generic ULLN based on the assumption that $q_t(z,\theta)$ satisfies a Lipshitz-type smoothness condition with respect to θ.[4] In contrast, the generic ULLN presented in this paper assumes that $q_t(z,\theta)$ satisfies a continuity-type condition jointly in z and θ. Therefore, the two ULLNs complement each other. Neither set of assumptions is weaker than the other. In essence, Andrew's ULLN imposes a stronger smoothness condition on the parameters, while the ULLN introduced here imposes a stronger smoothness condition on the data. As an illustration of the difference between the two ULLNs, we note that the former ULLN does not, e.g., cover all functions $q_t(z,\theta) \equiv q(z,\theta)$ where $q(z,\theta)$ is jointly continuous, whereas the latter ULLN does. We also note that our continuity-type smoothness condition should be easy to verify.

---

[3] In Pötscher and Prucha (1986b), we derive also a more general generic ULLN than that presented here, at the expense of a more complex catalog of assumptions and a more complex proof.

[4] The proofs of most ULLNs, including those mentioned above, are based on the finite approximation technique dating back to Wald (1949). This technique reduces the proof essentially to a verification of a single condition, sometimes referred to as the first moment continuity condition; compare e.g. Amemiya (1985), Hansen (1982), Newey (1987). In Andrews (1986), this condition is first formulated as an assumption; it is then shown that the Lipshitz-type smoothness condition implies this assumption.

2. A Uniform Law of Large Numbers

Let $(\Omega, \mathcal{A}, P)$ be a probability space and (Θ, ρ) a (non-empty) metric space with metric ρ.[5] Let $(z_t: t \in \mathbf{N})$ be a stochastic process defined on $(\Omega, \mathcal{A}, P)$ with values in a (non-empty) measurable space $(Z, \mathcal{Z})$. For ease of exposition we first take $(Z, \mathcal{Z})$ to be $(\mathbf{R}^s, \mathcal{B}^s)$; more general choices for Z will be discussed later. Our generic ULLN is based on the following assumptions:

Assumption 1: Θ is compact.

Assumption 2 (Smoothness condition): $q_t(z,\theta) = \Sigma_{k=1}^{K} r_{kt}(z) s_{kt}(z,\theta)$ where the r_{kt} are $\mathcal{Z}$-measurable real functions for all $t \in \mathbf{N}$ and $1 \leq k \leq K$, and the family $(s_{kt}(z,\theta): t \in \mathbf{N})$ of real functions is equicontinuous on $Z \times \Theta$ for all $1 \leq k \leq K$, i.e. $\sup_{t \in \mathbf{N}} |s_{kt}(z,\theta) - s_{kt}(z^0,\theta^0)| \longrightarrow 0$ for $(z,\theta) \longrightarrow (z^0,\theta^0)$ for all $(z^0,\theta^0) \in Z \times \Theta$ and each $1 \leq k \leq K$.[6]

Assumption 2 allows for discontinuities in the data z, given those discontinuities can be "separated" from the parameters θ. (As a consequence, this assumption covers, e.g., the case of the Tobit log-likelihood function where the r_{kt} represent indicator functions.) Assumption 2 is weaker than the assumption that the family $q_t(z,\theta)$ is equicontinuous. The latter assumption is contained as a special case and corresponds to, say, $K = 1$ and $r_{1t} \equiv 1$. However, as Newey (1987) points out, equicontinuity can often be obtained

---

[5] None of the subsequent results and conditions depend on the metric structure of (Θ, ρ), they only depend on the topology induced on Θ by the metric ρ. The choice of a fixed metric is only made for convenience.

[6] Equicontinuity on $Z \times \Theta$ is, of course, defined w.r.t. the product topology on $Z \times \Theta$, where $Z = \mathbf{R}^s$ carries the standard Euclidean topology.

through a suitable redefinition of the data.[7]

We introduce the following notation: $D_t(z) = \sup\{|q_t(z,\theta)| : \theta\in\Theta\}$ and $q_t^*(z,\theta,\tau) = \sup\{q_t(z,\theta'): \rho(\theta,\theta')<\tau\}$, $q_{t*}(z,\theta,\tau) = \inf\{q_t(z,\theta'): \rho(\theta,\theta')<\tau\}$. Under Assumptions 1 and 2 the functions $D_t(z)$, $q_t^*(z,\theta,\tau)$ and $q_{t*}(z,\theta,\tau)$ are real valued and Z-measurable. (Note that Θ is separable and $q_t(z,\theta)$ is continuous in θ).

