COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station

New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 907

AN EMPIRICAL PROCESS CENTRAL LIMIT THEOREM FOR DEPENDENT NON-IDENTICALLY DISTRIBUTED RANDOM VARIABLES

by

Donald W.K. Andrews

November 1988

Revised: May 1989

AN EMPIRICAL PROCESS CENTRAL LIMIT THEOREM FOR DEPENDENT NON-IDENTICALLY DISTRIBUTED RANDOM VARIABLES

Donald W. K. Andrews[1]
Cowles Foundation
Yale University

November, 1988

Revised: May, 1989

ABSTRACT

This paper establishes a central limit theorem (CLT) for empirical processes indexed by smooth functions. The underlying random variables may be temporally dependent and non-identically distributed. In particular, the CLT holds for near epoch dependent (i.e., functions of mixing processes) triangular arrays, which include strong mixing arrays, among others. The results apply to classes of functions that have series expansions. The proof of the CLT is particularly simple; no chaining argument is required. The results can be used to establish the asymptotic normality of semiparametric estimators in time series contexts. An example is provided.

JEL Classification Number: 211

Keywords: Central limit theorem, empirical process, Fourier series, near epoch dependence, semiparametric estimator, series expansion, Sobolev norm, stochastic equicontinuity, strong mixing.

1. INTRODUCTION

This paper establishes a central limit theorem (CLT) for empirical processes indexed by classes of smooth functions. The contribution of the paper is twofold. First, the results improve existing results by allowing the underlying random variables (rv's) to possess a more general form of temporal dependence and non–identical distributions than is available elsewhere. Second, the proof of the results is very simple relative to proofs in the literature.

The results of this paper apply to near epoch dependent (NED) (i.e., functions of mixing processes) triangular arrays of rv's. In contrast, existing results in the empirical process literature only consider independent rv's, while results in the Banach space–valued CLT literature only consider weakly stationary strong mixing sequences of rv's (e.g., see Dehling and Philipp (1982) and Dehling (1983)). As discussed below, NED triangular arrays include strong mixing triangular arrays, as well as a rich class of non–strong mixing triangular arrays.

The approach used here is to consider an index class of functions that have series expansions with respect to the same set of series functions. A prime example is a class of functions that have Fourier series expansions. Conditions are placed on the coefficients of the series expansions that are sufficient to obtain stochastic equicontinuity of the empirical process and a CLT for it. In particular applications, these conditions can be verified by using known results for series expansions. The case of differentiable functions on an open bounded subset of R^k (whose boundary is minimally smooth) is examined in detail. In this application, the required number of finite derivatives of the functions is the same as for Ossiander's (1987) results for iid rv's obtained using a metric entropy with bracketing condition.[2]

Note that the series expansion approach used here is different from approaches currently used in the empirical process literature, but is similar to an approach used in the

Banach space-valued CLT literature (see the references above).

An appealing feature of the series expansion approach is the simplicity with which one obtains the stochastic equicontinuity and CLT results. Elementary manipulations and inequalities suffice (see the proof of Theorem 1). No chaining argument is required. In consequence, the series expansion approach is conducive to obtaining results under quite weak assumptions regarding temporal dependence and non-identical distributions (as noted above). The drawback of this approach is that the variety of different classes of functions to which it applies is more restricted than with a bracketing approach (e.g., see Ossiander (1987) or Pollard (1989)) or a Vapnik–Cervonenkis approach (e.g., see Pollard (1988)).

The empirical process results of this paper have numerous applications in econometrics and statistics. For example, they are tailor-made for use with the results of Andrews (1989a,b,c), which establish the $\sqrt{n}$–consistency and asymptotic normality of semiparametric and parametric estimators and the asymptotic chi–squared distributions of Wald, Lagrange multiplier, and likelihood ratio–like test statistics that correspond to such estimators. Other applications of empirical processes to parametric estimators include those of Pollard (1984, 1985, 1988), Pakes and Pollard (1986), and Shorack and Wellner (1986). For applications of empirical processes to goodness of fit tests, see Pollard (1979), Shorack and Wellner (1986), and Andrews (1988a, b).

