Cowles Foundation Paper 810

Econometric Theory, **8**, 1992, 241–257. Printed in the United States of America.

GENERIC UNIFORM CONVERGENCE

Donald W.K. Andrews
Yale University

This paper presents several generic uniform convergence results that include generic uniform laws of large numbers. These results provide conditions under which pointwise convergence almost surely or in probability can be strengthened to uniform convergence. The results are useful for establishing asymptotic properties of estimators and test statistics.

The results given here have the following attributes, (1) they extend results of Newey [15] to cover convergence almost surely as well as convergence in probability, (2) they apply to totally bounded parameter spaces (rather than just to compact parameter spaces), (3) they introduce a set of conditions for a generic uniform law of large numbers that has the attribute of giving the weakest conditions available for i.i.d. contexts, but which apply in some dependent nonidentically distributed contexts as well, and (4) they incorporate and extend the main results in the literature in a parsimonious fashion.

1. INTRODUCTION

This paper extends several generic uniform convergence results and uniform laws of large numbers (LLN's) given in the literature by Andrews [1], Pötscher and Prucha [18], and Newey [15]. By definition, these "generic" results provide sufficient conditions to turn pointwise convergence (or pointwise LLN's) into uniform convergence (or uniform LLN's). As is well known, uniform convergence and uniform LLN's are used in standard proofs of consistency of extremum estimators, including parametric, semiparametric, and nonparametric estimators. In particular, they are most useful for consistency proofs for estimators that are only implicitly defined and are such that the criterion function is a function of partial sums. This includes nonlinear least squares estimators, *M*-estimators, generalized method of moments estimators, and maximum likelihood estimators, as well as various semiparametric and nonparametric extremum estimators. For examples of the use of uniform LLN's in establishing the consistency and asymptotic normality of extremum estimators, see Andrews [3,4], Gallant [8], Gallant and White [9], and Pötscher and Prucha [19].

The author thanks B.M. Pötscher, W.K. Newey, Y.J. Whang, the co-editor P.M. Robinson, and two referees for helpful comments and suggestions. The author gratefully acknowledges the financial support of the National Science Foundation via grant number SES-8821021.

 0266-4666/92 $5.00 + .00

The contribution of the present paper to the literature on generic uniform results is as follows: First, the paper follows the approach of Newey [15] and uses a stochastic equicontinuity (SE) condition. The SE condition of Newey [15] is modified, however, in a way that simplifies the condition itself and simplifies the proofs of the results. The modification also allows one to derive almost sure (a.s.) results in addition to the "in probability" results given by Newey [15]. As in Newey [15], the SE condition is shown to be a necessary condition for the generic uniform results. When replaced by more primitive conditions, the SE condition yields simpler conditions than those obtained by following the first moment continuity approach adopted in Andrews [1] and Pötscher and Prucha [18]. In some cases, the resultant conditions are also easier to verify.

Second, the results given here apply when the parameter space is noncompact. The compactness condition is replaced by a condition of total boundedness. (Recall that a subset of a metric space is totally bounded if it can be covered by a finite number of ϵ balls $\forall \epsilon > 0$.) A totally bounded and complete parameter space is compact, so total boundedness is weaker than compactness. For Euclidean parameter spaces, sets are totally bounded iff they are bounded, whereas sets are compact iff they are bounded and closed. In this case, the results of the paper remove the restriction that the parameter space is closed. See below for several reasons why the relaxation of this restriction is useful.[1]

Third, the two main sets of primitive conditions for generic uniform LLN's given by Andrews [1] and Pötscher and Prucha [18] have the drawback that they do not contain the best results available for the case when the random variables (rv's) are independent and identically distributed (i.i.d.) (e.g., see the uniform LLN of Jennrich [12, Theorem 2]). Here, we present a third set of primitive conditions for a uniform LLN (see Assumption TSE-1 of Section 3) that applies to dependent nonidentically distributed (d.n.i.d.) rv's. These conditions have the attribute that they obtain (and generalize) the best results available in the case where the rv's are i.i.d. They also contain a generic analog of Bierens' [5, Lemma 3] uniform LLN, as well as a new generic uniform LLN based on an absolute continuity condition (see Assumption TSE-1A below). This third set of primitive conditions does not dominate the conditions given in Andrews [1] or Pötscher and Prucha [18], however, since it places much more restrictive conditions on the extent to which the rv's may exhibit nonidentical distributions. Each of the three sets of primitive condi-

[1] The weakening of the compactness condition is not tied to the SE approach used here. A similar modification of the first moment continuity (FMC) condition used by Andrews [1] and Pötscher and Prucha [18] would allow their results to be applied to totally bounded parameter spaces. The modification required is that the FMC condition (see Assumption A3 of Andrews [1]) needs to hold uniformly over $\theta \in \Theta$ (i.e., one needs to add $\sup_{\theta \in \Theta}$ before $\lim_{\rho \to 0}$ in Assumption A3). Then, the same argument as in the proof of the Theorem of Andrews [1] can be used to obtain a uniform LLN, noting that $\rho(\theta)$ in (4) does not depend on θ, and hence, that the collection of ρ-balls $\{B(\theta,\rho) : \theta \in \Theta\}$ that covers Θ has a finite subcover under the total boundedness assumption.

tions involves a different tradeoff between conditions. Depending upon the context, any one of the three may be the most suitable.

Fourth, the results given here are simple and compact. They incorporate and extend the main results in the literature (or variants thereof) in a parsimonious fashion.

Next, we discuss three reasons why it is useful to replace the compactness assumption on the parameter space with the total boundedness assumption. First, in some cases the natural parameter space has one or more parameters restricted by a strict inequality. For example, a variance parameter may be restricted to be positive. In this case, the parameter space has to be artificially restricted if it is to be compact but need not be if it is to be totally bounded.

Second, to obtain asymptotic normality of parametric estimators, the true parameter must be an interior point of the parameter space. Hence, to obtain asymptotic normality of an estimator when the true parameter can be any point in the parameter space, the parameter space must be open (as is often assumed in asymptotic normality results in the literature). This assumption is incompatible with the standard compactness assumption used for uniform LLN's and consistency results. With a totally bounded parameter space, however, there is no incompatibility.

Third, the econometric literature on uniform LLN's and consistency in nonlinear models tends to indicate that compactness is a crucial feature of the methods used. The results given here show that it is not a crucial feature, but rather, just a convenient property.