Assumption 3 (Dominance condition): (i) $\sup_n n^{-1}\Sigma_{t=1}^n E[D_t(z_t)]^{1+\delta} < \infty$ for some $\delta > 0$. (ii) $\sup_n n^{-1}\Sigma_{t=1}^n E|r_{kt}(z_t)| < \infty$ for $1\le k\le K$.

If the r_{kt} represent, e.g., indicator functions Assumption 3(ii) is trivially satisfied. More primitive conditions implying Assumption 3 are $\sup_n n^{-1}\Sigma_{t=1}^n E|r_{kt}(z_t)|^{p(1+\delta)} < \infty$ and $\sup_n n^{-1}\Sigma_{t=1}^n E(\sup_{\theta\in\Theta}|s_{kt}(z_t,\theta)|^{q(1+\delta)}) < \infty$ for some $\delta>0$ and some $p>1$, $p^{-1}+q^{-1}=1$, and $1\le k\le K$.

Assumption 4 (Pointwise law of large numbers): For all $\theta\in\Theta$ there exists a sequence of positive numbers $\tau_i = \tau_i(\theta)$, $\tau_i \longrightarrow 0$, such that for each τ_i the random variables $q_t^*(z_t,\theta,\tau_i)$ and $q_{t*}(z_t,\theta,\tau_i)$ satisfy a strong law of large numbers, i.e. as $n \longrightarrow \infty$

$$n^{-1}\Sigma_{t=1}^n[q_t^*(z_t,\theta,\tau_i) - Eq_t^*(z_t,\theta,\tau_i)] \longrightarrow 0, \quad \text{P-a.s.},$$

$$n^{-1}\Sigma_{t=1}^n[q_{t*}(z_t,\theta,\tau_i) - Eq_{t*}(z_t,\theta,\tau_i)] \longrightarrow 0, \quad \text{P-a.s.}.$$

Clearly Assumption 4 is implied by the following stronger condition: Each sequence of random variables $\{f_t(z_t)\}$ with $|f_t(z_t)| \le D_t(z_t)$ satisfies a strong law of large numbers. We need, furthermore, either one of the

---

[7] For example, any function q_t as specified in Assumption 2 can be written as an equicontinuous function in new variables if the s_{kt} satisfy an additional condition: Define $\bar{z}_t = (w_{1t},\ldots,w_{Kt},z_t)$ with $w_{kt} = r_{kt}(z_t)$; then $\bar{q}_t(\bar{z}_t,\theta) = q_t(z_t,\theta)$ with $\bar{q}_t(\bar{z},\theta) = \Sigma_{k=1}^K w_k s_{kt}(z,\theta)$ and $\bar{z} = (w_1,\ldots w_K,z)$. Then $\bar{q}_t(\bar{z},\theta)$ is equicontinuous if the s_{kt} are equicontinuous and $\sup_t|s_{kt}(z,\theta)|<\infty$.

following two assumptions:

Assumption 5A: There exists an increasing function $h: [0,\infty) \longrightarrow [0,\infty)$ with $h(x) \longrightarrow \infty$ as $x \longrightarrow \infty$ such that $\sup_n n^{-1}\Sigma_{t=1}^n Eh(\|z_t\|) < \infty$.[8]

Assumption 5B: The process (z_t) is asymptotically stationary in the following sense: Let H_t denote the distribution of z_t on Z, then $H^n = n^{-1}\Sigma_{t=1}^n H_t$ converges weakly to some probability measure H on Z.