Note that empirical process results for triangular arrays of rv's, rather than sequences, are needed for applications to testing problems in which sequences of local alternatives are considered. For this reason, we deviate in this paper from the bulk of the empirical process literature and treat triangular arrays rather than sequences.

The remainder of this paper is organized as follows: Section 2 provides the main CLT results. Section 3 applies the CLT results to the case of indexing by a class of differentiable functions defined on an open bounded subset of R^k whose boundary is minimally smooth. Section 4 illustrates the use of the results in a semiparametric estimation problem.

2. AN EMPIRICAL PROCESS CLT

2.1. *Stochastic Equicontinuity*

Let $\{X_{ni} : 1 \leq i \leq n, n \geq 1\}$ be a triangular array of $\mathcal{X}$–valued random vectors (rv's) defined on a probability space $(\Omega,\mathcal{B},P)$, where $\mathcal{X} \subset R^k$. Let $m(\cdot,\cdot)$ be a real function defined on $\mathcal{X} \times \mathcal{T}$, where $\mathcal{T}$ is an index set that is a metric space with metric ρ (defined below). Assume $m(x,\tau)$ is Borel measurable in x for each fixed $\tau \in \mathcal{T}$. Define the empirical process $\nu_n(\cdot)$ by

$$\nu_n(\tau) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(m(X_{ni}, \tau) - Em(X_{ni}, \tau)) . \tag{2.1}$$

The process $\nu_n(\cdot)$ converges weakly to a Gaussian process if it is stochastically equicontinuous, if its finite dimensional distributions are asymptotically normal, and if $\mathcal{T}$ is totally bounded (see Pollard (1988, Weak Convergence Theorem) for a proof of this result). In view of this result, we give conditions in this subsection under which $\{\nu_n(\cdot) : n \geq 1\}$ is stochastically equicontinuous.

DEFINITION. $\{\nu_n(\cdot) : n \geq 1\}$ *is stochastically equicontinuous if* $\forall \epsilon > 0$ *and* $\eta > 0$, $\exists \delta > 0$ *such that*

$$\overline{\lim_{n\to\infty}}\, P^*(\sup_{\rho(\tau,\gamma)<\delta} |\nu_n(\tau) - \nu_n(\gamma)| > \eta) < \epsilon , \tag{2.2}$$

where P^* *denotes* P*–outer probability.*

The following implies $\{\nu_n(\cdot)\}$ is stochastically equicontinuous:

ASSUMPTION A. (i) (Series expansion) *For some sequence* $\{h_j(\cdot) : j \geq 1\}$ *of real or complex Borel measurable functions on* $\mathcal{X}$, $m(\cdot,\tau)$ *has a pointwise convergent series expansion for each* $\tau \in \mathcal{T}$: $m(x,\tau) = \sum_{j=1}^{\infty} c_j(\tau)h_j(x)$ $\forall x \in \mathcal{X}$, *where for each* $\tau \in \mathcal{T}$ $\{c_j(\tau) : j \geq 1\}$ *is a sequence of (real or complex) constants.*

(ii) (Smoothness) $\sum_{j=1}^{\infty} |c_j(\tau)| E|h_j(X_{ni})| < \infty$ $\forall i \leq n$, $n \geq 1$, $\tau \in \mathcal{T}$.

(iii) (Smoothness/weak dependence tradeoff) $\sup_{\tau \in \mathcal{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2/a_j \to 0$ *as* $J \to \infty$ *for some summable sequence of positive real constants* $\{a_j\}$ *for which* $\sum_{j=1}^{\infty} a_j\gamma_j < \infty$, *where* $\gamma_j = \sum_{s=-\infty}^{\infty} \gamma_j(s)$ *and* $\gamma_j(s) = \sup_{i \leq n-|s|, n \geq 1} |\mathrm{Cov}(h_j(X_{ni}), h_j(X_{ni+|s|}))|$.

