Statistics 25 (1994) 343–360
Reprints available directly from the publisher
Photocopying permitted by license only

ON THE FORMULATION OF UNIFORM LAWS OF LARGE NUMBERS: A TRUNCATION APPROACH

BENEDIKT M. PÖTSCHER and INGMAR R. PRUCHA

University of Vienna and University of Maryland, College Park

(Received 3 February 1992; in final form 10 March 1993)

Summary. The paper develops a general framework for the formulation of generic uniform laws of large numbers. In particular, we introduce a basic generic uniform law of large numbers that contains recent uniform laws of large numbers by Andrews [2] and Hoadley [9] as special cases. We also develop a truncation approach that makes it possible to obtain uniform laws of large numbers for the functions under consideration from uniform laws of large numbers for truncated versions of those functions. The point of the truncation approach is that uniform laws of large numbers for the truncated versions are typically easier to obtain. By combining the basic uniform law of large numbers and the truncation approach we also derive generalizations of recent uniform laws of large numbers introduced in Pötscher and Prucha [15, 16].

Key words: Uniform laws of large numbers, dependent processes.

1. INTRODUCTION[1]

Uniform laws of large numbers (ULLNs) are important tools for developing asymptotic theory in econometrics and statistics. For example, consistency proofs of estimators in nonlinear econometric models frequently employ ULLNs. In the early seventieths Hoadley [9] introduced a ULLN that applies to non-i.i.d. data processes. This ULLN (or some version of it) has been used widely in the econometrics literature, see, e.g., Bates and White [3], Domowitz and White [7], Levine [11], White [19] and White and Domowitz [20]. However, Andrews [2] and Pötscher and Prucha [14,15] point out that the equicontinuity assumption maintained by Hoadley's ULLN is severe and precludes the analysis of many estimators and models of interest.[2] These observations motivate interest in alternative ULLNs.

Consider the sum $n^{-1}\Sigma_{t=1}^{n}[q_t(z_t,\theta) - \mathrm{E}q_t(z_t,\theta)]$, where (z_t) denotes a stochastic process that takes its values in a set Z, θ is an element of a parameter space Θ, and q_t: $Z \times \Theta \rightarrow \mathbb{R}$. ULLNs then provide conditions under which the above sum converges to zero uniformly over the parameter space. Ideally, ULLNs should be applicable to a wide range of problems. They should be able to handle temporal dependence and heterogeneity of the stochastic process (z_t) as well as heterogeneity in the functions q_t. ULLNs that are aimed towards that goal have been introduced recently, e.g., by Andrews [2] and Pötscher and Prucha [15,16]. Those ULLNs are generic in the sense that they transform laws of large numbers (LLNs) for certain bracketing functions for $q_t(z_t,\theta)$ into corresponding uniform ones.

The proofs of most ULLNs, including those mentioned above, are based on an approximation technique that dates back to Wald [18], see also Jennrich [10]. This technique reduces the proof essentially to the verification of a single condition and also underlies the ULLNs in Amemiya [1], Bierens [4,5], Hansen [8], and Newey [13]. Several authors refer to that condition (formulated within contexts of varying generality) as the first moment continuity condition. Therefore, ULLNs that are based on Wald's approximation technique differ in essence only in the way how this first moment continuity condition is verified from basic assumptions on the functions $q_t(z_t, \theta)$.[3] For example, Andrews [2] imposes a Lipshitz-type smoothness condition which enables him to verify the first moment continuity condition by essentially separating the stochastic component from the parameters. Alternatively, Hoadley [9] imposes an almost sure equicontinuity condition, which—together with a standard domination condition—implies the first moment continuity condition. We note that it is precisely this almost sure equicontinuity condition that is restrictive in general.

The present paper is a revision and extension of Pötscher and Prucha [15]. In the latter paper we introduced a generic ULLN by (i) truncating the (argument z_t of the) functions $q_t(z_t, \theta)$ appropriately, (ii) by applying Hoadley's almost sure equicontinuity condition to the truncated functions to obtain a ULLN for the latter functions, and (iii) by recovering a ULLN for the untruncated functions from the ULLN for the truncated functions. The crucial point is that the almost sure equicontinuity condition is not restrictive when applied to the truncated functions. We note further that this truncation approach was also used implicitly in the proof of a recent ULLN given in Pötscher and Prucha [16].

In Section 2 of the present paper we first show that this truncation approach can be applied to any ULLN (and not only to Hoadley's ULLN). More specifically, we introduce a lemma that gives sufficient conditions under which a ULLN for the untruncated functions can be recovered from a ULLN for the truncated functions. Second, we derive a basic generic ULLN that unifies Andrews' [2] separation approach and Hoadley's [9] equicontinuity approach. We note that combining the lemma and this basic ULLN gives a general ULLN that contains a generalization of Andrews' [2] ULLN as well as Pötscher and Prucha's [15,16] ULLNs (and hence Hoadley's [9] ULLN) as special cases. In Section 3 we now derive Pötscher and Prucha's [15] generic ULLN as a special case and provide various simple sufficient conditions that imply the assumptions of this ULLN and also provide further interpretation to those assumptions. We also illustrate how specific ULLNs can be obtained from this generic ULLN by presenting a ULLN for α-mixing and ϕ-mixing processes. Various further points of discussion are provided in Section 4. In Section 5 we give an example for which it is possible to establish a ULLN under weaker conditions based on the results of this paper than is possible based on the results given in Andrews [2] and Pötscher and Prucha [16]. All proofs and some technical remarks are relegated to the appendices.

2. THE TRUNCATION APPROACH AND A BASIC UNIFORM LAW OF LARGE NUMBERS

2.1. *The truncation approach*

Let $(\Omega, \mathcal{A}, P)$ be a probability space, let $(Z, \mathcal{B})$ be a (non-empty) measurable space and (Θ, ρ) a (non-empty) metric space with metric ρ.[4] We refer to Θ as the

Remark 1.

(i) The sequence (K_m) will frequently (but not necessarily) be an increasing sequence of sets exhausting Z, i.e., $K_m \uparrow Z$.
(ii) Inspection of the proof shows that the condition of continuity of $n^{-1}\Sigma_{t=1}^{n} \mathrm{E}q_t(z_t, \theta)$ in part (b) of Lemma 1 can be dropped if part (a) of Assumption A is strengthened by replacing "limsup" with "sup".
(iii) Inspection of the proof shows further that part (b) of Assumption A is not needed for part (b) of Lemma 1.

2.2. *The basic uniform law of large numbers*

In preparation of our basic generic ULLN we introduce some further assumptions. Again, a discussion of those assumptions and, in particular, several sets of more primitive sufficient conditions will be given in the next section.

Assumption 1. Θ is compact.

We adopt the following notation: Let f_t: $Z \times \Theta \to \bar{\mathbb{R}}$ and $\tau > 0$, then

$$f_t^*(z, \theta', \tau) = \sup_{\rho(\theta, \theta') < \tau} f_t(z, \theta), \qquad f_{t^*}(z, \theta', \tau) = \inf_{\rho(\theta, \theta') < \tau} f_t(z, \theta).$$

Definition 2. The sequence $(f_t(z_t, \theta)\colon t \in \mathbb{N})$ is said to satisfy a strong LLN locally at θ', if there exists a sequence of positive numbers $(\tau_k)_{k \in \mathbb{N}}$, $\tau_k = \tau_k(\theta')$, with $\tau_k \to 0$ as $k \to \infty$, such that for each τ_k the sequences $(f_t^*(z_t, \theta', \tau_k)\colon t \in \mathbb{N})$ and $(f_{t^*}(z_t, \theta', \tau_k)\colon t \in \mathbb{N})$ satisfy a strong LLN, i.e.,

$$\lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} [f_t^*(z_t, \theta', \tau_k) - \mathrm{E} f_t^*(z_t, \theta', \tau_k)] = 0, \qquad \text{P-a.s.},$$

and

$$\lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} [f_{t^*}(z_t, \theta', \tau_k) - \mathrm{E} f_{t^*}(z_t, \theta', \tau_k)] = 0, \qquad \text{P-a.s.}.$$

We note that in the above definition it is implicitly understood that $f_t^*(z_t, \theta', \tau_k)$ and $f_{t^*}(z_t, \theta', \tau_k)$ are measurable P-a.s. and that the respective expectations exist and are finite.[7]

Assumption B. The sequence $(q_t(z_t, \theta)\colon t \in \mathbb{N})$ satisfies a strong LLN locally at θ' for all $\theta' \in \Theta$.

We note that Assumption B requires nothing else than strong LLNs to hold for the bracketing functions $q_t^*(z_t, \theta', \tau_k) = \sup_{\rho(\theta, \theta') < \tau} q_t(z_t, \theta)$ and $q_{t^*}(z_t, \theta', \tau) = \inf_{\rho(\theta, \theta') < \tau} q_t(z_t, \theta)$. Hence, this assumption is satisfied for many stochastic processes (z_t) like i.i.d. processes, stationary and ergodic processes, α-mixing processes, and ϕ-mixing processes given standard domination conditions.