Assumption 5A is a rather mild moment condition. Assumption 5B is clearly satisfied for any identically distributed process. Given the above catalog of assumptions, we have the following ULLN:

Theorem 1: If Assumptions 1-4,5A or 1-4,5B hold, then $Eq_t(z_t,\theta)$ exists and is continuous on Θ, and

(a) $$\lim_{n\to\infty} \sup_{\theta\in\Theta} \left|n^{-1}\Sigma_{t=1}^n [q_t(z_t,\theta) - Eq_t(z_t,\theta)]\right| = 0 \quad \text{P-a.s.},$$

(b) $$\left\{n^{-1}\Sigma_{t=1}^n Eq_t(z_t,\theta) \ : \ n\in N\right\} \text{ is equicontinuous on } \Theta.$$

All proofs are given in the Appendix. Theorem 1 may be viewed as a generic ULLN in the sense that it postulates in Assumption 4 the existence of a pointwise law of large numbers. Specific ULLNs can now be obtained simply by verifying Assumption 4 from more primitive conditions on the process (z_t) as, e.g., α-mixing or ϕ-mixing conditions. Those mixing conditions have been frequently used in the recent econometric literature. The following corollary corresponds to these mixing conditions. (For a definition of ϕ-mixing and α-mixing see e.g. McLeish (1975).)

Corollary: Given Assumptions 1,2,3(ii),5A or 1,2,3(ii),5B hold. If,

---

[8] Examples for h are: $h(x) = x^p$, $p>0$, or $h(x) = \ln(1+x)$.

furthermore the process (z_t) is ϕ-mixing of size $-r/(2r-1)$ with $r\geq 1$ [or α-mixing of size $-r/(r-1)$ with $r>1$], and $\sup_t E[D_t(z_t)]^{r+\delta} < \infty$ for some $\delta > 0$, then the conclusions of Theorem 1 hold.

Theorem 1 only ensures that $n^{-1}\Sigma_{t=1}^{n}[q_t(z_t,\theta)-Eq_t(z_t,\theta)]$ converges to zero uniformly on Θ. In certain cases a stronger result, namely that also $n^{-1}\Sigma_{t=1}^{n}q_t(z_t,\theta)$ converges to some finite limit uniformly on Θ, is useful. Clearly, in order to obtain this stronger result we need some kind of asymptotic stationarity of the process. The following generic ULLN generalizes results in Bierens (1981,1984) and Pötscher and Prucha (1986a).

<u>Theorem</u> <u>2</u>: Given Assumptions 1,3(i),4,5B hold and $q_t(z,\theta) \equiv q(z,\theta)$ is continuous on $Z\times\Theta$, then the conclusions of Theorem 1 hold. Furthermore $\int q(z,\theta)dH(z)$ exists and is continuous on Θ, and

$$\lim_{n\to\infty}\ \sup_{\theta\in\Theta}\ |n^{-1}\Sigma_{t=1}^{n}q_t(z_t,\theta) - \int q(z,\theta)dH(z)| = 0 \quad \text{P-a.s..}$$

Specific ULLNs can again be obtained from Theorem 2 if Assumption 4 is replaced by a suitable "mixing" condition.

<u>REMARK</u> <u>1</u>: An inspection of the proofs shows that Theorem 1 and the Corollary remain valid if Assumptions 5A and 5B are replaced by:

<u>Assumption</u> <u>5</u>: $\lim_{m\to\infty} \sup_n n^{-1}\Sigma_{t=1}^{n}P(z_t\notin K_m) = 0$ for some sequence of compact sets $K_m \subseteq Z$, i.e. the sequence $\{H^n\colon n\in N\}$ with $H^n = n^{-1}\Sigma_{t=1}^{n}H_t$ is tight.

In fact, the structure of the proofs is such that Assumptions 5A or 5B are used to verify that Assumption 5 holds for suitable sets K_m.

<u>REMARK</u> <u>2</u>:[9] So far we have assumed that the functions $q_t(z,\theta)$ are defined on $Z\times\Theta$ with $Z = \mathbf{R}^s$. Now assume that Z is a subset of $\mathbf{R}^s$ (equipped with the

---

[9] For a proof of the subsequent results see the Appendix.

induced Euclidean topology and the corresponding Borel σ-field).