Subsequent to the present paper, Pötscher and Prucha [20] have shown that an alternative method of generating uniform convergence results for a noncompact totally bounded metric parameter space is to extend all of the functions defined on this space to its (compact) completion and apply existing results for compact parameter spaces. Pötscher and Prucha [20] also provide a number of useful results comparing the conditions given in various uniform convergence and uniform LLN papers in the literature, including the present one.

The remainder of this paper is organized as follows: Section 2 provides the generic uniform convergence results for convergence in probability and convergence almost surely. Section 3 presents the generic uniform LLN's mentioned above.

2. GENERIC UNIFORM CONVERGENCE

We start by giving sufficient conditions for a sequence $\{G_n(\theta): n \geq 1\}$ to converge to zero in probability and almost surely uniformly over an index set Θ as $n \to \infty$. For the special case of a uniform LLN, $G_n(\theta)$ is of the form $(1/n)\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta))$, where Σ_1^n denotes $\sum_{t=1}^n$. Let $(\Omega,\mathcal{F},\mathbf{P})$ be a probability space and let Θ be a metric space with metric d. For each $\theta \in \Theta$ and $n \geq 1$, let $G_n(\theta)$ $(= G_n(\theta)_\omega)$ be a real measurable function on Ω.

Let $B(\theta,\delta)$ denote a closed ball in Θ of radius $\delta \geq 0$ centered at θ. All limits are taken as $n \to \infty$ unless stated otherwise. All suprema over subsets of Θ (or of $\Theta \times \Theta$) of random functions of θ (or of (θ,θ')) that are used below are assumed to be measurable.

First, we consider "in probability" results.

DEFINITION. *$\{G_n(\theta): n \geq 1\}$ is stochastically equicontinuous on Θ if $\forall \epsilon > 0$ $\exists \delta > 0$ such that*

$$\overline{\lim_{n\to\infty}} P\left(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |G_n(\theta') - G_n(\theta)| > \epsilon\right) < \epsilon. \qquad \blacksquare$$

This definition is the standard one in the literature. For example, it is the definition used in Pollard [16] and Andrews [3,4]. The concept of stochastic equicontinuity is quite old and appears in the literature under various guises. For example, it appears in Theorem 8.2 of Billingsley [7, p. 55], which is attributed to Prohorov [21], for the case of $C[0,1]$ random elements. Moreover, a nonasymptotic analog of stochastic equicontinuity arises in the even older literature on the existence of stochastic processes with continuous sample paths.

The assumptions below use the following abbreviations: BD for totally bounded, P-WCON for pointwise weak (i.e., "in probability") convergence to zero, SE for stochastic equicontinuity, and U-WCON for uniform weak convergence to zero.

Assumption BD. Θ is a totally bounded metric space with metric d.

Assumption P-WCON. $G_n(\theta) \xrightarrow{p} 0$ $\forall \theta \in \Theta$.

Assumption SE. $\{G_n(\theta): n \geq 1\}$ is stochastically equicontinuous on Θ.

Property U-WCON. $\sup_{\theta\in\Theta}|G_n(\theta)| \xrightarrow{p} 0$.

THEOREM 1. *(a) BD, P-WCON, & SE ⇒ U-WCON. (b) U-WCON ⇒ P-WCON & SE.*

Comments. 1. Theorem 1 is quite similar to Theorem 1 of Newey [15]. It does not require compactness of Θ, however, and it uses a somewhat different SE condition than that used by Newey [15]. See Section 4 of Newey [15] for two examples of the use of Theorem 1 in non-LLN contexts.

2. Theorem 1 can be obtained as a special case of a recent result of Pollard [16, Theorem 10.2] regarding the weak convergence of a sequence of stochastic processes to a limit process that is not necessarily degenerate. The proof of Theorem 1 given here, however, is very much simpler than that required for Pollard's result.

Proof of Theorem 1. For part (a), suppose BD, P-WCON, and SE hold. Given $\epsilon > 0$, take δ as in the definition of stochastic equicontinuity. Using

BD, let $\{B(\theta_j,\delta): j = 1,\ldots,J\}$ be a finite cover of Θ. Using P-WCON and SE,

$$\begin{aligned}\overline{\lim_{n\to\infty}} P\big(\sup_{\theta\in\Theta}|G_n(\theta)| > 2\epsilon\big) &\le \overline{\lim_{n\to\infty}} P\big(\max_{j\le J} \sup_{\theta'\in B(\theta_j,\delta)} (|G_n(\theta') - G_n(\theta_j)| \\ &\qquad + |G_n(\theta_j)|) > 2\epsilon\big) \\ &\le \overline{\lim_{n\to\infty}} P\big(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |G_n(\theta') - G_n(\theta)| > \epsilon\big) \\ &\quad + \overline{\lim_{n\to\infty}} P\big(\max_{j\le J}|G_n(\theta_j)| > \epsilon\big) < \epsilon. \qquad (2.1)\end{aligned}$$

For part (b), U-WCON ⇒ P-WCON is immediate and U-WCON ⇒ SE follows from

$$P\big(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |G_n(\theta') - G_n(\theta)| > \epsilon\big) \le P\big(2\sup_{\theta\in\Theta}|G_n(\theta)| > \epsilon\big) \to 0. \qquad \blacksquare \qquad (2.2)$$

Next, we consider "almost sure" results:

DEFINITION. *$\{G_n(\theta): n \ge 1\}$ is strongly stochastically equicontinuous on Θ if $\{\sup_{m\ge n}|G_m(\theta)| : n \ge 1\}$ is stochastically equicontinuous on Θ and $\sup_{n\ge 1}|G_n(\theta)| < \infty$ $\forall\theta \in \Theta$ a.s.* ■

The assumptions below use the abbreviations: P-SCON for pointwise strong (i.e., almost sure) convergence, SSE for strong stochastic equicontinuity, and U-SCON for uniform strong convergence.

Assumption P-SCON. $G_n(\theta) \to 0$ a.s. $\forall\theta \in \Theta$.

Assumption SSE. $\{G_n(\theta): n \ge 1\}$ is strongly stochastically equicontinuous on Θ.

Property U-SCON. $\sup_{\theta\in\Theta}|G_n(\theta)| \to 0$ a.s.

THEOREM 2. *(a) BD, P-SCON, & SSE ⇒ U-SCON. (b) U-SCON ⇒ P-SCON & SSE.*

Proof of Theorem 2. Theorem 2 follows from Theorem 1 and the well-known result (e.g., see Lukacs [13, Theorem 2.1.2, p. 30])

$$X_n \to 0 \text{ a.s.} \Leftrightarrow \sup_{m\ge n}|X_m| \xrightarrow{p} 0. \qquad \blacksquare \qquad (2.3)$$

We now introduce Lipschitz conditions that are sufficient for SE and SSE. For the special case of uniform LLN's, alternative sufficient conditions for SE and SSE are given in Section 3. The latter include both Lipschitz and non-Lipschitz conditions. (The "Lipschitz" conditions introduced here and in Section 3 are actually weaker than standard Lipschitz conditions, see Comment 2 following Theorem 3 below.)