In Section 3 below, Assumption A is verified with the functions $\{h_j(\cdot)\}$ being complex exponential functions. With these functions (and any other functions $\{h_j(\cdot)\}$ that are bounded uniformly over $j \geq 1$ and $x \in \mathcal{X}$), Assumption A(iii) implies Assumption A(ii) (assuming γ_j is bounded away from zero).

The series functions $\{h_j(\cdot)\}$ of Assumption A(i) need not be orthogonal or orthonormal, nor do the functions $\{m(\cdot,\tau)\}$ need to have unique series expansions in terms of $\{h_j(\cdot)\}$. All that is required is that one can identify a single sequence of coefficients $\{c_j(\tau)\}$ (out of many perhaps) with each function $m(\cdot,\tau)$.

Since each $\tau \in \mathcal{T}$ is identified with a particular sequence $\{c_j(\tau) : j \geq 1\}$ under Assumption A, the metric ρ on $\mathcal{T}$ can be taken to be

$$(2.3) \qquad \rho(\tau,\gamma) = \left[\sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2\right]^{1/2}, \ \forall \tau, \ \gamma \in \mathcal{T}.$$

THEOREM 1. *For* ρ *as defined in (2.3), Assumption A implies that* $\{\nu_n(\cdot)\}$ *is stochastically equicontinuous and* $\mathcal{T}$ *is totally bounded.*

PROOF OF THEOREM 1. First we show stochastic equicontinuity. By Assumptions A(i) and A(ii), we have

$$m(X_{ni}, \tau) - Em(X_{ni}, \tau) = \sum_{j=1}^{\infty} c_j(\tau)(h_j(X_{ni}) - Eh_j(X_{ni})) \tag{2.4}$$

for all $i \leq n$, $n \geq 1$, $\tau \in \mathcal{T}$. Thus, we get

$$\begin{aligned}
&\overline{\lim_{n\to\infty}}\, P^*\Big[\sup_{\rho(\tau,\gamma)<\delta} |\nu_n(\tau) - \nu_n(\gamma)| > \eta\Big] \\
&\leq \overline{\lim_{n\to\infty}}\, \eta^{-2} E^*\Big[\sup_{\rho(\tau,\gamma)<\delta} \Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{\infty}(c_j(\tau) - c_j(\gamma))(h_j(X_{ni}) - Eh_j(X_{ni}))\Big|\Big]^2 \\
&= \overline{\lim_{n\to\infty}}\, \eta^{-2} E^*\Big[\sup_{\rho(\tau,\gamma)<\delta} \Big|\sum_{j=1}^{\infty}\frac{(c_j(\tau) - c_j(\gamma))}{\sqrt{a_j}}\cdot\sqrt{a_j}\,\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(h_j(X_{ni}) - Eh_j(X_{ni}))\Big|\Big]^2 \\
&\leq \eta^{-2} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty}\frac{|c_j(\tau) - c_j(\gamma)|^2}{a_j}\, \overline{\lim_{n\to\infty}}\, E^* \sum_{j=1}^{\infty} a_j \Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(h_j(X_{ni}) - Eh_j(X_{ni}))\Big|^2 \\
&\leq \eta^{-2} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty}|c_j(\tau) - c_j(\gamma)|^2/a_j \cdot \sum_{j=1}^{\infty} a_j\gamma_j,
\end{aligned} \tag{2.5}$$

where the first inequality uses Chebyshev's inequality and (2.4), the second inequality uses the Cauchy–Schwartz inequality, and the third uses standard manipulations of the variance of a sum of correlated rv's.