Assumption C. For all $\theta' \in \Theta$ we have: Let $(\tau_k)_{k \in \mathbb{N}}$, $\tau_k = \tau_k(\theta')$, with $\tau_k \to 0$ as $k \to \infty$, be the sequence defined implicitly in Assumption B. Then there exists a p with $1 \le p \le \infty$ and a $k_0 = k_0(\theta')$ such that $\rho(\theta, \theta') < \tau_{k_0}$ implies

$$|q_t(z_t, \theta) - q_t(z_t, \theta')| \leq b_t(z_t, \theta')\, h_t(z_t, \theta', \theta)$$

for all $\omega \in \Omega \setminus N_{\theta'}$, $\mathsf{P}(N_{\theta'}) = 0$, where $b_t(z_t, \theta')$ and $h_t(z_t, \theta', \theta)$ are finite, non-negative and $\mathscr{A}$-measurable (for given θ' and all $\theta \in \Theta$) and satisfy the following:

(a) $\sup_t h_t(z_t, \theta', \theta) \to 0$ if $\rho(\theta, \theta') \to 0$ for all $\omega \in \Omega \setminus N_{\theta'}$.
(b) The functions $h_t^*(z_t, \theta', \tau_k) = \sup_{\rho(\theta, \theta') < \tau_k} h_t(z_t, \theta', \theta)$ are P-a.s. measurable for $\tau_k \leq \tau_{k_0}$.
(c) $\sup_n n^{-1}\Sigma_{t=1}^n \mathsf{E}[b_t(z_t, \theta')^p] < \infty$ if $1 \leq p < \infty$, and $\sup_t \|b_t(z_t, \theta')\|_\infty < \infty$ if $p = \infty$; furthermore, $\{n^{-1}\Sigma_{t=1}^n h_t^*(z_t, \theta', \tau_{k_0})^q : n \in \mathbb{N}\}$ is uniformly integrable if $1 < p \leq \infty$, and $\sup_t \|h_t^*(z_t, \theta', \tau_k)\|_\infty \to 0$ as $k \to \infty$ if $p = 1$, where $q^{-1} + p^{-1} = 1$.[8]

In the above $\|.\|_\infty$ denotes the essential supremum of $|.|$ w.r.t. P.

The above assumptions allow us now to introduce a basic generic ULLN. The idea behind the proof can be outlined as follows: Compactness of Θ postulated in Assumption 1 and the conditions in Assumption C are used to show that $n^{-1}\Sigma_{t=1}^n [q_t(z_t, \theta) - \mathsf{E}q_t(z_t, \theta)]$ can be approximated arbitrarily closely by finitely many functions of the form $n^{-1}\Sigma_{t=1}^n [q_t^*(z_t, \theta_i', \tau_{k_i}) - \mathsf{E}q_t^*(z_t, \theta_i', \tau_{k_i})]$ and $n^{-1}\Sigma_{t=1}^n [q_{t^*}(z_t, \theta_i', \tau_{k_i}) - \mathsf{E}q_{t^*}(z_t, \theta_i', \tau_{k_i})]$.[9] The strong LLNs for these functions postulated in Assumption B are then used to deliver the ULLN.

Theorem 1. *Let Assumptions 1, B and C be satisfied. Then:*

(*a*) *The sequence* $(q_t(z_t, \theta) : t \in \mathbb{N})$ *satisfies a ULLN.*
(*b*) $\{n^{-1}\Sigma_{t=1}^n \mathsf{E}q_t(z_t, \theta) : n \in \mathbb{N}\}$ *is equicontinuous on* Θ.

Remark 2. (i) If Assumption C holds for $h_t(z_t, \theta', \theta) = h(\rho(\theta, \theta'))$ with $h(x)\downarrow 0$ for $x \downarrow 0$ and if we put $p = 1$, then Theorem 1 reduces to a recently introduced ULLN by Andrews [2], which is based on a Lipshitz-type smoothness condition. If Assumption C holds for $p = \infty$, $h_t(z_t, \theta', \theta) = |q_t(z_t, \theta) - q_t(z_t, \theta')|$ and $b_t(z_t, \theta') \equiv 1$, then Theorem 1 reduces to a generic version of Hoadley's [9, Theorem A5] ULLN. (In this latter case Assumption C(a) becomes Hoadley's almost sure equicontinuity condition.) Hence, by applying Theorem 1 to the functions $q_{t,m}$ and then using Lemma 1 we can obtain further generalizations of both ULLNs. In Section 3 we show exemplarily how we can obtain a useful ULLN by generalizing the rather restrictive ULLN of Hoadley [9] via this truncation approach.

(ii) Also the ULLN given in Pötscher and Prucha [16] is essentially a special case of Theorem 1 combined with Lemma 1; for a more detailed discussion see comment (iv) in Section 4.

3. GENERALIZATION OF HOADLEY'S UNIFORM LAW OF LARGE NUMBERS AND A DISCUSSION OF SUFFICIENT CONDITIONS

It is important to recognize that, based on the approach developed in Section 2, we can obtain a very general ULLN by maintaining that Assumptions 1 and A hold, by postulating Assumptions B and C for the truncated functions $q_{t,m}$, and by appealing to Lemma 1 and Theorem 1. In the following we illustrate the approach further by considering ULLNs corresponding to sets of sufficient conditions for this general catalog of assumptions. In particular, we derive a generic ULLN that generalizes Hoadley's [9] ULLN. This generic ULLN was first presented in an earlier version of this paper, see Pötscher and Prucha [15]. In the following we first state the catalog of assumptions and present the theorem. We then give a discussion of the maintained assumptions. We will also provide sufficient conditions for the respective assumptions that can be readily checked. Based on those sufficient conditions we present two ULLNs for α-mixing and ϕ-mixing processes as corollaries.

Assumption 2. (a) For all $m \in \mathbb{N}$ and all $\theta' \in \Theta$ we have: $\{q_{t,m}(z_t, \theta): t \in \mathbb{N}\}$ is equicontinuous at θ', P-a.s..

(b) For all $t \in \mathbb{N}$ and $\theta' \in \Theta$ we have that $q_t(z_t, \theta)$ is continuous at θ', P-a.s..

A slightly stronger assumption (where the exceptional null set is now not allowed to depend on θ') is the following:

Assumption 2′. (a) For all $m \in \mathbb{N}$ we have: $\{q_{t,m}(z_t, \theta): t \in \mathbb{N}\}$ is equicontinuous on Θ, P-a.s..

(b) For all $t \in \mathbb{N}$ we have that $q_t(z_t, \theta)$ is continuous on Θ, P-a.s..

Assumption 3. The function $d_t(z_t)$ is P-a.s. measurable with $\mathrm{E}d_t(z_t)$ finite. Furthermore:

(a) $\lim_{m\to\infty} \{ \limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathrm{E}d_{t,m}^{c}(z_t)\} = 0.$

(b) The family of functions $\{n^{-1}\Sigma_{t=1}^{n} d_{t,m}(z_t): n \in \mathbb{N}\}$ is uniformly integrable, for each $m \in \mathbb{N}$.

Assumption 4. For all $m \in \mathbb{N}$:

(a) The sequence $(q_{t,m}(z_t, \theta): t \in \mathbb{N})$ satisfies a strong LLN locally at θ', for all $\theta' \in \Theta$ and all $m \in \mathbb{N}$.

(b) The sequence $(d_{t,m}^{c}(z_t): t \in \mathbb{N})$ satisfies a strong LLN for all $m \in \mathbb{N}$.[10]

We introduce the following generic ULLN.

Theorem 2. *Given Assumption* 1, *and given Assumptions* 2, 3, *and* 4 *are satisfied for some sequence* $(K_m)_{m \in \mathbb{N}}$ *with* $K_m \in \beta$. *Then*:

(a) The sequence $(q_t(z_t, \theta): t \in \mathbb{N})$ *satisfies a ULLN; and*
(b) $\{n^{-1}\Sigma_{t=1}^{n} \mathrm{E}q_t(z_t, \theta): n \in \mathbb{N}\}$ *is equicontinuous on* Θ.

Remark 3. (i) We note that for $K_m \equiv Z$ Theorem 2 reduces to a (slightly generalized) generic version of Hoadley's [9] ULLN. The equicontinuity condition in Hoadley [9] is identical to Assumption 2′ for $K_m \equiv Z$. As explained in more detail in Andrews [2] and Pötscher and Prucha [15], Hoadley's assumption is restrictive since it will, loosely speaking, be violated if there is unbounded variation in the variables z_t. The basic idea behind Theorem 2 is to avoid this restrictiveness by requesting only that this condition holds for certain approximations to the functions $q_t(z_t, \theta)$, i.e., for the functions $q_{t,m}(z_t, \theta)$. For simplicity of the argument assume for a moment that $Z = \mathbb{R}^S$. If the K_m are chosen as increasing bounded sets, then those approximations are obtained by truncating the functions $q_t(z_t, \theta)$ for "large" arguments z_t. We note that for these truncated functions the effective domain of the argument z_t is bounded. Consequently, and as demonstrated in more detail below, we find that the equicontinuity condition is not restrictive, when applied to the truncated functions.

(ii) We note that for the proof of part (a) of Theorem 2 part (b) of Assumption 2 has not been used. Also, if the sets K_m exhaust Z, i.e. , $\cup_{m=1}^{\infty} K_m = Z$, it is readily seen that part (b) of Assumption 2 is already implied by part (a). (The same is also true for Assumption 2′.)