Assumption SE-1

(a) $G_n(\theta) = \hat{Q}_n(\theta) - Q_n(\theta)$, where $Q_n(\theta)$ is a nonrandom function that is continuous in θ uniformly over $\theta \in \Theta$ and $n \geq 1$.
(b) $|\hat{Q}_n(\theta') - \hat{Q}_n(\theta)| \leq B_n h(d(\theta',\theta)) \forall \theta', \theta \in \Theta$ a.s. for some random variable B_n and some nonrandom function h such that $h(y) \downarrow 0$ as $y \downarrow 0$, where d is the metric on Θ.
(c) $B_n = O_p(1)$.

Assumption SSE-1

(a) SE-1(a) and (b) hold.
(b) $B_n = O(1)$ a.s.

A sufficient condition for SE-1(c) is $\sup_{n\geq 1} EB_n < \infty$. Sufficient conditions for SSE-1(b) are $\sup_{n\geq 1} EB_n < \infty$ and $B_n - EB_n \to 0$ a.s.

LEMMA 1. *(a) SE-1 ⇒ SE. (b) BD, P-SCON, & SSE-1 ⇒ SSE.*

Comment. If $G_n(\theta)$ satisfies a Lipschitz condition, one can take $Q_n(\theta) = 0$ in SE-1 and SSE-1.

Proof of Lemma 1. Let $1(\cdot)$ denote the indicator function, which equals one when $\cdot$ occurs and equals zero otherwise. BD, P-SCON, and SSE-1 imply $\sup_{n\geq 1}|G_n(\theta)| < \infty \ \forall \theta \in \Theta$ a.s. by writing $|G_n(\theta)| \leq |G_n(\theta) - G_n(\theta_j)| + |G_n(\theta_j)|$ for some θ_j in a suitably chosen finite set $\{\theta_1, \ldots, \theta_J\} \subset \Theta$. SSE-1 ⇒ SSE then follows from

$$\begin{aligned} &\overline{\lim_{n\to\infty}} P\Big(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} \Big|\sup_{m\geq n}|G_m(\theta')| - \sup_{m\geq n}|G_m(\theta)|\Big| > \epsilon\Big) \\ &\quad\leq \overline{\lim_{n\to\infty}} P\left(\sup_{m\geq n} B_m h(\delta) > \frac{\epsilon}{2}\right) \\ &\qquad + 1\left(\sup_{n\geq 1} \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |Q_n(\theta') - Q_n(\theta)| > \frac{\epsilon}{2}\right) < \epsilon \qquad (2.4) \end{aligned}$$

for δ sufficiently small since SSE-1(b) implies $\sup_{m\geq n} B_m = O_p(1)$. The proof of SE-1 ⇒ SE is analogous. ∎

3. GENERIC UNIFORM LAWS OF LARGE NUMBERS

3.1. Preliminaries

Throughout this section, we take $G_n(\theta) = (1/n)\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta))$, where Z_t is assumed to be a $\mathcal{Z}$-valued random variable (rv) for some measurable space $(\mathcal{Z},\mathcal{B})$ and $q_t(z,\theta)$ is a (Borel) measurable function from $\mathcal{Z}$ to R for each $\theta \in \Theta$ and each $t \geq 1$. For $G_n(\theta)$ as above, we give primitive conditions on $\{Z_t : t \geq 1\}$ and $\{q_t(z,\theta) : t \geq 1\}$ such that SE and SSE hold. Then, Theorems 1 and 2 yield generic weak and strong LLN's.

To illustrate what the function $q_t(Z_t,\theta)$ is in an example, consider the use of a uniform LLN in establishing the consistency of the least-squares (LS) estimator in a nonlinear regression model. The model is $Y_t = g(X_t,\theta_0) + U_t$ for $t = 1,\ldots,n$, where $Y_t \in R$ and $X_t \in R^p$ are observed, $g(\cdot,\cdot)$ is a known real-valued function, $\theta_0 \in \Theta \subset R^m$ is an unknown parameter, and U_t is an unobserved error with $E(U_t|X_t) = 0$ a.s. The LS estimator $\hat{\theta}_n$ of θ_0 minimizes $\Sigma_1^n(Y_t - g(X_t,\theta))^2$, or equivalently, minimizes $(1/n)\Sigma_1^n[(Y_t - g(X_t,\theta))^2 - U_t^2]$. In this case, we take $Z_t = (Y_t, X_t', U_t)'$ and $q_t(Z_t,\theta) = (Y_t - g(X_t,\theta))^2 - U_t^2 = (g(X_t,\theta) - g(X_t,\theta_0))^2 - 2(g(X_t,\theta) - g(X_t,\theta_0))U_t$. By applying the results of the present paper (or some other uniform LLN), one can establish the consistency of $\hat{\theta}_n$ when Θ is bounded (and not necessarily compact) and when the errors and regressors $(U_t, X_t')'$ are "asymptotically weakly dependent" nonidentically distributed rv's with the regressors being fixed or random (under additional assumptions to those outlined above). In comparison with the results of Jennrich [12] and Hannan [10], for example, one does not need identically distributed errors, weakly stationary errors, finite error variances, independence of the errors and regressors, or a compact parameter space.

For $G_n(\theta)$ as above, P-WCON, P-SCON, U-WCON, and U-SCON correspond to a pointwise weak LLN (P-WLLN), a pointwise strong LLN (P-SLLN), a uniform weak LLN (U-WLLN), and a uniform strong LLN (U-SLLN). For reasons of convention, we adopt the latter terminology in this section and define:

Assumption P-WLLN. $\frac{1}{n}\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta)) \xrightarrow{p} 0 \ \forall\theta \in \Theta.$

Assumption P-SLLN. $\frac{1}{n}\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta)) \to 0$ a.s. $\forall\theta \in \Theta.$

Property U-WLLN. $\sup_{\theta\in\Theta}\left|\frac{1}{n}\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta))\right| \xrightarrow{p} 0.$

Property U-SLLN. $\sup_{\theta\in\Theta}\left|\frac{1}{n}\Sigma_1^n(q_t(Z_t,\theta) - Eq_t(Z_t,\theta))\right| \to 0$ a.s.