Stochastic equicontinuity follows from (2.5) and

$$\lim_{\delta\to 0} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2/a_j = 0\,. \tag{2.6}$$

To obtain (2.6), suppose $\epsilon > 0$ is given. By Assumption A(iii), one can choose J sufficiently large that $\sup_{\tau\in\mathcal{T}} \sum_{j=J+1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2/a_j < \epsilon/2$. Take, $\delta = \Big[\epsilon\Big/\Big[2\sum_{j=1}^{J} 1/a_j\Big]\Big]^{1/2}$. By definition of $\rho(\cdot,\cdot)$, this gives

$$(2.7)\qquad \sup_{\tau\in\mathcal{T}} \sum_{j=1}^{J} |c_j(\tau) - c_j(\gamma)|^2/a_j \le \sum_{j=1}^{J} \delta^2/a_j = \epsilon/2\ ,$$

which establishes (2.6).

Next, we show that Assumption A implies that $\mathcal{T}$ is totally bounded. By Assumption A(iii), given any $\epsilon > 0$, there exists an integer $J < \infty$ such that $\sup_{\tau\in\mathcal{T}} \sum_{j=J+1}^{\infty} |c_j(\tau)|^2 < \epsilon/2$. Thus, identifying τ with the function $m(\cdot,\tau)$, each function $m(\cdot,\tau)$ for $\tau \in \mathcal{T}$ is within $\epsilon/2$ under the metric ρ of a function $g(\cdot) \equiv \sum_{j=1}^{\infty} c_j h_j(\cdot)$ for which $c_j = 0$ $\forall j > J$. So, to obtain a finite cover of $\mathcal{T}$ comprised of ϵ–balls, it suffices to take as centers of the ϵ–balls points τ that correspond to functions for which $c_j = 0$ $\forall j > J$. Given that $\sup_{\tau\in\mathcal{T}} |c_j(\tau)|^2 < K$ for some $K < \infty$ $\forall j = 1, \ldots, J$, it is straightforward to choose a finite number of such centers so that the union of the ϵ–balls with these centers covers $\mathcal{T}$. □

2.2. *Fidi Convergence*

We now provide a set of primitive conditions under which the finite dimensional distributions of $\nu_n(\cdot)$ are asymptotically normal. To this end, we state a CLT of Wooldridge (1986) for triangular arrays of rv's that are *near–epoch dependent* (NED) (also referred to in the literature as "functions of mixing processes"). The NED condition is one of asymptotically weak temporal dependence. It was introduced by Ibragimov (1962) and results utilizing it were developed by Billingsley (1968, p. 182) and McLeish (1975a, b). The NED condition is quite general. It allows for non–identical distributions and covers (i) square integrable strong mixing rv's, (ii) square integrable general linear processes (with non–identically distributed strong mixing innovations if desired) including autoregressive and autoregressive–moving average processes (which are not necessarily strong mixing, e.g., see Andrews (1985)), and (iii) various nonlinear autoregressions and dynamic simultaneous equations (see Bierens (1981, Ch. 5), Gallant (1987, pp. 481, 502, 539), and

Gallant and White (1988, p. 29)).

First we define strong mixing double arrays and NED triangular arrays of rv's. Let $\{V_{ni} : i = 0, \pm 1, \ldots; n \geq 1\}$ be a double array of rv's on $(\Omega, \mathcal{B}, P)$. (V_{ni} may be $\mathcal{V}_{ni}$–valued for any measurable space $\mathcal{V}_{ni}$, but usually $\mathcal{V}_{ni} = \mathcal{V}$ and $\mathcal{V} \subset R^v$ or $\mathcal{V} \subset \mathcal{C}^v$ for some $v \geq 1$, where $\mathcal{C}$ denotes the complex plane.) Let $\mathcal{F}^j_{n,i}$ $(\subset \mathcal{B})$ denote the σ–field generated by $(V_{ni}, \ldots, V_{nj})$ for $-\infty \leq i \leq j \leq \infty$. Let $E^j_{n,i}(\cdot)$ denote the conditional expectation operator $E(\cdot \,|\, \mathcal{F}^j_{n,i})$ given the σ–field $\mathcal{F}^j_{n,i}$.