(iii) Furthermore, if "limsup" in Assumption 3(a) is replaced by "sup", then Theorem 2 in its entirety also holds without postulating part (b) of Assumption 2, cp. also Remark 1(ii).

In the following remarks we discuss simple sufficient conditions for Assumptions 2, 3 and 4.

Remark 4. (Sufficient conditions for Assumption 2) The following discussion is based on Lemma A3 in Appendix A. As remarked above, Assumption 2 clearly holds if Assumption 2′ is satisfied. A stronger condition, that is often easier to verify, and implies Assumptions 2 and 2′, is the following:

(I) (a) $\sup_t \sup_{z \in K_m} |q_t(z, \theta) - q_t(z, \theta')| \to 0$ if $\rho(\theta, \theta') \to 0$, for all $\theta' \in \Theta$, $m \in \mathbb{N}$,

(b) $q_t(z, \theta)$ is continuous on Θ for each $z \in Z$ and $t \in \mathbb{N}$.

(Of course, if $K_m = \phi$ the supremum in I(a) is to be interpreted as zero.) If the K_m exhaust Z then part (b) of condition (I) is implied by part (a). If $Z = \mathbb{R}^s$ (or more generally, Z is a metrizable space), then condition (I) is implied by:

(II) $\{q_t(z, \theta): t \in \mathbb{N}\}$ is equicontinuous on $Z \times \Theta$, and the sets K_m are compact.

We note that in condition (II), in contrast to Hoadley's equicontinuity condition, the arguments of q_t do not depend on t. Hence this condition is (in a practical sense) far less restrictive, even so q_t is assumed to be jointly continuous in both arguments. For the important special case $q_t \equiv q$ condition (II) reduces to the condition that $q(z, \theta)$ is continuous on $Z \times \Theta$ and that the sets K_m are compact. As a point of interest, we note that both conditions (I) and (II) do not refer to properties of the stochastic process (z_t). Therefore, both conditions (I) and (II) imply that Assumptions 2 and 2′ are satisfied, regardless of the nature of the stochastic process (z_t).

Remark 5. (Sufficient conditions for Assumption 3) This discussion is based on Lemma A4 in Appendix A, and proceeds under the assumption that $d_t(z_t)$ is P-a.s. measurable.

(i) The following two conditions are sufficient for Assumption 3 to hold; they can be checked quite easily:

(III) $\sup_n n^{-1} \Sigma_{t=1}^{n} \mathrm{E}[d_t(z_t)^{1+\delta}] < \infty$ for some $\delta > 0$,

(IV) $\lim_{m \to \infty} \{ \limsup_{n \to \infty} n^{-1} \Sigma_{t=1}^{n} \mathrm{P}(z_t \notin K_m) \} = 0$.

parameter space. Let $(z_t)_{t\in\mathbb{N}}$ be a stochastic process defined on $(\Omega, \mathcal{A}, \mathsf{P})$ taking its values in Z. Let q_t: $Z \times \Theta \to \mathbb{R}$ be such that $q_t(z,\theta)$ is β-measurable for each $\theta \in \Theta$ and $t \in \mathbb{N}$.

It proves helpful to introduce the following definitions: For a sequence of sets $(K_m)_{m\in\mathbb{N}}$ with $K_m \in \beta$ let $q_{t,m}(z,\theta) = q_t(z,\theta)\mathbf{1}_{K_m}(z)$ and $q^c_{t,m}(z,\theta) = q_t(z,\theta)\mathbf{1}_{Z-K_m}(z)$, where $\mathbf{1}_A$ denotes the indicator function of a set A; also let $d_{t,m}(z) = \sup_{\theta\in\Theta}|q_{t,m}(z,\theta)|$, $d^c_{t,m}(z) = \sup_{\theta\in\Theta}|q^c_{t,m}(z,\theta)|$ and $d_t(z) = \sup_{\theta\in\Theta}|q_t(z,\theta)|$, which may take on the value $+\infty$. (Note that $d_{t,m}(z) = d_t(z)\mathbf{1}_{K_m}(z)$ and $d^c_{t,m}(z) = d_t(z)\mathbf{1}_{Z-K_m}(z)$ if we adopt the convention $\infty \cdot 0 = 0$.)

Definition 1. (a) The sequence $(q_t(z_t,\theta): t \in \mathbb{N})$ is said to satisfy a (strong) ULLN if and only if $\sup_{\theta\in\Theta}|n^{-1}\Sigma_{t=1}^n[q_t(z_t,\theta) - \mathsf{E}q_t(z_t,\theta)]| \to 0$, P-a.s., as $n \to \infty$.

(b) A sequence of random variables $(x_t: t \in \mathbb{N})$ is said to satisfy a strong LLN if and only if $n^{-1}\Sigma_{t=1}^n[x_t - \mathsf{E}x_t] \to 0$, P-a.s., as $n \to \infty$.

Note, that in this definition it is implicitly understood that the respective expectations exist and are finite.

We now introduce a basic lemma that makes it possible to generalize existing ULLNs via a "truncation principle". More specifically, the lemma gives basic conditions under which ULLNs for the sequences of truncated functions $(q_{t,m}(z_t,\theta)$: $t \in \mathbb{N})$ can be carried over into a ULLN for the sequence of untruncated functions $(q_t(z_t,\theta): t \in \mathbb{N})$. A discussion of these conditions and, in particular, several sets of more primitive sufficient conditions will be given in the next section. We emphasize that the point of the lemma is that it will be typically easier to verify the conditions of a given ULLN for the truncated functions then for the original functions. This fact will also be illustrated in more detail in the next section.

Assumption A. For a sequence of sets $(K_m)_{m\in\mathbb{N}}$ with $K_m \in \beta$ the corresponding functions $d^c_{t,m}(z_t)$ are P-a.s. measurable with $\mathsf{E}d^c_{t,m}(z_t)$ finite and furthermore:[5]

(a) $\lim_{m\to\infty}\{\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^n \mathsf{E}d^c_{t,m}(z_t)\} = 0$.[6]

(b) The sequences $(d^c_{t,m}(z_t): t \in \mathbb{N})$ satisfy a strong LLN for all $m \in \mathbb{N}$.

Lemma 1. *Let Assumption A be satisfied. Then:*

(a) The sequence $(q_t(z_t,\theta): t \in \mathbb{N})$ satisfies a ULLN if for all $m \in \mathbb{N}$ the sequences $(q_{t,m}(z_t,\theta): t \in \mathbb{N})$ satisfy a ULLN.

(b) $\{n^{-1}\Sigma_{t=1}^n \mathsf{E}q_t(z_t,\theta): n \in \mathbb{N}\}$ is equicontinuous on Θ, if $\{n^{-1}\Sigma_{t=1}^n \mathsf{E}q_{t,m}(z_t,\theta)$: $n \in \mathbb{N}\}$ is equicontinuous on Θ for each $m \in \mathbb{N}$ and if $n^{-1}\Sigma_{t=1}^n \mathsf{E}q_t(z_t,\theta)$ is continuous on Θ for each $n \in \mathbb{N}$.

In part (b) of the above lemma it is again implicitly understood that the expectations $\mathsf{E}q_{t,m}(z_t,\theta)$ exist and are finite; hence in view of Assumption A also the expectations $\mathsf{E}q_t(z_t,\theta)$ exist and are finite.

Condition (III) is a domination condition typical for ULLNs. Condition (IV) is a kind of asymptotic tightness condition for the average distribution of z_t, i.e. , for $n^{-1}\Sigma_{t=1}^{n} H_t$, where H_t is the distribution of z_t. Condition (IV) is, e.g., implied by either one of the following three simple conditions (Va), (Vb) or (Vc):

(Va) $Z = \mathbb{R}^s$, $K_m \uparrow \mathbb{R}^s$ is a sequence of Borel measurable convex sets (e.g., a sequence of closed or open balls), and $\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathsf{E}h(\|z_t\|) < \infty$ where h: $[0, \infty) \to [0, \infty)$ is a monotone function such that, $\lim_{x\to\infty} h(x) = \infty$.[11]

(Vb) $Z = \mathbb{R}^s$, $K_m \uparrow \mathbb{R}^s$ is a sequence of Borel measurable convex sets (e.g. a sequence of closed or open balls), and (z_t) is asymptotically stationary in the sense that $n^{-1}\Sigma_{t=1}^{n} H_t$ converges weakly to some probability measure H.

(Vc) (z_t) is identically distributed and $K_m \uparrow Z$.

More general formulations of (Va) and (Vb) are given after Lemma A4.

(ii) We note that Assumption 3 is also implied by, e.g.,

(VI) $\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathsf{E}d_t(z_t) = 0$, or

(VII) (z_t) is identically distributed, $K_m \uparrow Z$, $d_t \equiv d$ and $\mathsf{E}d(z_t) < \infty$.

Furthermore, Assumption 3(a) is trivially satisfied in the case $z_t \in K_0$, P-a.s., for all t and $K_m = K_0$ for all m.