For example, Assumptions P-WLLN and P-SLLN can be verified for dependent nonidentically distributed rv's using results of Andrews [2], Hannan [10], Hansen [11], or McLeish [14].

Next, we introduce two useful continuity properties of $(1/n)\Sigma_1^n Eq_t(Z_t,\theta)$ in θ. The second property is actually an asymptotic continuity (ACTY) property. One or another of these two properties is implied by each of the sufficient conditions given below for a U-WLLN or a U-SLLN.

Property CTY. $\frac{1}{n}\Sigma_1^n Eq_t(Z_t,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$ and $n \geq 1$.

Property ACTY. $\overline{\lim}_{n\to\infty} \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} \left| \frac{1}{n}\Sigma_1^n Eq_t(Z_t,\theta') - \frac{1}{n}\Sigma_1^n Eq_t(Z_t,\theta)\right| \to 0$ as $\delta \to 0$.

Property ACTY is weaker than CTY, since it equals CTY if $\overline{\lim}_{n\to\infty}$ is replaced by $\sup_{n\geq 1}$. Nevertheless, property ACTY is strong enough for use in consistency proofs of extremum estimators.

3.2. Generic Uniform LLN's Via a Lipschitz Condition

The first set of sufficient conditions for a U-WLLN and U-SLLN utilizes a Lipschitz condition, as in Andrews [1]:

Assumption W-LIP

(a) $|q_t(Z_t,\theta') - q_t(Z_t,\theta)| \leq B_t(Z_t)h(d(\theta',\theta)) \forall\theta', \theta \in \Theta$ a.s. for some measurable functions $\{B_t : t \geq 1\}$ and some nonrandom function h that satisfies $h(y) \downarrow 0$ as $y \downarrow 0$, where d is the metric on Θ.

(b) $\sup_{n\geq 1} \frac{1}{n}\Sigma_1^n EB_t(Z_t) < \infty$.

Assumption S-LIP

(a) W-LIP holds.

(b) $\frac{1}{n}\Sigma_1^n(B_t(Z_t) - EB_t(Z_t)) \to 0$ a.s.

LEMMA 2. *(a) W-LIP ⇒ SE & CTY. (b) BD, P-SLLN, & S-LIP ⇒ SSE & CTY.*

Proof of Lemma 2. S-LIP ⇒ W-LIP ⇒ CTY is immediate. Also, W-LIP ⇒ SE-1 and S-LIP ⇒ SSE-1 hold with $\hat{Q}_n(\theta) = (1/n)\Sigma_1^n q_t(Z_t,\theta)$, $Q_n(\theta) = (1/n)\Sigma_1^n Eq_t(Z_t,\theta)$, and $B_n = (1/n)\Sigma_1^n B_t(Z_t)$. Lemma 1 now gives the result. ■

Theorems 1 and 2 and Lemma 2 give the following U-WLLN and U-SLLN results:

THEOREM 3. *(a) BD, P-WLLN, & W-LIP ⇒ U-WLLN & CTY. (b) BD, P-SLLN, & S-LIP ⇒ U-SLLN & CTY.* ■

Comments. 1. Theorem 3 is similar to Corollary 2 of Andrews [1] except that Θ is allowed to be totally bounded rather than compact and P-WLLN, P-SLLN, and S-LIP(b) arise in place of conditions of pointwise weak and strong LLN's for the supremum and infimum of $q_t(Z_t,\theta)$ over small neighborhoods of θ $\forall\theta \in \Theta$. The conditions P-WLLN, and so on, are neater and more elegant than the latter, but are usually not any more general. Theorem 3(a) is also very similar to Corollary 3.1 of Newey [15]. It differs only in that it allows Θ to be totally bounded, whereas Newey's U-WLLN requires Θ to be compact.

2. In spite of the description of W-LIP and S-LIP as Lipschitz conditions, these conditions are not smoothness conditions. The reason is that $h(\cdot)$ is arbitrary (provided $h(y) \downarrow 0$ as $y \downarrow 0$). For example, W-LIP and S-LIP are implied by continuity of $q_t(z,\theta)$ in θ uniformly over $\theta \in \Theta$, $z \in \mathcal{Z}$, and $t \geq 1$ and the latter is not a smoothness condition.[2] (A consequence of this result is that when Θ and $\mathcal{Z}$ are compact, Pötscher and Prucha's [18] equicontinuity condition on $\{q_t(z,\theta) : t \geq 1\}$ implies W-LIP and S-LIP and, in turn, is close to being implied by the latter. This is not true, of course, in the important case of noncompact $\mathcal{Z}$.)

3. In many cases, uniform LLN's are used in proofs of consistency and asymptotic normality of extremum estimators. Standard proofs of the asymptotic normality of such estimators rely on the differentiability of the functions $\{q_t(z,\theta) : t \geq 1\}$ in θ and on a uniform WLLN for the Hessian matrix $\{(\partial^2/\partial\theta\partial\theta')q_t(z,\theta) : t \geq 1\}$. An attribute of the conditions W-LIP and S-LIP on $\{q_t(z,\theta) : t \geq 1\}$ is that they are implied by such assumptions. Some other sets of sufficient conditions for uniform LLN's, including those of Pötscher and Prucha [18] and Assumption TSE-2A below, impose continuity assumptions on $q_t(z,\theta)$ in (z,θ) that are not implied by the differentiability in θ assumptions used to obtain asymptotic normality, although these continuity conditions may be (but are not necessarily) implied by the assumptions used to obtain a uniform WLLN for the Hessian matrix.

With regard to the choice of assumptions to establish uniform convergence of the Hessian matrix, neither a Lipschitz condition nor an equicontinuity condition in (z,θ) is unambiguously superior. These two conditions provide different tradeoffs in the assumptions.