DEFINITION. *The double array* $\{V_{ni}\}$ *of rv's is strong mixing if* $\alpha(s) \downarrow 0$ *as* $s \to \infty$, *where*

$$\alpha(s) = \sup_{i=0,\pm 1,\ldots;n \geq 1} \ \sup_{A \in \mathcal{F}^i_{n,-\infty},\, B \in \mathcal{F}^\infty_{n,i+s}} |P(A \cap B) - P(A)P(B)| \quad \text{for } s \geq 1 . \tag{2.8}$$

$\{V_{ni}\}$ *is strong mixing of size* $-\beta$ *if* $\alpha(s) = O(s^{-\beta-\epsilon})$ *for some* $\epsilon > 0$.

Let $\{Z_{ni} : i \leq n, n \geq 1\}$ be a triangular array of R^k– or $\mathcal{C}^k$–valued rv's on $(\Omega, \mathcal{B}, P)$. Let $\|\cdot\|$ denote the Euclidean norm.

DEFINITION. *The triangular array* $\{Z_{ni}\}$ *is near–epoch dependent (NED) on* $\{V_{ni}\}$ *if* $E\|Z_{ni}\|^2 < \infty$ $\forall i \leq n$, $\forall n \geq 1$, *and* $\eta(s) \downarrow 0$ *as* $s \to \infty$, *where*

$$\eta(s) = \sup_{s < i \leq n-s, n \geq 1} \left[E\|Z_{ni} - E^{i+s}_{n,\,i-s} Z_{ni}\|^2 \right]^{1/2} \quad \text{for } s = 0, 1, \ldots . \tag{2.9}$$

$\{Z_{ni}\}$ *is NED of size* $-\beta$ *on* $\{V_{ni}\}$ *if* $E\|Z_{ni}\|^2 < \infty$ $\forall i \leq n$, $n \geq 1$ *and* $\eta(s) = O(s^{-\beta-\epsilon})$ *for some* $\epsilon > 0$.

(It is easy to see that the R^k– or $\mathcal{C}^k$–valued array $\{Z_{ni}\}$ is NED of size $-\beta$ on $\{V_{ni}\}$ if and only if the real– or complex–valued array $\{Z_{ni\ell}\}$ is NED of size $-\beta$ for $\ell = 1, \ldots, k$, where $Z_{ni} = (Z_{ni1}, \ldots, Z_{nik})'$.)

PROPOSITION 1 (Wooldridge (1986, Ch. 2, Thm. 3.13)). *Let* $\{Z_{ni} : i \leq n, n \geq 1\}$ *be a triangular array of real-valued rv's that satisfies*

(i) $\lim_{n\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_{ni}\right] = \sigma^2$,

(ii) $\sup_{i\leq n, n\geq 1} E|Z_{ni}|^r < \infty$ *for some* $r > 2$, *and*

(iii) $\{Z_{ni}\}$ *is NED of size* -1 *on* $\{V_{ni}\}$, *where* $\{V_{ni}\}$ *is a strong mixing double array of rv's of size* $-2r/(r-2)$.

Then, $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Z_{ni} - EZ_{ni}) \xrightarrow{d} N(0, \sigma^2)$ *as* $n \to \infty$.

Comment: Wooldridge's proof uses the approach of Withers (1981).

DEFINITION. *Let* $\mathrm{Lip}(\alpha,C,\mathcal{X})$ *denote the class of real or complex functions* g *on* $\mathcal{X} \subset R^k$ *that satisfy the Lipschitz condition*

$$(2.10) \qquad |g(x) - g(y)| \leq C\|x-y\|^{\alpha} \;\; \forall x , \; y \in \mathcal{X} ,$$

where $0 < C < \infty$ *and* $0 < \alpha \leq 1$.