Remark 6. (Sufficient conditions for Assumption 4) In this remark we show that Assumption 4 is satisfied if the underlying process has a sufficiently short memory. This is expressed by mixing properties. In particular we consider processes (z_t) that are α-mixing or ϕ-mixing. (For a definition of α-mixing and ϕ-mixing processes see, e.g., Domowitz and White [7]; for a definition of size see McLeish [12].) We introduce the following assumption:

Assumption 5. The process (z_t) is ϕ-mixing [α-mixing] with mixing coefficients of size $-r/(2r-1)$ where $r \geq 1$ [of size $-r/(r-1)$ where $r > 1$] and $\sup_t \mathsf{E}[d_t(z_t)^{r+\delta}] < \infty$ for some $\delta > 0$.

Of course it is implicitly assumed in Assumption 5 that $d_t(z_t)$ is P-a.s. measurable. In Lemma A5 in Appendix A we show that, given Assumption 1 and 2′(b), Assumption 5 implies Assumption 4.

From Remarks 4, 5, and 6 we can put together various sets of sufficient conditions that imply the assumptions of Theorem 2. We illustrate this by two corollaries:

Corollary 1. *Suppose Assumptions* 1, 2′ *and* 5 *hold. If furthermore* $\lim_{m\to\infty} \{\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathsf{P}(z_t \notin K_m)\} = 0$, *then the conclusions* (*a*) *and* (*b*) *of Theorem* 2 *hold.*

Corollary 2. *Suppose* $Z = \mathbb{R}^s$ *and Assumptions* 1 *and* 5 *hold. If* $q_t(z, \theta)$ *is equicontinuous on* $Z \times \Theta$, *and if* $\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathsf{E}(\|z_t\|^p) < \infty$ *for some* $p > 0$, *then conclusions* (*a*) *and* (*b*) *of Theorem* 2 *hold.*

Hoadley's [9] ULLN is a special case of Corollary 1 for $K_m \equiv Z$ (which, of course, implies that $\mathsf{P}(z_t \notin K_m) = 0$.) Corollary 2 is related to a ULLN reported in Pötscher and Prucha [16].

4. ADDITIONAL COMMENTS

(i) In situations where we wish to apply the above ULLNs to functions that depend on leads and/or lags we can think of the vector z_t as already containing those leads and/or lags. Cp. also Pötscher and Prucha [17], Section 5.

(ii) The assumption that Z does not depend on t can be made without loss of generality. Since no particular structure was postulated for Z in the general discussion (apart from being a measurable space) situations where the range space for the variables z_t depends on t can always be accommodated by defining Z as the Cartesian product of the respective range spaces and by redefining z_t and q_t in an obvious way.

(iii) Clearly any finite linear combination of functions individually satisfying a ULLN also satisfies a ULLN. One practical implication of this observation is the following: Although a function $q_t(z_t, \theta)$ may not satisfy the assumptions of an existing ULLN, it may be possible to decompose that function such that existing ULLNs can be applied to each of the summands in the decomposition. For example, there may exist functions $f_t(z_t)$ and $g_t(\theta)$ such that $f_t(z_t)$ satisfies a LLN and are such that $q_t(z_t, \theta) - f_t(z_t) - g_t(\theta)$ can be readily shown to satisfy a ULLN. Clearly, then also $q_t(z_t, \theta)$ satisfies an ULLN. As an illustration consider the case where $g_t(\theta) = \mathrm{E}q_t(z_t, \theta)$ is a trend increasing with t. Then the dominance condition may be violated for q_t but not for $q_t - \mathrm{E}q_t$. It may also prove useful in certain circumstances to subtract the value of q_t at a certain parameter θ_0, i.e. $f_t(z_t) = q_t(z_t, \theta_0)$, or a finite linear combination thereof. Of course one has then to prove a LLN for $q_t(z_t, \theta_0)$, but this is a much simpler task.

(iv) In Pötscher and Prucha [16] we prove a ULLN for functions of the form $q_t(z_t, \theta) = \Sigma_{k=1}^{K} r_{kt}(z_t)\, s_{kt}(z_t, \theta)$ where r_{kt} is measurable and s_{kt} jointly equicontinuous. This ULLN can be viewed as a special case of the results of Section 2 of the present paper (abstracting from the slight differences in the maintained local LLN conditions). We note that Assumption 3(ii) of Pötscher and Prucha [16] can be replaced by the slightly weaker but more complex assumption $\sup_n n^{-1}\Sigma_{t=1}^{n} \mathrm{E}[\,|r_{kt}(z_t)|\,1_{K_m}(z_t)] < \infty$ for all $m \in \mathbb{N}$ without changing the proof. We note further that if the functions r_{kt} satisfy the stronger condition $\sup_t |r_{kt}(z_t)|\, 1_{K_m}(z_t) < \infty$ P-a.s. for all $m \in \mathbb{N}$, then the ULLN of Pötscher and Prucha [16] can also be obtained as a special case of Theorem 2 of the present paper (again abstracting from the slight differences in the maintained local LLN conditions).

(v) In this paper we give conditions under which $n^{-1}\Sigma_{t=1}^{n}[q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]$ converges uniformly to zero. For a further discussion of conditions such that not only a ULLN holds, but such that also $n^{-1}\Sigma_{t=1}^{n} q_t(z_t, \theta)$ converges uniformly to a fixed limit, see e.g. Pötscher and Prucha [14,15,16].

(vi) Lemma 1, Theorem 1 and Theorem 2 remain valid if q_t and z_t also depend on n, given the following modifications are made: Assumptions A and B are to be applied accordingly to the triangular array, say, $q_t^n(z_t^n, \theta)$; in Assumption C the functions b_t and h_t may now also depend on n and the "$\sup_t$" in Assumptions C(a) and C(c) is to be replaced by "$\sup_n \sup_{t \le n}$". In Section 3 the equicontinuity conditions postulated in Assumptions 2(a) or 2′(a) have now to hold for the families $\{q_{t,m}^n(z_t^n, \theta): t \le n,\ n \in \mathbb{N}\}$. Similarly the continuity condition in Assumptions 2(b) or 2′(b) has now to hold for $q_t^n(z_t^n, \theta)$. Assumptions 3 and 4 must again be applied to the triangular array $q_t^n(z_t^n, \theta)$. Of course, sufficient conditions similar to the ones presented in Remarks 4–6 can again be derived in a completely analogous way.

We note further that the situation where a norming factor c_n^{-1} other than n^{-1} is used can be immediately incorporated into the framework of this paper by redefining the functions q_t^n as $(n/c_n)q_t^n$.

(vii) If in the assumptions for Lemma 1, Theorems 1 and 2 the respective strong LLNs and strong local LLNs are replaced by weak versions thereof (and all other assumptions are left unchanged), then the lemma and the theorems still hold with the strong ULLN statements replaced by weak ones. (Of course, one then has to assume the P-a.s. measurability of quantities like $\sup_\Theta |n^{-1}\Sigma[q_t(z_t,\theta) - \mathrm{E}q_t(z_t,\theta)]|$, or one has to try to circumvent this problem by the use of outer probabilities.)

5. EXAMPLE

We next give a simple example for which, by using Theorem 2, it is possible to establish a ULLN under weaker conditions than is possible by using the ULLN in Andrews [2] based on a Lipshitz-type condition or the ULLN in Pötscher and Prucha [16].

Let $\Theta = [-a, a]$ with $a > 0$, let $Z = \mathbb{R}^2$ and $z_t = (y_t, x_t)$, let $q_t(z,\theta) = q(z,\theta) = \min\{c, |y - g(x)\theta|\}s(|x|)$ with $c > 0$ and where g: $\mathbb{R} \to \mathbb{R}$ with $|g(x)| = g(|x|)$ monotonically increasing to infinity as $|x| \to \infty$ and s: $[0,\infty) \to [b, \underline{b}]$ with $0 < b < \underline{b} < \infty$ are Borel measurable functions. Assume that (y_t) is generated from the model $y_t = g(x_t)\theta_0 + u_t$ with $\theta_0 \in \Theta$, x_t and u_t (contemporaneously) independent, and where the distributions of x_t and u_t put positive probability on every nondegenerate interval. Assume furthermore that the process (z_t) is α-mixing with mixing coefficients of size $-r/(r-1)$ for some $r > 1$, and that (z_t) is asymptotically stationary.

Within the above setup we may interpret $n^{-1}\Sigma_{t=1}^n q(z_t,\theta)$ as the objective function of a minimization estimator for θ_0. For $s(|x|) \equiv 1$ and for c "large" this estimator reduces essentially to the least absolute deviation estimator. The function $s(|x|)$ allows different observations to enter the objective function with different weights.[12]

To establish a ULLN for the above example via Theorem 2 we need to check the validity of Assumptions 1-4. Assumption 1 is trivially satisfied. Choose $K_m = \{(y,x)\colon |y| \le m,\ |g(x)| \le m\}$, then clearly $K_m \uparrow Z$ and K_m convex as $|g|$ is monotone in $|x|$. Since $|\min\{b, |v_1|\} - \min\{b, |v_2|\}| \le ||v_1| - |v_2|| \le |v_1 - v_2|$ for all v_1, $v_2 \in \mathbb{R}$ it follows that $\sup_{z\in K_m}|q(z,\theta) - q(z,\theta')| \le m\underline{b}|\theta - \theta'|$.