(a) If Z is a closed subset of $\mathbf{R}^s$ then Theorem 1 and 2, the Corollary and Remark 1 remain valid without modification. (Clearly, all assumptions are now to be interpreted to hold on Z rather than on $\mathbf{R}^s$.) Moreover for Z compact Assumption 5 and 5A are automatically satisfied. If Z is an open subset of $\mathbf{R}^s$ then the parts of Theorem 1 and the Corollary corresponding to Assumption 5B, as well as Theorem 2 and Remark 1 remain valid without modification.

(b) If Z is an arbitrary subset of $\mathbf{R}^s$ then Theorem 1 and the Corollary still hold given Assumption 5A (5B) is replaced by either Assumption 5 or 5A′ (5B′) given below, and Theorem 2 remains valid if Assumption 5B is replaced by Assumption 5B′.[10]

<u>Assumption</u> <u>5A′</u>: There exists a sequence of compact sets $K_m \subseteq Z$, a $\mathcal{Z}$-measurable function $g: Z \longrightarrow [0,\infty]$ and real numbers $r_m \longrightarrow \infty$ such that:

(i) $\{z \in Z: g(z) < r_m\} \subseteq K_m$ and (ii) $\sup_n n^{-1}\Sigma_{t=1}^n Eg(z_t) < \infty$.

<u>Assumption</u> <u>5B′</u>: The process (z_t) is asymptotically stationary in the sense that the probability measures $H^n = n^{-1}\Sigma_{t=1}^n H_t$ converge weakly to some probability measure H on Z. Furthermore, H and each of the H^n are tight.[11]

The common feature of Assumptions 5, 5A, 5B, 5A′ and 5B′ is that they exclude situations where some mass of the average probability distribution H^n escapes a sequence of compact sets K_m in Z. This is achieved by Assumptions 5A, 5A′ by placing bounds on moments of certain functions of z_t, and by Assumptions 5B, 5B′ by requiring that the H^n converge to H (and all measures have σ-compact support.). Clearly, if Z is a closed subset of $\mathbf{R}^s$ then Assumption

---

[10] We note that this sentence actually remains valid for any metrizable space Z with corresponding Borel σ-field.

[11] I.e., H and each of the H^n have a σ-compact support. This is, e.g., trivially satisfied if Z is σ-compact or if Z is a separable and completely metrizable space, compare Theorem 1.4 in Billingsley (1968).

5A implies Assumption 5A' upon choosing $K_m = Z \cap \{z \in \mathbf{R}^s: \|z\| \leq m\}$, $g(z) = h(\|z\|)$ and $r_m = h(m)$. If Z is a closed or open subset of $\mathbf{R}^s$ then Assumption 5B' reduces to 5B by Theorem 1.4 in Billingsley (1968) and observing that a closed or open subset of $\mathbf{R}^s$ is separable and completely metrizable.

(c) If the functions $q_t(z,\theta)$ can be extended to $W \times \Theta$, $Z \subseteq W \subseteq \mathbf{R}^s$, such that Assumption 2 holds on W, we can substitute W for Z and perform the analysis on W rather than Z. This may be helpful if, e.g., the discussion in (a) applies to W but not to Z.

<u>REMARK</u> <u>3</u>: "Weak" ULLNs corresponding to the above "strong" ULLNs can be obtained by replacing in Assumption 4 almost sure convergence by convergence in probability.

Appendix

Lemma A1: Let $s_t(z,\theta)$ be a sequence of equicontinuous, real functions on $Z\times\Theta$ where Z is a metrizable space and (Θ,ρ) is a compact metric space; let $\emptyset\neq K\subseteq Z$ be compact. Then for each $\theta\in\Theta$ and each sequence $\tau_i\longrightarrow 0$, $\tau_i>0$, we have: $\sup_{t\in N}\sup_{z\in K}\sup_{\theta'\in B(\theta,\tau_i)}|s_t(z,\theta')-s_t(z,\theta)|\longrightarrow 0$ as $i\longrightarrow\infty$, where $B(\theta,\tau_i)=\{\theta'\in\Theta: \rho(\theta,\theta')<\tau_i\}$.