3.3 Generic Uniform WLLN's via Termwise Stochastic Equicontinuity

Here we introduce a condition called termwise stochastic equicontinuity (TSE) which, together with a domination (DM) condition, is sufficient for SE. The TSE condition is not primitive, since it involves an interaction between the functions $\{q_t(z,\theta) : t \geq 1\}$ and the marginal distributions of

[2] To see this, suppose $q_t(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$, $z \in \mathcal{Z}$, and $t \geq 1$. Then, given $\epsilon > 0$ $\exists\delta(\epsilon) > 0$ such that $\sup_{t\geq 1} \sup_{z\in\mathcal{Z}} |q_t(z,\theta') - q_t(z,\theta)| < \epsilon$ $\forall\theta' \in B(\theta,\delta(\epsilon))$, $\forall\theta \in \Theta$. We can take $\delta(\epsilon)$ such that $\delta(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$. Define the inverse function $\delta^{-1}(y) = \inf\{\epsilon > 0 : \delta(\epsilon) \geq y\}$. Note that $\epsilon \leq \delta^{-1}(\delta(\epsilon))$. Now, given any θ', $\theta \in \Theta$, take $\epsilon > 0$ such that $d(\theta',\theta) \in [\delta(\epsilon/2), \delta(2\epsilon)]$. Then, $\sup_{t\geq 1} \sup_{z\in\mathcal{Z}} |q_t(z,\theta') - q_t(z,\theta)| < 2\epsilon \leq 4\delta^{-1}(\delta(\epsilon/2)) \leq 4\delta^{-1}(d(\theta',\theta))$. Hence, W-LIP and S-LIP hold with $B_t(Z_t) = 4$ and $h(y) = \delta^{-1}(y)$.

$\{Z_t : t \geq 1\}$. It is easy to obtain primitive sufficient conditions that imply TSE, however, and a variety is given below.

Assumption TSE. $\lim_{\delta\to 0} \overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n P(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |q_t(Z_t,\theta') - q_t(Z_t,\theta)| > \epsilon) = 0$, $\forall \epsilon > 0$.

Assumption DM. $\lim_{M\to\infty} \overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n E d_t(Z_t) 1(d_t(Z_t) > M) = 0$ for some nonrandom measurable functions $\{d_t : t \geq 1\}$ that satisfy $d_t(z) \geq \sup_{\theta\in\Theta} |q_t(z,\theta)| \, \forall z \in \mathcal{Z}$, $\forall t \geq 1$.

Assumption DM is implied by $\overline{\lim}_{n\to\infty}(1/n)\Sigma_1^n E d_t^\gamma(Z_t) < \infty$ for some $\gamma > 1$. Also, for identically distributed $\{Z_t : t \geq 1\}$, if $d_t(z)$ can be taken to be independent of t (as occurs, for example, if $q_t(z,\theta)$ does not depend on t), then DM is implied by $Ed(Z_t) < \infty$.

LEMMA 3. *DM & TSE ⇒ SE & ACTY.*

Proof of Lemma 3. Let $Y_{t\delta} = \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |q_t(Z_t,\theta') - q_t(Z_t,\theta)|$ and $d_t = d_t(Z_t)$. Given $\epsilon > 0$, take $M < \infty$ and $\delta > 0$ such that $\overline{\lim}_{n\to\infty}(2/n)\Sigma_1^n E d_t 1(d_t > M/2) < \epsilon^2/6$ and $\overline{\lim}_{n\to\infty}(1/n)\Sigma_1^n P(Y_{t\delta} > \epsilon^2/6) < \epsilon^2/(6M)$. Then, we have

$$\lim_{n\to\infty} P(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |G_n(\theta') - G_n(\theta)| > \epsilon)$$

$$\leq \overline{\lim}_{n\to\infty} P\left(\frac{1}{n} \Sigma_1^n (Y_{t\delta} + EY_{t\delta}) > \epsilon\right)$$

$$\leq \frac{2}{\epsilon} \overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n EY_{t\delta}$$

$$= \frac{2}{\epsilon} \overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n \left(EY_{t\delta} 1\left(Y_{t\delta} \leq \frac{\epsilon^2}{6}\right) + EY_{t\delta} 1\left(\frac{\epsilon^2}{6} < Y_{t\delta} \leq M\right) + EY_{t\delta} 1(Y_{t\delta} > M)\right)$$

$$\leq \frac{2}{\epsilon}\left(\frac{\epsilon^2}{6} + \overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n MP\left(Y_{t\delta} > \frac{\epsilon^2}{6}\right) + \overline{\lim}_{n\to\infty} \frac{2}{n} \Sigma_1^n E d_t 1(2d_t > M)\right) \leq \epsilon. \tag{3.1}$$

Thus, SE holds. Since $\overline{\lim}_{n\to\infty}(1/n)\Sigma_1^n EY_{t\delta} < \epsilon^2/2$, ACTY also holds. ■

Theorem 1 and Lemma 3 combine to give the following U-WLLN result:

THEOREM 4. *BD, P-WLLN, DM, & TSE ⇒ U-WLLN & ACTY.* ■

We now introduce a series of assumptions that are primitive and that imply TSE. These assumptions will be shown to satisfy

$$
\begin{array}{ccccccc}
 & & & \nearrow & \text{TSE-1A} & \searrow & \\
\text{TSE-1D} & \rightarrow & \text{TSE-1C} & & & & \text{TSE-1}\\
 & & & \searrow & \text{TSE-1B} & \nearrow & \\
 & & & & & & \searrow \\
 & & & & \text{TSE-2A} \rightarrow & \text{TSE-2} \rightarrow & \text{TSE}\\
 & & & & & \text{W-LIP} & \nearrow
\end{array} \qquad (3.2)
$$

when Θ is totally bounded.

Let μ_t denote the probability distribution or distribution function of Z_t for $t \geq 1$. Let $\bar{\mu}_n = (1/n)\Sigma_1^n \mu_t$ for $n \geq 1$.

Assumption TSE-1

(a) $q_t(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$ and $t \geq 1$ $\forall z \in \mathcal{Z}$.
(b) For every sequence of measurable sets $\{A_m \subset \mathcal{Z} : m \geq 1\}$ such that $A_m \downarrow \phi$ as $m \to \infty$, $\overline{\lim}_{n\to\infty} \frac{1}{n} \Sigma_1^n P(Z_t \in A_m) \to 0$ as $m \to \infty$, where ϕ denotes the null set.

Assumption TSE-1A

(a) $q_t(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$ and $t \geq 1$ $\forall z \in \mathcal{Z}$.
(b) μ_t is absolutely continuous with respect to some σ-finite measure μ $\forall t \geq 1$ and $\int [\sup_{n\geq 1}(1/n)\Sigma_1^n f_t(z)]\, d\mu(z) < \infty$, where $f_t(z)$ denotes the density of μ_t with respect to μ.

Assumption TSE-1B

(a) $q_t(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$ and $t \geq 1$ $\forall z \in \mathcal{Z}$.
(b) $\bar{\mu}_n \to \mu$ properly setwise for some probability measure μ, that is, for all measurable sets $A \subset \mathcal{Z}$, $(1/n)\Sigma_1^n P(Z_t \in A) \to \mu(A)$.