Consequently condition (I) and hence Assumption 2′ is satisfied. Condition (III) holds trivially since q is bounded; condition (IV) is satisfied since clearly condition (Vb) holds. Consequently Assumption 3 is satisfied. Assumption 5 holds since (y_t, x_t) is assumed to be α-mixing and q is bounded. It then follows from Lemma A5 that also Assumption 4 is satisfied. Consequently the example satisfies all of the respective assumptions of Theorem 2 and hence $(q(z_t,\theta)\colon t \in \mathbb{N})$ satisfies a ULLN.

In order to be able to apply Andrews' [2] ULLN to the above example the following Lipshitz-type condition must hold instead of Assumptions 2 and 3:[13]

For each $\theta' \in \Theta$ there is a constant $\tau > 0$ such that $|\theta - \theta'| \le \tau$ implies

$$|q(z_t,\theta) - q(z_t,\theta')| \le B_t(z_t)\, h(|\theta - \theta'|) \tag{1}$$

a.s. for every t, where B_t: $Z \to [0, \infty)$ and h: $[0, \infty) \to [0, \infty)$ such that $B_t(z_t)$ is a random variable, $\limsup_{n\to\infty} n^{-1}\Sigma_{t=1}^{n} \mathrm{E}B_t(z_t) < \infty$, $h(\eta) \downarrow h(0) = 0$ as $\eta \to 0$.

Condition (1) has to hold, in particular, for $\theta' = \theta_0$ and any admissible θ_0; we consider for simplicity the case $\theta' = \theta_0 = 0$. Then

$$|q(z_t,\theta) - q(z_t,\theta')| = |\min\{c, |u_t - g(x_t)\theta|\} - \min\{c, |u_t|\}|s(|x_t|) \quad (2)$$

and $h(|\theta - \theta'|) = h(|\theta|)$.

For expositional reasons we first analyze the case $h(|\theta|) = |\theta|$. Consider some $\tau > 0$ and some realization $(u_t(\omega), x_t(\omega))$ for which (1) holds and for which $c/4 < |u_t(\omega)| < c/2$ (where t is some given index). Clearly there exists a θ with $0 < |\theta| \le \tau$ such that $|g(x_t(\omega))\theta| < |u_t(\omega)|$. Consequently $|q(z_t(\omega),\theta) - q(z_t(\omega),\theta')| = |g(x_t(\omega))|\,|\theta|s(|x_t(\omega)|) \le B_t(z_t(\omega))|\theta|$, and hence $b|g(x_t(\omega))| \le B_t(z_t(\omega))$. This shows that (1) implies $b|g(x_t)|\mathbf{1}_{\{c/4<|u_t|<c/2\}} \le B_t(z_t)\mathbf{1}_{\{c/4<|u_t|<c/2\}}$ a.s., and therefore $b\mathrm{E}[|g(x_t)|]\mathrm{P}(c/4 < |u_t| < c/2) \le \mathrm{E}[B_t(z_t)\mathbf{1}_{\{c/4<|u_t|<c/2\}}] \le \mathrm{E}[B_t(z_t)] < \infty$ in view of condition (1). The probability on the l.h.s. of the inequality is positive; consequently for Andrews' [2] Lipshitz-type condition to be satisfied in the case $h(|\theta|) = |\theta|$ it is necessary that $\mathrm{E}[|g(x_t)|] < \infty$ holds in addition to the assumptions maintained so far for the example. If, for example, $|x| \le |g(x)|$, as is the case if $g(x) = x$ or $g(x) = -g(-x) = i$ for $i - 1 < x \le i$, $i \in \mathbb{N}$, and $g(0) = 0$, then it is necessary that $\mathrm{E}|x_t| < \infty$.

In Appendix B we consider the case of general functions $h(|\theta|)$. We show by similar but somewhat more involved argumentation that e.g.

$$\mathrm{E}[[1/h(c'/|g(x_t)|)]\mathbf{1}_{\{|x_t|>M\}}] < \infty$$

(with $c' = c/4$) for all M sufficiently large is a necessary condition for Andrews' [2] Lipshitz-type condition to be satisfied. Observing that $1/h(c'/|g(x)|)\uparrow\infty$ as $|x|\to\infty$, this condition clearly limits in general the class of admissible processes (x_t). If, for example, $|x| \le |g(x)|$ and $h(|\theta|) = |\theta|^\gamma$ we obtain $\mathrm{E}|x_t|^\gamma < \infty$ as a necessary condition.

We have thus demonstrated that in order to be able to apply Andrews' [2] ULLN to the example we have to make stronger assumptions concerning the distribution of x_t than is necessary if we apply the ULLN presented in this paper. Furthermore, to apply Pötscher and Prucha's [16] ULLN we have to assume that g is continuous. The example thus shows that the ULLN presented in this paper is neither dominated by Andrews' [2] nor Pötscher and Prucha's [16] ULLN.[14] The example also illustrates how the sufficient conditions discussed in Section 3 can be conveniently utilized to check whether Assumptions 1–4 hold in a specific application.

APPENDIX A

In the appendix all summations are to be taken over $t = 1, \ldots, n$ unless indicated otherwise.

Lemma A1. *Let $h_n(\theta)$ for $n \in \mathbb{N}$ be continuous real functions on the metric space (Θ, ρ). If for each $m \in \mathbb{N}$ the family of real functions $\{h_{m,n}: n \in \mathbb{N}\}$ is equicontinuous on Θ and if $\lim_{m\to\infty} \limsup_{n\to\infty} \sup_{\theta\in\Theta}|h_{m,n}(\theta) - h_n(\theta)| = 0$, then also the family $\{h_n: n \in \mathbb{N}\}$ is equicontinuous on Θ.*

Proof. Choose $\varepsilon > 0$. Then there exists an $m_0 = m_0(\varepsilon)$ such that $\limsup_{n\to\infty} \sup_{\theta\in\Theta} |h_{m_0,n}(\theta) - h_n(\theta)| < \varepsilon/3$; hence, for some $n_0 = n_0(m_0(\varepsilon), \varepsilon)$ we have $\sup_{\theta\in\Theta} |h_{m_0,n}(\theta) - h_n(\theta)| < \varepsilon/3$ for all $n \geq n_0$. Consider an arbitrary θ', then there exists a positive $\delta_0 = \delta_0(m_0(\varepsilon), \varepsilon)$ such that $\sup_n |h_{m_0,n}(\theta) - h_{m_0,n}(\theta')| < \varepsilon/3$ if $\rho(\theta, \theta') < \delta_0$, since the family $\{h_{m_0,n}: n \in \mathbb{N}\}$ is equicontinuous on Θ. It now follows that for all $n \geq n_0$ and $\rho(\theta, \theta') < \delta_0$: $|h_n(\theta) - h_n(\theta')| \leq |h_n(\theta) - h_{m_0,n}(\theta)| + |h_{m_0,n}(\theta) - h_{m_0,n}(\theta')| + |h_{m_0,n}(\theta') - h_n(\theta')| < \varepsilon$. For $1 \leq n \leq n_0$ we can find a $\delta_1 > 0$ such that $|h_n(\theta) - h_n(\theta')| < \varepsilon$ for $\rho(\theta, \theta') < \delta_1$ since $h_n(\theta)$ is continuous. Let $\delta = \min(\delta_0, \delta_1)$ then for all $n \in \mathbb{N}$: $|h_n(\theta) - h_n(\theta')| < \varepsilon$ for $\rho(\theta, \theta') < \delta$, which is the desired result. ∎

We note in connection with Remark 1(ii) that the lemma remains valid if the condition that h_n be continuous is dropped but the condition involving the "limsup" is strengthened by replacing the "limsup" with the respective "sup".

Proof of Lemma 1. To prove (a) observe that $\mathrm{E}q_{t,m}(z_t, \theta)$ and $\mathrm{E}q^c_{t,m}(z_t, \theta)$ exist and are finite by the assumptions of the lemma and hence also $\mathrm{E}q_t(z_t, \theta)$ exists and is finite. Furthermore for all $m \in \mathbb{N}$

$$\sup_{\theta\in\Theta} |n^{-1}\Sigma[q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]| \leq A_{n,m} + B_{n,m}$$

where

$$A_{n,m} = \sup_{\theta\in\Theta} |n^{-1}\Sigma[q^c_{t,m}(z_t, \theta) - \mathrm{E}q^c_{t,m}(z_t, \theta)]|$$

and

$$B_{n,m} = \sup_{\theta\in\Theta} |n^{-1}\Sigma[q_{t,m}(z_t, \theta) - \mathrm{E}q_{t,m}(z_t, \theta)]|.$$

Since by assumption a ULLN holds for $q_{t,m}(z_t, \theta)$ we have $\lim_{n\to\infty} B_{n,m} = 0$, P-a.s. Since $A_{n,m} \leq n^{-1}\Sigma[d^c_{t,m}(z_t) - \mathrm{E}d^c_{t,m}(z_t)] + 2n^{-1}\Sigma\mathrm{E}d^c_{t,m}(z_t)$ it follows from Assumption A that $0 \leq \limsup_{m\to\infty} \limsup_{n\to\infty} A_{n,m} \leq \lim_{m\to\infty} \limsup_{n\to\infty} 2n^{-1}\Sigma\mathrm{E}d^c_{t,m}(z_t) = 0$, P-a.s.. This proves part (a) of Lemma 1. Part (b) follows (given the maintained set of assumptions) from Lemma A1 and observing that $\lim_{m\to\infty} \limsup_{n\to\infty} \sup_{\theta\in\Theta} |n^{-1} \Sigma[\mathrm{E}q_{t,m}(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]| \leq \lim_{m\to\infty} \limsup_{n\to\infty} n^{-1}\Sigma\mathrm{E}d^c_{t,m}(z_t) = 0.$ ∎