The following assumption implies fidi convergence for $\nu_n(\cdot)$:

ASSUMPTION B (Fidi convergence).

(i) $S(\tau,\gamma) = \lim_{n\to\infty} \mathrm{Cov}(\nu_n(\tau), \nu_n(\gamma))$ *exists* $\forall \tau , \; \gamma \in \mathcal{T}$,

(ii) $\sup_{i\leq n, n\geq 1} E\|X_{ni}\|^r < \infty$ *for some* $r > 2$,

(iii) $\{X_{ni} : i \leq n, n \geq 1\}$ *is a NED triangular array of size* -1 *on* $\{V_{ni}\}$, *where* $\{V_{ni} : i = 0, \pm1, \ldots; n \geq 1\}$ *is some strong mixing double array of size* $-2r/(r-2)$, *and*

(iv) $m(\cdot,\tau) \in \mathrm{Lip}(1,C,\mathcal{X})$ $\forall \tau \in \mathcal{T}$ *for some* $C < \infty$.

THEOREM 2. *Under Assumption B, for each finite subset* $(\tau_1, \ldots, \tau_v)$ *of* $\mathcal{T}$, $(\nu_n(\tau_1), \ldots, \nu_n(\tau_v))'$ *converges weakly to a* $N(\underset{\sim}{0}, S_v)$ *rv, where* S_v *is a* $v \times v$ *covariance matrix with* (s,t)–th *element* $S(\tau_s, \tau_t)$.

The proof of Theorem 2 uses the following result:

LEMMA 1. *If* $g \in \mathrm{Lip}(1,C,\mathcal{X})$ *and* $\{Z_{ni} : i \leq n, n \geq 1\}$ *is a NED triangular array of* $\mathcal{X}$*-valued rv's of size* $-\xi$ *on a strong mixing double array* $\{V_{ni}\}$ *of size* $-\beta$ *for* ξ, $\beta > 0$, *then* $\{g(Z_{ni}) : i \leq n, n \geq 1\}$ *is a NED triangular array of real- or complex-valued rv's of size* $-\xi$ *on the same strong mixing double array* $\{V_{ni}\}$. *In addition, the NED numbers* $\{\eta_g(s)\}$ *of* $\{g(Z_{ni})\}$ *satisfy* $\eta_g(s) \leq C\eta(s)$ *for all* $s \geq 0$, *where* $\{\eta(s)\}$ *are the NED numbers of* $\{Z_{ni}\}$.

This lemma is similar to Theorem 4.2 of Gallant and White (1988, p. 48). It differs from the latter in that Gallant and White consider more general functions than those in $\mathrm{Lip}(1,C,\mathcal{X})$, but their result does not obtain the same size of the NED numbers of $\{g(Z_{ni})\}$ as those of $\{Z_{ni}\}$ and their result imposes stronger moment conditions.

PROOF OF LEMMA 1. For arbitrary fixed $x \in \mathcal{X}$,

$$\begin{aligned} E|g(Z_{ni})|^2 &\leq E(|g(Z_{ni}) - g(x)| + |g(x)|)^2 \\ &\leq E(C\|Z_{ni} - x\| + |g(x)|)^2 < \infty \end{aligned} \tag{2.11}$$

for all $i \leq n$, $n \geq 1$, since $|g(x)| < \infty$ and $E\|Z_{ni}\|^2 \leq \infty$.