Assumption TSE-1C

(a) $q_t(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$ and $t \geq 1$ $\forall z \in \mathcal{Z}$.
(b) $\{Z_t : t \geq 1\}$ are identically distributed.

Assumption TSE-1D

(a) $q_t(z,\theta) = q(z,\theta)$ $\forall t \geq 1$ and $q(z,\theta)$ is continuous in θ $\forall \theta \in \Theta$, $\forall z \in \mathcal{Z}$.
(b) Θ is compact.
(c) $\{Z_t : t \geq 1\}$ are identically distributed.

Assumption TSE-2

(a) $q_t(z,\theta) = \sum_{k=1}^K r_{kt}(z)s_{kt}(z,\theta)$, $r_{kt}(z)$ and $s_{kt}(z,\theta)$ are measurable in z $\forall \theta \in \Theta$, $\forall 1 \leq k \leq K$, $\forall t \geq 1$ and $s_{kt}(z,\theta)$ is continuous in θ uniformly over $\theta \in \Theta$, $z \in C_j$, and $t \geq 1$, $\forall 1 \leq k \leq K$, $\forall j \geq 1$, where the sequence of measurable sets $\{C_j \subseteq \mathcal{Z} : j \geq 1\}$ satisfies $\lim_{j\to\infty} \inf_{n\geq 1}(1/n)\Sigma_1^n P(Z_t \in C_j) = 1$.
(b) $\sup_{n\geq 1}(1/n)\Sigma_1^n E|r_{kt}(Z_t)| < \infty$ $\forall 1 \leq k \leq K$.

Assumption TSE-2A

(a) $q_t(z,\theta) = \Sigma_{k=1}^K r_{kt}(z)s_{kt}(z,\theta)$, $r_{kt}(z)$, and $s_{kt}(z,\theta)$ are measurable in z $\forall\theta \in \Theta$, $\forall 1 \le k \le K$, $\forall t \ge 1$, and $\{s_{kt}(z,\theta): t \ge 1\}$ is equicontinuous on $\mathcal{Z} \times \Theta$ $\forall 1 \le k \le K$, i.e., $\sup_{t\ge 1}|s_{kt}(z',\theta') - s_{kt}(z,\theta)| \to 0$ as $(z',\theta') \to (z,\theta) \forall (z,\theta) \in \mathcal{Z} \times \Theta$, $\forall 1 \le k \le K$.
(b) Θ is compact.
(c) TSE-2(b) holds and $\{\bar{\mu}_n : n \ge 1\}$ is tight (i.e., $\lim_{j\to\infty} \inf_{n\ge 1}(1/n)\Sigma_1^n P(Z_t \in C_j) = 1$ for some sequence of compact sets $\{C_j \subset \mathcal{Z} : j \ge 1\}$).

The condition TSE-1 has not been considered previously in the literature. We view it as a third alternative to the Lipschitz condition of Andrews [1] (i.e., W-LIP or S-LIP) and the equicontinuity condition of Pötscher and Prucha [18] (i.e., TSE-2 or TSE-2A). Roughly speaking, TSE-1(b) requires that the standard continuity property of the distribution of an rv Z_t (i.e., $A_m \downarrow \phi \Rightarrow P(Z_t \in A_m) \downarrow 0$ as $m \to \infty$) holds uniformly over $t \ge 1$. Of course, with identically distributed rv's, this condition holds automatically. Also, in the case where $\{Z_t : t \ge 1\}$ takes on at most a finite number of different distributions (e.g., when "fixed in repeated samples" asymptotics are used), this condition automatically holds. As mentioned in the introduction, an attribute of TSE-1 is that it applies with d.n.i.d. rv's, yet when the rv's are identically distributed, it yields the best available sufficient conditions for a generic U-WLLN.

On the other hand, TSE-1 is considerably more restrictive than W-LIP or TSE-2 with regard to the degree of nonidentical distributions of $\{Z_t : t \ge 1\}$ that it allows. Pötscher and Prucha [20, Remark 4.5] provide the following example where TSE-1(b) fails. Suppose $P(Z_t = a_t) \ge \epsilon > 0$ $\forall t \ge 1$ for some constants $\{a_1, a_2, \ldots,\}$ such that $a_s \ne a_t$ for $s \ne t$. By taking $A_m = \{a_m, a_{m+1}, \ldots\}$, one sees that TSE-1(b) fails. This example shows that there may be many circumstances where TSE-1 is too restrictive and that W-LIP or TSE-2 is more appropriate.

TSE-1D corresponds to Jennrich's [12, Theorem 2] conditions for a uniform LLN for i.i.d. rv's.[3] TSE-1C generalizes the Jennrich-type conditions to allow for a totally bounded parameter space, functions $q_t(z,\theta)$ that depend on t, and rv's $\{Z_t : t \ge 1\}$ that are non-Euclidean-valued. A nice feature of TSE is that the sufficiency of TSE-1C and TSE-1D for TSE follows immediately by the monotone convergence theorem. TSE-1B corresponds to Bierens' [5, Lemma 3] conditions for a uniform LLN (but generalized in the same way that TSE-1C generalizes TSE-1D).[4] TSE-1A is a new condition. It does not require convergence of the average distributions of $\{Z_t\}$ as in TSE-1B.

[3]We note that Jennrich [12] also considers nonidentically distributed rv's in his treatment of the fixed regressor nonlinear regression model. For this model, he makes use of the concept of tail products.
[4]As stated, Lemma 3 of Bierens [5] only assumes proper convergence of $\bar{\mu}_n$ to μ. As noted in Bierens [6, Remark 2 to Theorem 2.6], however, proper setwise convergence is actually needed, as in TSE-1B.

TSE-2A corresponds to Pötscher and Prucha's [18] conditions for a generic uniform LLN, while TSE-2 has been considered in Pötscher and Prucha [17, conditions I and IV].[5]

LEMMA 4. *Suppose BD holds. Then,*

(a) TSE-1D → TSE-1C ↗ TSE-1A ↘ ; ↘ TSE-1B ↗ TSE-1 → TSE.

(b) TSE-2A → TSE-2 → TSE.

(c) W-LIP → TSE.

Comments. 1. TSE-1 is weaker than TSE-2 with regard to the assumptions on $\{q_t(z,\theta): t \geq 1\}$. It is stronger than TSE-2, however, regarding the assumptions on $\{Z_t: t \geq 1\}$. TSE-2 uses the additional assumptions on $\{q_t(z,\theta): t \geq 1\}$ to identify a particular sequence of sets $\{A_m: m \geq 1\}$ that satisfies $A_m \downarrow \phi$ as $m \to \infty$ and for which $\overline{\lim}_{n\to\infty}(1/n)\Sigma_1^n P(Z_t \in A_m) \to 0$ as $m \to \infty$ implies TSE (see the proof of Lemma 4).