Proof of Theorem 1. The existence and finiteness of $\mathrm{E}q_t(z_t, \theta)$ follows from the existence and finiteness of $\mathrm{E}q^*_t(z_t, \theta, \tau_k(\theta))$ and $\mathrm{E}q_{t^*}(z_t, \theta, \tau_k(\theta))$, which in turn is implied by Assumption B and Definition 2. Consider an arbitrary $\theta' \in \Theta$, let $(\tau_k(\theta'))_{k\in\mathbb{N}}$ be the sequence defined implicitly by Assumption B and Definition 2, and let $N_{\theta'}$ be the exceptional null set in Assumption C. We first show that for $\tau_k = \tau_k(\theta') \to 0$:

$$A(\theta', \tau_k(\theta')) = \sup_n |n^{-1}\Sigma[\mathrm{E}q^*_t(z_t, \theta', \tau_k(\theta')) - \mathrm{E}q_t(z_t, \theta')]| \to 0, \quad \text{(A.1)}$$

$$B(\theta', \tau_k(\theta')) = \sup_n |n^{-1}\Sigma[\mathrm{E}q_{t^*}(z_t, \theta', \tau_k(\theta')) - \mathrm{E}q_t(z_t, \theta')]| \to 0. \quad \text{(A.2)}$$

Upon taking appropriate suprema and upon applying Hölder's inequality twice to the inequality given in Assumption C we obtain for $\tau_k \leq \tau_{k_0}$:

$$A(\theta', \tau_k(\theta')) \le \sup_n n^{-1}\Sigma \mathrm{E}[b_t(z_t, \theta')h_t^*(z_t, \theta', \tau_k(\theta'))] \tag{A.3}$$
$$\le \{\sup_n n^{-1}\Sigma \mathrm{E}[b_t(z_t, \theta')^p]\}^{1/p}\{\sup_n n^{-1}\Sigma \mathrm{E}[h_t^*(z_t, \theta', \tau_k(\theta'))^q]\}^{1/q},$$

where for $p = \infty$ the first expression on the r.h.s. has to be replaced by $\sup_t \|b_t(z_t, \theta')\|_\infty$ and where for $p = 1$ the second expression on the r.h.s. has to be replaced by $\sup_t \|h_t^*(z_t, \theta', \tau_k(\theta'))\|_\infty$. Assumption C(c) implies that the first expression on the r.h.s. of (A.3) if finite for $1 \le p \le \infty$. Hence to prove (A.1) it is sufficient to show that the second expression on the r.h.s. of (A.3) goes to zero as $\tau_k(\theta') \to 0$. For $p = 1$ this is directly implied by Assumption C(c). Next consider the case $1 < p \le \infty$: Assumption C(a) implies that $\sup_t h_t^*(z_t, \theta', \tau_k(\theta')) \to 0$ and hence $\sup_n n^{-1}\Sigma h_t^*(z_t, \theta', \tau_k(\theta'))^q \to 0$ for $\omega \in \Omega \setminus N_{\theta'}$ and $\tau_k(\theta') \to 0$. The family $\{n^{-1}\Sigma h_t^*(z_t, \theta', \tau_k(\theta'))^q: n \in \mathbb{N}, \tau_k \le \tau_{k_0}\}$ is uniformly integrable as a consequence of Assumption C(c) since $\tau_k \le \tau_{k_0}$ for k large enough. Applying Theorem A3 of Hoadley [9] completes the proof of (A.1).[15] A similar argument proves (A.2). Clearly for $\rho(\theta, \theta') < \tau_k(\theta')$:

$$\sup_n |n^{-1}\Sigma[\mathrm{E}q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta')]| \le \max[A(\theta', \tau_k(\theta')), B(\theta', \tau_k(\theta'))].$$

Part (b) of Theorem 1 then follows since the r.h.s. goes to zero as $\tau_k(\theta') \to 0$. To prove part (a) of Theorem 1 observe that using (A.1) and (A.2) we can find for every $\varepsilon > 0$ and $\theta' \in \Theta$ a $\tau(\varepsilon, \theta') = \tau_{k(\varepsilon, \theta')}(\theta')$ such that $A(\theta', \tau(\varepsilon, \theta')) < \varepsilon$ and $B(\theta', \tau(\varepsilon, \theta')) < \varepsilon$. Consequently, for all $\theta \in \Theta$ with $\rho(\theta, \theta') < \tau(\varepsilon, \theta')$, for all $n \in \mathbb{N}$ and $\omega \in \Omega$:

$$\begin{aligned}
-2\varepsilon + n^{-1}\Sigma[q_{t^*}(z_t, \theta', \tau(\varepsilon, \theta')) - \mathrm{E}q_{t^*}(z_t, \theta', \tau(\varepsilon, \theta'))]&\\
\le n^{-1}\Sigma[q_{t^*}(z_t, \theta', \tau(\varepsilon, \theta')) - \mathrm{E}q_t^*(z_t, \theta', \tau(\varepsilon, \theta'))]&\\
\le n^{-1}\Sigma[q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]&\\
\le n^{-1}\Sigma[q_t^*(z_t, \theta', \tau(\varepsilon, \theta')) - \mathrm{E}q_{t^*}(z_t, \theta', \tau(\varepsilon, \theta'))]&\\
\le n^{-1}\Sigma[q_t^*(z_t, \theta', \tau(\varepsilon, \theta')) - \mathrm{E}q_t^*(z_t, \theta', \tau(\varepsilon, \theta'))] + 2\varepsilon.&
\end{aligned}$$

Now cover Θ by the collection of balls $\{O(\theta', \tau(\theta', \varepsilon)): \theta' \in \Theta\}$ where $O(\theta', \tau) = \{\theta \in \Theta: \rho(\theta, \theta') < \tau\}$. Since Θ is compact there exists a finite subcover of balls from this collection centered at $\theta_1', \ldots, \theta_r'$. Let $\tau_i = \tau(\varepsilon, \theta_i')$, then we have for all $\theta \in \Theta$, all $n \in \mathbb{N}$ and all $\omega \in \Omega$:

$$\begin{aligned}
-2\varepsilon + \min_{i=1,\ldots,r} n^{-1}\Sigma[q_{t^*}(z_t, \theta_i', \tau_i) - \mathrm{E}q_{t^*}(z_t, \theta_i', \tau_i)]&\\
\le n^{-1}\Sigma[q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]&\\
\le \max_{i=1,\ldots,r} n^{-1}\Sigma[q_t^*(z_t, \theta_i', \tau_i) - \mathrm{E}q_t^*(z_t, \theta_i', \tau_i)] + 2\varepsilon.&
\end{aligned}$$

Because of Assumption B it follows further that $\limsup_{n\to\infty} \sup_{\theta\in\Theta} |n^{-1}\Sigma[q_t(z_t, \theta) - \mathrm{E}q_t(z_t, \theta)]| \le 2\varepsilon$, P-a.s.. This clearly implies part (a) of Theorem 1. ∎

Lemma A2. *Under Assumptions* 1 *and* 2′(*b*) *the functions* $d_t(z_t)$, $d_{t,m}^c(z_t)$, $q_{t,m}^*(z_t, \theta', \tau)$, $q_{t,m^*}(z_t, \theta', \tau)$, $q_t^*(z_t, \theta', \tau)$ *and* $q_{t^*}(z_t, \theta', \tau)$ *are measurable* P-*a.s.*.

Proof. Observe that for $\omega \notin N$, the null set defined in Assumption 2′(b), $q_t(z_t(\omega), \theta)$ is continuous on Θ. Hence for all $\omega \notin N$ and any countable dense set $\Theta_0 \subseteq \Theta$ we have $d_t(z_t(\omega)) = \sup_{\theta\in\Theta_0}|q_t(z_t(\omega), \theta)|$. The latter expression is a countable supremum of measurable functions and hence it is measurable. Since $d_t(z_t)$ coincides with this expression except on a set of measure zero, $d_t(z_t)$ is measurable P-a.s.. The proof for the remaining functions is analogous. ■

Proof of Theorem 2. We first show that the assumptions of the theorem imply that the assumptions of Theorem 1 hold for $q_{t,m}$ for each $m \in \mathbb{N}$. Assumption B holds for $q_{t,m}$ because of Assumption 4. To verify Assumption C for $q_{t,m}$ put $b_t(z_t, \theta') \equiv 1$ and $h_t(z_t, \theta', \theta) = |q_{t,m}(z_t, \theta) - q_{t,m}(z_t, \theta')|$, let $\tau_k(\theta')$ be the sequence defined implicitly in Assumption 4 (corresponding to $q_{t,m}$), put $k_0(\theta') = 1$ and $p = \infty$. Then part (a) of Assumption C follows from Assumption 2(a). To verify part (b) observe that $h_t^*(z_t, \theta', \tau_k(\theta')) = \max\{q_{t,m}^*(z_t, \theta', \tau_k(\theta')) - q_{t,m}(z_t, \theta'), q_{t,m}(z_t, \theta') - q_{t,m*}(z_t, \theta', \tau_k(\theta'))\}$ is P-a.s. measurable for all $k \geq 1$ in view of the P-a.s. measurability assumption implicit in Assumption 4. Part (c) of Assumption C follows clearly from Assumption 3(b) for $p = \infty$, using the triangle inequality and bounding h_t^* by $2d_{t,m}$. Consequently, $q_{t,m}$ satisfies (a) and (b) of Theorem 1. Now Assumption A is satisfied in view of Assumptions 3(a) and 4(b). Hence part (a) of Theorem 2 follows from Lemma 1 and the just established ULLN for $q_{t,m}$. Part (b) also follows from Lemma 1 if we can establish existence, finiteness and continuity of $n^{-1}\Sigma\, \mathrm{E}q_t(z_t, \theta)$. But since $\mathrm{E}d_t(z_t) < \infty$ by Assumption 3 existence and finiteness are obvious; since $q_t(z_t, \theta)$ is P-a.s. continuous at each $\theta' \in \Theta$ by Assumption 2(b), we get continuity of $\mathrm{E}q_t(z_t, \theta)$ at each $\theta' \in \Theta$ using the Dominated Convergence Theorem.