Proof: First note that the sup operators can be interchanged. Suppose the lemma is not true, then for some $\epsilon>0$ and some $\theta^0\in\Theta$ we can find a sequence $\theta^i\longrightarrow\theta^0$ for which $\sup_t\sup_{z\in K}|s_t(z,\theta^i)-s_t(z,\theta^0)|>\epsilon$. But then there exists for each $i\in N$ a point $z^i\in K$ such that $\sup_t|s_t(z^i,\theta^i)-s_t(z^i,\theta^0)|>\epsilon$. Since K is compact we can find a subsequence $z^{i(k)}$ converging to some $z^0\in K$. The last inequality then implies $\epsilon<\sup_t|s_t(z^{i(k)},\theta^{i(k)})-s_t(z^{i(k)},\theta^0)|\leq$ $\sup_t|s_t(z^{i(k)},\theta^{i(k)})-s_t(z^0,\theta^0)| + \sup_t|s_t(z^0,\theta^0)-s_t(z^{i(k)},\theta^0)|$. Since under equicontinuity both expressions on the r.h.s. converge to zero as $k\longrightarrow\infty$ this yields a contradiction. ∎

Proof of Theorem 1: Clearly, $Eq_t^*(z_t,\theta,\tau_i)$, $Eq_{t*}(z_t,\theta,\tau_i)$ and $Eq_t(z_t,\theta)$ exist by Assumption 3, the latter expectation is also continuous by Assumptions 2 and 3 and dominated convergence. Let K_m be any sequence of nonempty compact sets in $Z=\mathbf{R}^s$ and $B_i(\theta)=\{\theta'\in\Theta: \rho(\theta,\theta')<\tau_i\}$ with $\tau_i=\tau_i(\theta)$ as in Assumption 4. We first show that for each $\theta\in\Theta$

$$\text{(A.1)}\qquad \lim_{i\to\infty}\sup_n n^{-1}\Sigma_{t=1}^n E\sup\{|q_t(z_t,\theta')-q_t(z_t,\theta)|: \theta'\in B_i(\theta)\}=0.$$

Clearly, the l.h.s. of (A.1) is for $m\in N$ not greater than A_m+B_m where

$$A_m=\overline{\lim_{i\to\infty}}\sup_n n^{-1}\Sigma_{t=1}^n E[1_{K_m}(z_t)\sup\{|q_t(z_t,\theta')-q_t(z_t,\theta)|: \theta'\in B_i(\theta)\}]\text{ and}$$

$$B_m=\overline{\lim_{i\to\infty}}\sup_n n^{-1}\Sigma_{t=1}^n E[1_{Z-K_m}(z_t)\sup\{|q_t(z_t,\theta')-q_t(z_t,\theta)|: \theta'\in B_i(\theta)\}].$$

Using the definition of q_t it is readily seen that $A_m \le$ $\Sigma_{k=1}^{K} \overline{\lim}_{i\to\infty} \sup_n n^{-1}\Sigma_{t=1}^{n} E[1_{K_m}(z_t)|r_{kt}(z_t)| \sup\{|s_{kt}(z_t,\theta')-s_{kt}(z_t,\theta)| : \theta'\in B_i(\theta)\}] \le$ $\Sigma_{k=1}^{K} \left[\overline{\lim}_{i\to\infty} \sup_t \sup_{z\in K_m} \sup\{|s_{kt}(z,\theta')-s_{kt}(z,\theta)| : \theta'\in B_i(\theta)\}\right]\left[\sup_n n^{-1}\Sigma_{t=1}^{n} E(1_{K_m}(z_t)|r_{kt}(z_t)|)\right]$. The first expression in brackets on the r.h.s. of the last inequality is zero by Lemma A1, the second expression is bounded by Assumption 3(ii); hence $A_m = 0$. By Assumption 3(i) and the triangle inequality we see that $B_m \le 2\sup_n n^{-1}\Sigma_{t=1}^{n} E[1_{Z-K_m}(z_t)D_t(z_t)]$. Applying twice Hölder's inequality with $p=1+\delta$ and $q=1+\delta^{-1}$ we get further: $B_m \le 2\left[\sup_n n^{-1}\Sigma_{t=1}^{n} E(D_t(z_t))^p\right]^{1/p}\left[\sup_n n^{-1}\Sigma_{t=1}^{n} P(z_t\notin K_m)\right]^{1/q}$. Since the first term in brackets is finite by Assumption 3(i), B_m goes to zero as $m \to \infty$ if