Next, note that the conditional expectation $E_{n,i-s}^{i+s} g(Z_{ni})$ minimizes $E_{n,i-s}^{i+s}|g(Z_{ni}) - Y|^2$ over all $\mathcal{F}_{n,i-s}^{i+s}$-measurable rv's Y. With this result and the Lipschitz condition on g, we get

$$\begin{aligned} \eta_g(s) &= \sup_{s<i\leq n-s, n\geq 1} \left[E E_{n,i-s}^{i+s} |g(Z_{ni}) - E_{n,i-s}^{i+s} g(Z_{ni})|^2 \right]^{1/2} \\ &\leq \sup_{s<i\leq n-s, n\geq 1} \left[E E_{n,i-s}^{i+s} |g(Z_{ni}) - g(E_{n,i-s}^{i+s} Z_{ni})|^2 \right]^{1/2} \\ &\leq C \sup_{s<i\leq n-s, n\geq 1} \left[E\|Z_{ni} - E_{n,i-s}^{i+s} Z_{ni}\|^2 \right]^{1/2} \\ &= C\eta(s) \end{aligned} \tag{2.12}$$

and the conclusion follows. □

PROOF OF THEOREM 2. Let $\underset{\sim}{\tau} = (\tau_1, \ldots, \tau_v)' \in \mathcal{T}^v$, $m(\cdot, \underset{\sim}{\tau}) = (m(\cdot, \tau_1), \ldots, m(\cdot, \tau_v))'$, and $\lambda \in R^v$ with $\|\lambda\| = 1$. It suffices to show that conditions (i), (ii), and (iii) of Proposition 1 hold with $Z_{ni} = \lambda' m(X_{ni}, \underset{\sim}{\tau})$ and $\sigma^2 = \lambda' S_v \lambda$, $\forall \underset{\sim}{\tau} \in \mathcal{T}^v$, $\forall \lambda \in R^v$ with $\|\lambda\| = 1$, and $\forall v \geq 1$.

Assumption B(i) implies condition (i) of Proposition 1. By the same argument as in the proof of Lemma 1, Assumptions B(ii) and (iv) imply that $\sup_{i \leq n, n \geq 1} E|m(X_{ni}, \tau)|^r < \infty$. That is, condition (ii) of Proposition 1 holds. Assumptions B(iii) and (iv) and Lemma 1 with $g(\cdot) = \lambda' m(\cdot, \underset{\sim}{\tau})$ imply that $\{\lambda' m(X_{ni}, \underset{\sim}{\tau})\}$ is NED of size −1 on $\{V_{ni}\}$. That is, condition (iii) of Proposition 1 holds. □

2.3. Stochastic Equicontinuity with NED Random Variables

Assumption A given above imposes a condition of asymptotic weak dependence in A(iv). (It requires $\sum_{s=-\infty}^{\infty} \gamma_j(s) < \infty$, $\forall j \geq 1$.) Since Assumption A is to be used in conjunction with Assumption B, it is useful to determine a replacement for Assumption A(iv) that uses the NED condition that is utilized in B. The following assumption does this.

ASSUMPTION A1. (i) (Series expansion) *Assumption A(i) holds.*

(ii) (Smoothness of series functions) $h_j(\cdot) \in \mathrm{Lip}(1, B_j, \mathcal{X})$ *for some* $B_j < \infty$ $\forall j \geq 1$ *and* $\sup_{j \geq 1} \sup_{i \leq n, n \geq 1} E|h_j(X_{ni})|^r < \infty$ *for some* $r > 2$.

(iii) (Smoothness of $m(\cdot, \tau)$) $\sup_{\tau \in \mathcal{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2 / a_j \to 0$ *as* $J \to \infty$ *for some summable sequence of positive constants* $\{a_j\}$ *for which* $\sum_{j=1}^{\infty} a_j B_j < \infty$.

(iv) (Weak dependence) $\{X_{ni} : i \leq n, n \geq 1\}$ *is a NED triangular array of size −1 on* $\{V_{ni}\}$, *where* $\{V_{ni}\}$ *is some strong mixing double array of size* $-2r/(r-2)$.

THEOREM 3. *Assumption A1 implies Assumption A.*

To prove Theorem 3 we use the following inequality for the autocovariances of NED rv's. Let $\|\cdot\|_p$ denote the $L^p(P)$ semi–norm for $p \geq 1$.