Lemma A3. (*i*) *Condition* (*I*) *implies Assumption* 2′ *and hence Assumption* 2. (*ii*) *If Z is a metrizable space then condition* (*II*) *implies condition* (*I*).

Proof. Part (i) of the lemma follows trivially since $\sup_t |q_{t,m}(z_t, \theta) - q_{t,m}(z_t, \theta')| \leq \sup_t \sup_{z \in K_m} |q_t(z, \theta) - q_t(z, \theta')|$. Part (ii) of the lemma follows from Lemma A1 in Pötscher and Prucha [16] if $K_m \neq \phi$ and is trivially true if $K_m = \phi$.

Lemma A4. *Assume* $d_t(z_t)$ *is measurable* P-*a.s.*.

(*i*) *Conditions* (*III*) *and* (*IV*) *imply Assumption* 3.
(*ii*) *Condition* (*IV*) *is implied by either one of* (*Va*), (*Vb*) *or* (*Vc*).
(*iii*) *Conditions* (*VI*) *or* (*VII*) *also imply Assumption* 3.

Proof. (i) Condition (III) clearly implies $\mathrm{E}[d_t(z_t)^{1+\delta}] < \infty$ and hence $\mathrm{E}d_t(z_t) < \infty$. Since the family of functions in Assumption 3(b) is clearly dominated by the family of functions $\{n^{-1}\Sigma d_t(z_t)\colon n \in \mathbb{N}\}$ it is sufficient to prove uniform integrability for the latter. However, this is readily seen since $\sup_n \mathrm{E}\{n^{-1}\Sigma d_t(z_t)\}^{1+\delta} \leq \sup_n\{n^{-1}\Sigma \mathrm{E}[d_t(z_t)^{1+\delta}]\}$, which is finite as a consequence of condition (III). Next define $p = 1 + \delta$ and $p^{-1} + q^{-1} = 1$, then by applying Hölder's inequality twice we see that $n^{-1}\Sigma \mathrm{E}d_{t,m}^c(z_t) \leq n^{-1}\Sigma\{\mathrm{E}[d_t(z_t)^p]\}^{1/p}\{\mathrm{E}1_{Z-K_m}(z_t)\}^{1/q} \leq \{n^{-1}\Sigma \mathrm{E}[d_t(z_t)^p]\}^{1/p}\{n^{-1}\Sigma \mathrm{P}(z_t \notin K_m)\}^{1/q}$. Therefore $0 \leq \limsup_{m\to\infty}\{\limsup_{n\to\infty} n^{-1}\Sigma \mathrm{E}d_{t,m}^c(z_t)\} \leq \sup_n\{n^{-1}\Sigma \mathrm{E}[d_t(z_t)^p]\}^{1/p}\{\lim_{m\to\infty}\limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(z_t \notin K_m)\}^{1/q} = 0$ because of conditions (III) and (IV).

(ii) Condition (IV) is trivially implied by (Vc). Next consider conditions (Va) and (Vb). We can then find open balls $B(0, r_m) = \{z\colon \|z\| < r_m\}$ such that $r_m \to \infty$ and $B(0, r_m) \subseteq K_m$ for all large enough m, say $m > m_0$. (This follows since every ball

$B(0,r)$ is contained in the convex hull of a finite number of suitably chosen points, since these points will be contained in K_m for large enough m and since K_m is convex.) From Markov's inequality we have for m large enough:

$$\begin{aligned}0 &\le \limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(z_t \notin K_m) \le \limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(z_t \notin B(0,r_m))\\ &= \limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(\|z_t\| \ge r_m) \le [\limsup_{n\to\infty} n^{-1}\Sigma \mathrm{E}h(\|z_t\|)]/h(r_m).\end{aligned}$$

Observing that condition (Va) implies that the r.h.s. goes to zero as $m \to \infty$ we see that (Va) implies (IV). Furthermore for $m > m_0$:

$$\begin{aligned}0 &\le \limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(z_t \notin K_m) \le \limsup_{n\to\infty} n^{-1}\Sigma \mathrm{P}(z_t \notin B(0,r_m))\\ &= \limsup_{n\to\infty} \int \mathbf{1}_{Z-B(0,r_m)}(z)\,\mathrm{d}(n^{-1}\Sigma H_t) \le \int \mathbf{1}_{Z-B(0,r_m)}(z)\,\mathrm{d}H\end{aligned}$$

by Theorem 2.1 in Billingsley [6]. Observing that the r.h.s. goes to zero as $m \to \infty$ shows that (Vb) implies (IV).

(iii) Condition (VI) clearly implies $\mathrm{E}d_t(z_t) < \infty$ and Assumption 3(a). Since $d_t(z_t) \ge 0$ holds, condition (VI) is equivalent to L_1-convergence of $n^{-1}\Sigma d_t(z_t)$. Therefore this sequence is uniformly integrable, which in turn implies Assumption 3(b). Finally, Assumption 3 is trivially implied by condition (VII). ∎

We note that Lemma A4(ii) also holds if the following generalizations are made in (Va) or (Vb). In (Va) Z can be taken to be an arbitrary set and the sequence K_m is taken such that there exists a β-measurable function $g\colon Z \to [0,\infty]$ and real numbers $r_m \to \infty$ satisfying (i) $\{z \in Z\colon g(z) < r_m\} \subseteq K_m$ and (ii) $\limsup_{n\to\infty} n^{-1}\Sigma \mathrm{E}g(z_t) < \infty$. In (Vb) (Z,β) can be taken to be a metrizable space with its Borel field and the sequence K_m $(\in \beta)$ is taken to satisfy $\operatorname{int} K_m \uparrow Z$ where $\operatorname{int} K_m$ denotes the interior of K_m. We also note that given Z is a metrizable space, condition (IV) holds for some sequence of compact sets K_m, if the sequence $n^{-1}\Sigma H_t$ is tight. In particular this follows if $n^{-1}\Sigma H_t$ converges weakly to a probability measure H and if each measure $n^{-1}\Sigma H_t$ as well as H are tight, cf. Billingsley [6, Theorem 8, Appendix III] and Pötscher and Prucha [16].

Lemma A5. *Assumptions* 1, 2′(*b*) *and* 5 *imply Assumption* 4.

Proof. Similarly as in the proof of Lemma A2 it follows that

$$q^*_{t,m}(z_t,\theta',\tau_k(\theta')) = g_{t,m}(z_t), \qquad \text{P-a.s.},$$

where

$$g_{t,m}(z) = \sup\{q_{t,m}(z,\theta)\colon \theta \in \Theta_0,\ \rho(\theta,\theta') < \tau_k(\theta')\}$$

and $\Theta_0 \subseteq \Theta$ is an arbitrary countable dense set. Clearly $g_{t,m}$ is β-measurable and takes its values in $\mathbb{R} \cup \{+\infty\}$. By Assumption 5 (z_t) is ϕ-mixing [α-mixing]. The measurability of $g_{t,m}$ implies that $g_{t,m}(z_t) - \mathrm{E}g_{t,m}(z_t)$ is ϕ-mixing [α-mixing] with mixing coeffients of the same size. Furthermore $\sup_t \mathrm{E}|g_{t,m}(z_t) - \mathrm{E}g_{t,m}(z_t)|^{r+\delta} < \infty$. Observing that McLeish's [12] definition of mixing coefficients is slightly weaker than the usual definition employed in this paper it follows from his Theorem 2.10 that $g_{t,m}(z_t)$ and hence $q^*_{t,m}(z_t,\theta',\tau_k(\theta'))$ satisfies a strong law of large numbers. Analogously the same is established for $q_{t,m*}(z_t,\theta',\tau_k(\theta'))$ and $d^c_{t,m}(z_t)$.