(A.2) $$C_m = \sup_n n^{-1}\Sigma_{t=1}^{n} P(z_t\notin K_m) \to 0 \text{ as } m \to \infty.$$

To see that Assumption 5A implies (A.2) take $K_m=\{z\in\mathbf{R}^s : \|z\|\le m\}$. Then by Markov's inequality $C_m \le \sup_n n^{-1}\Sigma_{t=1}^{n} P(\|z_t\| \ge m) \le \sup_n n^{-1}\Sigma_{t=1}^{n} Eh(\|z_t\|)/h(m) \to 0$. To see that Assumption 5B implies (A.2) observe from Theorem 6.2 in Billingsley (1968) that the sequence H^n is tight, which is just (A.2) for some sequence of compact sets K_m. This proves (A.1), which implies part (b) of Theorem 1. Given (A.1) the proof of part (a) is now standard: Observe that (A.1) implies for $\epsilon > 0$ and all $\theta\in\Theta$ the existence of an index $i(\theta,\epsilon)$ such that for $i=i(\theta,\epsilon)$:

(A.3a) $$\sup_n |n^{-1}\Sigma_{t=1}^{n} E(q_t^*(z_t,\theta,\tau_i)-q_t(z_t,\theta))| < \epsilon ,$$

(A.3b) $$\sup_n |n^{-1}\Sigma_{t=1}^{n} E(q_{t*}(z_t,\theta,\tau_i)-q_t(z_t,\theta))| < \epsilon .$$

The family $\{B_{i(\theta,\epsilon)}(\theta) : \theta\in\Theta\}$ is an open cover of Θ, hence by Assumption 1 there exist finitely many $\theta(\nu)$, $1\le\nu\le L$, such that the corresponding balls $B(\nu) = B_{i(\theta(\nu),\epsilon)}(\theta(\nu))$ cover Θ. It then follows using (A.3) that for all $\theta\in\Theta$: $-2\epsilon + \min_\nu n^{-1}\Sigma_{t=1}^{n}[q_{t*}(z_t,\theta(\nu),\tau(\nu))-Eq_{t*}(z_t,\theta(\nu),\tau(\nu))] \le n^{-1}\Sigma_{t=1}^{n}[q_t(z_t,\theta)-Eq_t(z_t,\theta)] \le \max_\nu n^{-1}\Sigma_{t=1}^{n}[q_t^*(z_t,\theta(\nu),\tau(\nu))-Eq_t^*(z_t,\theta(\nu),\tau(\nu))] + 2\epsilon$ where $\tau(\nu) = \tau_{i(\theta(\nu),\epsilon)}(\theta(\nu))$. Part (a) of the theorem now follows from

Assumption 4 since ϵ is arbitrary. ■

Proof of Corollary 2: Assumptions 3(i) is directly implied by the Assumptions of the corollary. Assumption 4 is readily seen to follow from Theorem 2.10 of McLeish (1975): Observe that q_t^* and q_{t*} are Z-measurable; hence $q_t^*(z_t,\theta,\tau_i)-Eq_t^*(z_t,\theta,\tau_i)$ and $q_{t*}(z_t,\theta,\tau_i)-Eq_{t*}(z_t,\theta,\tau_i)$ are α-mixing (ϕ-mixing) processes with mixing coefficients no greater than those of (z_t). ■

Proof of Theorem 2: Clearly, all assumptions of Theorem 1 are satisfied (taking K = 1 and $r_{1t}\equiv 1$). It hence suffices to prove that $\int q(z,\theta)dH$ exists and that it is the uniform limit of $n^{-1}\Sigma_{t=1}^{n}Eq(z_t,\theta) = \int q(z,\theta)dH^n$. Assumption 3 implies that $\sup_n\int|q(z,\theta)|^{1+\delta}dH^n < \infty$ for all $\theta\in\Theta$. Since $H^n \longrightarrow H$ weakly it follows analogously as in Theorem 5.4 of Billinglsley (1968) that $\int q(z,\theta)dH$ exists and that $\int q(z,\theta)dH^n \longrightarrow \int q(z,\theta)dH$ for each $\theta\in\Theta$; part (b) of Theorem 1 implies that the family $\{\int q(z,\theta)dH^n\colon n\in\mathbf{N}\}$ is equicontinuous on Θ. The uniform convergence now follows from the theorem of Ascoli-Arzela. ■