2. Part (a) of TSE-1B, TSE-1C, and TSE-1D only needs to hold a.s. $[\mu_s]$ $\forall s \geq 1$, rather than $\forall z \in \mathcal{Z}$, in order for each to imply TSE. (Although Lemma 4(a) no longer holds with this change.)

Proof of Lemma 4. TSE-1D ⇒ TSE-1C, TSE-1C ⇒ TSE-1B, and TSE-1C ⇒ TSE-1A are immediate. TSE-1B ⇒ TSE-1 follows from

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P(Z_t \in A_m) = \mu(A_m) \to 0 \quad \text{as} \quad m \to \infty \tag{3.3}$$

whenever $A_m \in \mathcal{B}$ $\forall m \geq 1$ and $A_m \downarrow \phi$ as $m \to \infty$. TSE-1A ⇒ TSE-1 follows from

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P(Z_t \in A_m) \leq \int 1(z \in A_m)\left[\sup_{n\geq 1} \frac{1}{n} \Sigma_1^n f_t(z)\right] d\mu(z) \to 0 \quad \text{as} \quad m \to \infty \tag{3.4}$$

whenever $A_m \in \mathcal{B}$ $\forall m \geq 1$ and $A_m \downarrow \phi$ as $m \to \infty$ by the dominated convergence theorem. TSE-1 ⇒ TSE is obtained by setting $A_m = \{z \in \mathcal{Z}: \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,1/m)} \sup_{t\geq 1} |q_t(z,\theta') - q_t(z,\theta)| > \epsilon\}$. By BD and TSE-1(a), A_m is measurable $\forall m \geq 1$ (this is the only place BD is used in Lemma 4) and by TSE-1(a), $A_m \downarrow \phi$ as $m \to \infty$. Hence, TSE-1(b) gives the desired result.

TSE-2A ⇒ TSE-2 with the sets $\{C_j: j \geq 1\}$ of TSE-2(a) given by those of TSE-2A(c), since continuous functions on compact sets are uniformly continuous. To show TSE-2 ⇒ TSE, take $K = 1$ *w*log. Let $Y_{t\delta}(z) = \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |s_{kt}(z,\theta') - s_{kt}(z,\theta)|$. Given $\epsilon > 0$, take j and $\delta > 0$ such that

[5]An error in the statement of TSE-2 and TSE-2A in an earlier version of this paper was pointed out by Pötscher and Prucha [20]. These assumptions have been altered to correct this error. See Pötscher and Prucha [20, Assumption 4.1] for a close analog of TSE-2.

$\overline{\lim}_{n\to\infty}(1/n)\Sigma_1^n P(Z_t \notin C_j) < \epsilon/2$ and $\sup_{z\in C_j}\sup_{t\geq 1} Y_{t\delta}(z) < \epsilon^2/(4R)$, where $R = \sup_{n\geq 1}(1/n)\Sigma_1^n E|r_{kt}(Z_t)|$. Then, we have

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P\big(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |q_t(Z_t,\theta') - q_t(Z_t,\theta)| > \epsilon\big)$$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P\left(|r_{kt}(Z_t)| Y_{t\delta}(Z_t) 1(Z_t \in C_j) > \frac{\epsilon}{2}\right)$$

$$+ \overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P\left(|r_{kt}(Z_t)| Y_{t\delta}(Z_t) 1(Z_t \notin C_j) > \frac{\epsilon}{2}\right)$$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P\left(|r_{kt}(Z_t)| \frac{\epsilon^2}{4R} > \frac{\epsilon}{2}\right) + \overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n P(Z_t \notin C_j)$$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \Sigma_1^n E|r_{kt}(Z_t)| \frac{\epsilon}{2R} + \frac{\epsilon}{2} \leq \epsilon. \tag{3.5}$$

Lastly, W-LIP ⇒ TSE is immediate by substitution. ■

Theorem 1 and Lemmas 3 and 4 combine to give the following U-WLLN result:

THEOREM 5. *BD, P-WLLN, DM, & any one of the TSE conditions ⇒ U-WLLN & ACTY.* ■

3.4. Generic Uniform SLLN's via Termwise Stochastic Equicontinuity

To obtain a uniform strong LLN via the TSE condition, the following additional condition is needed.

Assumption P-SLLN2. $(1/n)\Sigma_1^n(Y_{t\delta_j} - EY_{t\delta_j}) \to 0$ a.s. $\forall j \geq 1$, where $\{\delta_j : j \geq 1\}$ is some sequence of positive constants that converges to zero as $j \to \infty$,

$$Y_{t\delta} = \sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |q_t(Z_t,\theta') - q_t(Z_t,\theta)|$$

and $Y_{t\delta_j}$ is measurable $\forall j \geq 1$.

LEMMA 5. *BD, P-SLLN, DM, TSE, & P-SLLN2 ⇒ SSE & ACTY.*

Proof of Lemma 5. ACTY holds by Lemma 3. Next we show that $\sup_{n\geq 1}|G_n(\theta)| < \infty$ $\forall\theta \in \Theta$ a.s. Using BD, let $\{\theta_1, \ldots, \theta_J\}$ denote the centers of a finite number of δ-balls that cover Θ for $\delta > 0$. Given $\theta \in \Theta$, let j be such that $\theta \in B(\theta_j, \delta)$. Write

$$\begin{aligned}\left|\frac{1}{n}\Sigma_1^n(q_t(\theta) - Eq_t(\theta))\right| &\le \left|\frac{1}{n}\Sigma_1^n(q_t(\theta_j) - Eq_t(\theta_j))\right| \\ &\quad + \left|\frac{1}{n}\Sigma_1^n(q_t(\theta) - q_t(\theta_j))\right| \\ &\quad + \left|\frac{1}{n}\Sigma_1^n(Eq_t(\theta) - Eq_t(\theta_j))\right| \\ &\le \left|\frac{1}{n}\Sigma_1^n(q_t(\theta_j) - Eq_t(\theta_j))\right| \\ &\quad + \frac{1}{n}\Sigma_1^n(Y_{t\delta} - EY_{t\delta}) + \frac{2}{n}\Sigma_1^n EY_{t\delta},\end{aligned}$$

where $q_t(\theta)$ abbreviates $q_t(Z_t, \theta)$. P-SLLN, P-SLLN2, and (3.1) imply that the sup over $n \ge 1$ of the right-hand side is finite a.s. as desired.