APPENDIX B

Consider the example of Section 5 and assume that Andrews' [2] Lipshitz-type condition (1) holds. As remarked in the text this condition has to hold in particular for $\theta' = \theta_0 = 0$. Consider some $\tau > 0$, some $M > 0$ such that $c/[4g(M)] \le \tau$ and $g(M) > 0$. Let $(u_t(\omega), x_t(\omega))$ be some realization for which (1) holds for all θ with $|\theta| \le \tau$ and for which $c/4 < |u_t(\omega)| < c/2$ and $|x_t(\omega)| > M$. Choose $\theta = c/[4g(x_t(\omega))]$ and observe that for this choice $0 < |\theta| \le \tau$. Utilizing (2) it is readily seen that condition (1) implies $|q(z_t(\omega), \theta) - q(z_t(\omega), \theta')| = ||u_t(\omega) - c/4| - |u_t(\omega)||s(|x_t(\omega)|) = (c/4)s(|x_t(\omega)|) \le B_t(z_t(\omega))h(c/|4g(x_t(\omega))|)$ and hence $c_* \le B_t(z_t(\omega))h(c/|4g(x_t(\omega))|)$ with $c_* = bc/4$ observing that s is bounded from below by b. Consequently $[1/h(c/|4g(x_t)|)]\mathbf{1}_{\{|x_t|>M\}}\mathbf{1}_{\{c/4<|u_t|<c/2\}} \le (1/c_*)B_t(z_t)\mathbf{1}_{\{|x_t|>M\}}\mathbf{1}_{\{c/4<|u_t|<c/2\}}$ a.s., and therefore $\mathsf{E}[[1/h(c/|4g(x_t)|)]\mathbf{1}_{\{|x_t|>M\}}]\,\mathsf{P}(c/4<|u_t|<c/2) \le (1/c_*)\mathsf{E}[B_t(z_t)] < \infty$. That the first factor on the l.h.s. has to be finite follows since $\mathsf{P}(c/4<|u_t|<c/2)$ is assumed to be positive.

Endnotes

1 We would like to thank Donald Andrews, Charles Bates, Manfred Deistler, Ian Domowitz, Ronald Gallant, Nanhua Hu, Harry Kelejian, Whitney Newey, Peter Phillips, Ching-Zong Wei, Halbert White, and Ernest Zampelli for helpful comments. We assume responsibility, however, for any errors. Some of the results presented in this paper were circulated earlier in Pötscher and Prucha [15]. The present paper is both an extension and revision of this earlier paper.

2 We note that Hoadley [9] does not use his ULLN in his consistency proof, hence his consistency result is not affected by the restrictiveness of his equicontinuity assumption. We note further that the proofs of theorems regarding consistency in the papers by Bates, Domowitz, Levine and White are such that Hoadley's ULLN can be replaced by some alternative ULLN. I.e., the theorems can be rectified and/or restored to their intended generality by use of an alternatively ULLN accompanied by a corresponding change in the catalogs of assumptions. For a more detailed discussion of this issue see, e.g., Pötscher and Prucha [15].

3 Of course, rather than to specify a set of assumptions that implies the first moment continuity condition one could maintain the latter as an assumption. It seems that the usefulness of such a result is limited, apart from emphasizing the structure of the proof.

4 All of the subsequent conditions and results do not depend, in an essential way, on the metric structure of (Θ, ρ), but only on the metrizability of the topology on Θ. The choice of a fixed metric is made only for convenience.

5 We call a function measurable P-a.s. if it coincides with a measurable function on a set of P-measure one. Clearly integrals of such functions remain well defined and every measurable function is measurable P-a.s.. Compare also footnote 10.

6 We note that the limit on the l.h.s. of the equation in Assumption A(a) exists automatically if the sequence (K_m) is monotonically increasing. In fact, (for general sequences (K_m)) the formally weaker condition where "liminf" replaces "lim" in A(a) can always be reduced to the condition as given in A(a) by passing to a suitable subsequence of K_m. We also note that in this paper often conditions which are satisfied for K_m are also satisfied for the corresponding monotonized sequence $K'_m = \cup_{i=1}^{m} K_i$; e.g. Assumption A(a) holds for K'_m if it holds for K_m.

7 This also implies that $f_t^*(z_t, \theta', \tau_k)$ and $f_{t*}(z_t, \theta', \tau_k)$ are P-a.s. finite, hence the Césaro-sums in Definition 2 are P-a.s. well defined. Note that these Césaro-sums are even well defined for all $\omega \in \Omega$ if the expectations exist, are finite and if f_t has as its range $\mathbb{R}$, since then the range of f_t^* and f_{t*} is, respectively, $\mathbb{R} \cup \{+\infty\}$ and $\mathbb{R} \cup \{-\infty\}$.

8 We note that also p may depend on θ'. Furthermore, Theorem 1 clearly remains valid if b_t and h_t in Assumption C are specified as $b_t(\omega, \theta')$ and $h_t(\omega, \theta', \theta)$, respectively, and not as composite functions of z_t.

9 Assumption B enters this argument also insofar as the existence and finiteness of the expectations $\mathsf{E}q_t(z_t, \theta)$, etc., is derived from that assumption.

10 Note that the existence and finiteness of $\mathsf{E}q^*_{t,m}(z_t, \theta', \tau_k)$, and $\mathsf{E}q_{t,m*}(z_t, \theta', \tau_k)$, and $\mathsf{E}d^c_{t,m}(z_t)$,

which is implicitly assumed in Assumption 4, is automatically implied by Assumption 3. Furthermore, it follows from Lemma A2 in Appendix A that under Assumptions 1 and 2′(b) the P-a.s. measurability conditions postulated explicitly and implicitly in Assumptions 3 and 4 are automatically satisfied. Also, note that under Assumptions 1 and 2′(b) the functions $d_t(z_t)$ are measurable P-a.s., but not necessarily measurable. By modifying $d_t(z_t)$ on an appropriate ω-set of measure zero, we could always obtain a measurable function of the argument ω. However, this modified function need not be expressible as a composite function of z_t. It is for this reason that we only assume P-a.s. measurability rather than measurability.

11 For example $h(x) = x^p$, $p > 0$, or $h(x) = \ln(1+x)$.

12 As it is of interest to analyze the properties of feasible GLS estimators under classical (OLS) assumptions, we allow for different weights for different observations in the objective function defining the estimator, although the spread of u_t is assumed not to depend on x_t.

13 Compare Assumption A4 in Andrews [2, p. 1467] or Remark 2 above, and observe that $q_t \equiv q$ and that the z_t's are identically distributed.

14 From the discussion in Andrews [2] and Pötscher and Prucha [15] it is also evident that Hoadley's [9] ULLN does not apply to this example in general.

15 Theorem A3 in Hoadley [9] is formulated for measureble functions. Since we apply that theorem only to a *sequence* of P-a.s. measurable functions, we can modify them on a common null set so that they are measurable and then apply Hoadley's Theorem A3 to the modified functions.

References

[1] Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press. Cambridge, MA.

[2] Andrews, D. W. K. (1987). Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica*, **55**, 1465–1471.

[3] Bates, C. and White, H. (1985). A unified theory of consistent estimation for parametric models. *Econometric Theory*, **1**, 151–178.

[4] Bierens, H. J. (1982). A uniform weak law of large numbers under ϕ-mixing with applications to nonlinear least squares. *Statistica Neerlandica*, **36**, 81–86.

[5] Bierens, H. J. (1989). Least squares estimation of linear and nonlinear ARMAX models under data heterogeneity. Department of Econometrics, Free University, Amsterdam.

[6] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[7] Domowitz, I. and White, H. (1982). Misspecified models with dependent observations. *Journal of Econometrics*, **20**, 35–58.

[8] Hansen, L. P. (1982). Large sample properties of generalized method of moment estimators. *Econometrica*, **50**, 1029–1054.

[9] Hoadley, B. (1971). Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case. *Ann. Math. Statist.*, **42**, 1977–1991.

[10] Jennrich, R. I. (1969). Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.*, **40**, 633–643.

[11] Levine, D. (1983). A remark on serial correlation in maximum likelihood. *Journal of Econometrics*, **23**, 337–342.

[12] McLeish, D. L. (1975). A maximal inequality and dependent strong laws. *Ann. Prob.*, **3**, 829–839.

[13] Newey, W. K. (1987). Expository notes on uniform laws of large numbers. Department of Economics, Princeton University.

[14] Pötscher, B. M. and Prucha, I. R. (1986). A class of partially adaptive one-step *M*-estimators for the nonlinear regression model with dependent observations. *Journal of Econometrics*, **32**, 219–251.

[15] Pötscher, B. M. and Prucha, I. R. (1986). 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.

[16] Pötscher, B. M. and Prucha, I. R. (1989). A uniform law of large numbers for dependent and heterogeneous data processes. *Econometrica*, **57**, 675–583.

[17] Pötscher, B. M. and Prucha, I. R. (1991) Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I: Consistency and approximation concepts. *Econometric Reviews*, **10**, 125–216.

[18] Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.*, **20**, 595–600.

[19] White, H. (1980). Nonlinear regression on cross-section data. *Econometrica*, **48**, 721–746.

[20] White, H. and Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, **52**, 143–161.

B. M. Pötscher
Institut für Statistik,
Operations Research und Computerverfahren
Universität Wien
A-1010 Wien
Austria

I. R. Prucha
Department of Economics
University of Maryland
College Park, MD 20742
U.S.A.