We note that the proofs of Theorem 1, 2, and the Corollary do not utilize the structure of $Z = \mathbf{R}^s$ beyond the fact that it is a metrizable space - except in the verification of (A.2) via Assumptions 5A or 5B. Therefore the following lemma proofs the nontrivial portions of Remark 2(b). (Observe that Assumptions 5, 5A′ and 5B′ all imply that the K_m are nonempty for large m.) Remark 2(a) follows from Remark 2(b).

Lemma A2: Let Z be a metrizable space and Z its Borel σ-field. Then Assumption 5A′ as well as Assumption 5B′ imply Assumption 5.

Proof: For the sets K_m of Assumption 5A′ we have that $\sup_n n^{-1}\Sigma_{t=1}^{n}P(z_t\notin K_m) \le \sup_n n^{-1}\Sigma_{t=1}^{n}P(g(z_t)\ge r_m) \le \sup_n n^{-1}\Sigma_{t=1}^{n}Eg(z_t)/r_m$. This proves the first claim. Under Assumption 5B′, Theorem 8 in Appendix III of Billingsley (1968) implies that $\{H^n\colon n\in\mathbf{N}\}$ is tight. This proves the second claim. ■

References

Andrews, D.W.K. (1986): "Consistency in Nonlinear Econometric Models: A Generic Uniform Law of Large Numbers," Econometrica, forthcoming. (Cowles Foundation for Research in Economics, Yale University, Discussion Paper No. 790, revised August 1986).

Bates, C., and H. White (1985): "A Unified Theory of Consistent Estimation for Parametric Models," Econometric Theory, 1, 151-178.

Bierens, H.J. (1981): "Robust Methods and Asymptotic Theory in Nonlinear Econometrics," Lecture Notes in Economics and Mathematical Systems, Vol. 192. Berlin: Springer Verlag.

Bierens, H.J. (1984): "Model Specification Testing of Time Series Regression," Journal of Econometrics, 26, 323-353.

Billingsley, P. (1968): Convergence of Probability Measures. New York: Wiley.

Domowitz, I., and H. White (1982): "Misspecified Models with Dependent Observations," Journal of Econometrics, 20, 35-58.

Hoadley, B. (1971): "Asymptotic Properties of Maximum Likelihood Estimators for the Independent Not Identically Distributed Case," Annals of Mathematical Statistics, 42, 1977-1991.

Levine, D. (1983): "A Remark on Serial Correlation in Maximum Likelihood," Journal of Econometrics, 23, 337-342.

McLeish, D.L. (1975): "A Maximal Inequality and Dependent Strong Laws," Annals of Probability, 3, 829-839.

Newey, W.K. (1987): "Expository Notes on Uniform Laws of Large Numbers," Princeton University, mimeo.

Pötscher, B.M., and I.R. Prucha (1986a): "A Class of Partially Adaptive One-Step M-Estimators for the Nonlinear Regression Model with Dependent Observations," Journal of Econometrics, 32, 219-251.

Pötscher, B.M., and I.R. Prucha (1986b): "Consistency in Nonlinear Econometrics: A Generic Uniform Law of Large Numbers and Some Comments on Recent Results," Department of Economics, University of Maryland, Working Paper No. 86-9, June 1986.

Wald, A. (1949): "Note on the Consistency of the Maximum Likelihood Estimate," Annals of Mathematical Statistics, 20, 595-600.

White, H. (1980): "Nonlinear Regression on Cross-Section Data," Econometrica, 48, 721-746.

White, H., and I. Domowitz (1984): "Nonlinear Regression with Dependent Observations," Econometrica, 52, 143-161.