Now we show that $\{\sup_{m\ge n}|G_m(\theta)| : n \ge 1\}$ is stochastically equicontinuous. Given $\epsilon \in (0,\frac{1}{2})$, take δ as in the proof of Lemma 3 and such that $\delta = \delta_j$ for some j. Then, we have

$$\begin{aligned}&\overline{\lim_{n\to\infty}} P\Big(\sup_{\theta\in\Theta}\ \sup_{\theta'\in B(\theta,\delta)} \big|\sup_{m\ge n}|G_m(\theta')| - \sup_{m\ge n}|G_m(\theta)|\big| > \epsilon\Big) \\ &\le \overline{\lim_{n\to\infty}} P\Big(\sup_{\theta\in\Theta}\ \sup_{\theta'\in B(\theta,\delta)}\ \sup_{m\ge n}|G_m(\theta') - G_m(\theta)| > \epsilon\Big) \\ &\le \overline{\lim_{n\to\infty}} P\left(\sup_{m\ge n}\frac{1}{m}\Sigma_1^m(Y_{t\delta} + EY_{t\delta}) > \epsilon\right) \\ &\le \overline{\lim_{n\to\infty}} P\left(\sup_{m\ge n}\left|\frac{1}{m}\Sigma_1^m(Y_{t\delta} - EY_{t\delta})\right| > \frac{\epsilon}{2}\right) \\ &\quad + \overline{\lim_{n\to\infty}}\, 1\left(\sup_{m\ge n}\frac{1}{m}\Sigma_1^m 2EY_{t\delta} > \frac{\epsilon}{2}\right) \\ &= 0 + 1\left(\overline{\lim_{n\to\infty}}\frac{1}{n}\Sigma_1^n EY_{t\delta} > \frac{\epsilon}{4}\right) = 0, \qquad \textbf{(3.6)}\end{aligned}$$

where the second last equality uses the characterization of a.s. convergence referred to in Section 2 and the last equality uses (3.1) of the proof of Lemma 3. ■

Theorem 2 and Lemmas 4 and 5 yield the following U-SLLN result:

THEOREM 6. *BD, P-SLLN, DM, any one of the TSE conditions, & P-SLLN2 ⇒ U-SLLN & ACTY.*

Comments. 1. Since Theorem 6 imposes Assumption P-SLLN2, its pointwise SLLN assumptions are no simpler than those of Andrews [1] and Pötscher and Prucha [18]. For a uniform SLLN that avoids imposing P-SLLN2 and uses an alternative to TSE, see Pötscher and Prucha [20, Corollary 4.3].

2. Theorem 6 adds the following to the literature on generic uniform strong LLN's. First, it gives generic U-SLLN's that do not require the parameter space to be compact. Second, it gives a third alternative sufficient condition (viz., TSE-1) to the Lipschitz and equicontinuity conditions of Andrews [1] and Pötscher and Prucha [18]. ■

The results of this section have been concerned with sequences of rv's $\{Z_t : t \geq 1\}$ and functions $\{q_t(z,\theta) : t \geq 1\}$. They can be extended straightforwardly to triangular arrays $\{Z_{nt} : t \leq n,\ n \geq 1\}$ and $\{q_{nt}(z,\theta) : t \leq n,\ n \geq 1\}$ by adding the subscript n in the appropriate places. (Note, however, that pointwise strong LLN's hold for triangular arrays only under quite restrictive assumptions.) For example, uniform LLN's for triangular arrays are useful for establishing consistency and asymptotic normality results in the context of sequences of local alternatives.

REFERENCES

1. Andrews, D.W.K. Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica* 55 (1987):1465–1471.
2. Andrews, D.W.K. Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4 (1988):458–467.
3. Andrews, D.W.K. Asymptotics for semiparametric econometric models: I. Estimation and testing. Cowles Foundation Discussion Paper No. 908R, Yale University, 1988.
4. Andrews, D.W.K. Asymptotics for semiparametric econometric models: II. Stochastic equicontinuity and nonparametric kernel estimation. Cowles Foundation Discussion Paper No. 909R, Yale University, 1989.
5. Bierens, H.J. Model specification testing of time series regressions. *Journal of Econometrics* 26 (1984):323–353.
6. Bierens, H.J. "Least Squares Estimation of Linear and Nonlinear ARMAX Models under Data Heterogeneity," unpublished manuscript, Department of Econometrics, Free University, Amsterdam.
7. Billingsley, P. *Convergence of Probability Measures.* Wiley: New York, 1968.
8. Gallant, A.R. *Nonlinear Statistical Models.* Wiley: New York, 1987.
9. Gallant, A.R. & H. White. *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* Basil Blackwell: Oxford, U.K., 1988.
10. Hannan, E.J. Nonlinear time series regression. *Journal of Applied Probability* 8 (1971):767–780.
11. Hansen, B.E. Strong laws for dependent heterogeneous processes. *Econometric Theory* 7 (1991):213–221.
12. Jennrich, R.I. Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics* 40 (1969):633–643.
13. Lukacs, E. *Stochastic Convergence*, 2nd ed. (1969): Academic Press: New York, 1975.
14. McLeish, D.L. A maximal inequality and dependent strong laws. *Annals of Probability* 3 (1975):829–839.

15. Newey, W.K. Uniform convergence in probability and stochastic equicontinuity. *Econometrica* 59 (1991):1161-1167.
16. Pollard, D. *Empirical Processes: Theory and Applications.* CBMS Regional Conference Series in Probability and Statistics, Volume 2, Institute of Mathematical Statistics: Hayward, CA, 1990.
17. Pötscher, B.M. & I.R. Prucha. Consistency in nonlinear econometrics: A generic uniform law of large numbers and some comments on recent results, working paper no. 86-9, Department of Economics, University of Maryland, 1986.
18. Pötscher, B.M. & I.R. Prucha. A uniform law of large numbers for dependent and heterogeneous data processes. *Econometrica* 57 (1989):675-683.
19. Pötscher, B.M. & I.R. Prucha. Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I: Consistency and approximation concepts. *Econometric Reviews* 10 (1991):125-216.
20. Pötscher, B.M. & I.R. Prucha. Generic uniform convergence and equicontinuity concepts for random functions: An exploration of the basic structure. Working paper no. 90-14, Department of Economics, University of Maryland, 1990.
21. Prohorov, Yu. V. Convergence of random processes and limit theorems in probability theory. *Theory of Probability and Its Applications* 1 (1956):157-214.