LEMMA 2. *Let* $\{Z_{ni}\}$ *be a NED triangular array on* $\{V_{ni}\}$ *(with NED numbers* $\{\eta(s) : s \geq 0\}$ *), where* $\{V_{ni}\}$ *is a strong mixing double array (with mixing numbers* $\{\alpha(s) : s \geq 1\}$ *). Then, for all* $r > 2$ *,* $|\mathrm{Cov}(Z_{ni}, Z_{ni-s})| \leq \|Z_{ni-s}\|_2 \eta(a) + 6\|Z_{ni}\|_2 \|Z_{ni-s}\|_r \alpha(b)^{(r-2)/(2r)}$ *for all* $0 \leq s < i$ *and* $n \geq 1$ *, where* a *and* b *are positive integers for which* $a+b \leq s$.

PROOF OF LEMMA 2. We have

$$\begin{aligned}
|\mathrm{Cov}(Z_{ni}, Z_{ni-s})| &= |EZ_{ni}(Z_{ni-s} - EZ_{ni-s})| \\
&= |E(Z_{ni} - E^{i+a}_{n,\, i-a} Z_{ni})(Z_{ni-s} - EZ_{ni-s}) \\
&\quad + E(E^{i+a}_{n,\, i-a} Z_{ni}) E^{i+a}_{n,\, i-a}(Z_{ni-s} - EZ_{ni-s})| \\
&\leq \left[E|Z_{ni} - E^{i+a}_{n,\, i-a} Z_{ni}|^2\right]^{1/2} \mathrm{Var}^{1/2}(Z_{ni-s}) \qquad (2.13) \\
&\quad + \left[E|E^{i+a}_{n,\, i-a} Z_{ni}|^2\right]^{1/2} \left[E|E^{i+a}_{n,\, i-a} Z_{ni-s} - EZ_{ni-s}|^2\right]^{1/2} \\
&\leq \eta(a)\|Z_{ni-s}\|_2 + \|Z_{ni}\|_2 6\alpha(b)^{(r-2)/(2r)} \|Z_{ni-s}\|_r
\end{aligned}$$

using the Cauchy–Schwartz inequality, the conditional Jensen's inequality, and a strong mixing inequality of McLeish (1975a, Lemma 2.1). □

PROOF OF THEOREM 3. Assumptions A1(i) and A(i) are equivalent. By Assumption A1(ii), Assumption A(ii) holds if $\sum_{j=1}^{\infty} |c_j(\tau)| < \infty \ \forall \tau \in \mathcal{T}$. This holds under Assumption A1(iii) because $\sum_{j=1}^{\infty} |c_j(\tau)|^2 / a_j < \infty$ for some summable sequence $\{a_j \geq 0\}$ only if

$\sum_{j=1}^{\infty} |c_j(\tau)| < \infty$. To prove the latter, suppose $\sum_{j=1}^{\infty} |c_j(\tau)| = \infty$. Then,

$$\infty = \sum_{j=1}^{\infty} \frac{|c_j(\tau)|}{\sqrt{a_j}} \sqrt{a_j} \leq \left[\sum_{j=1}^{\infty} |c_j(\tau)|^2/a_j\right]^{1/2} \left[\sum_{j=1}^{\infty} a_j\right]^{1/2} \tag{2.14}$$

and either $\sum_{j=1}^{\infty} |c_j(\tau)|^2/a_j = \infty$ or $\sum_{j=1}^{\infty} a_j = \infty$.

Next we show that Assumption A1 implies A(iii). By Assumptions A1(ii) and (iv) and Lemma 1, $\{h_j(X_{ni})\}$ is a NED triangular array on $\{V_{ni}\}$ with NED numbers $\{\eta_{h_j}(s)\}$ that satisfy $\eta_{h_j}(s) \leq B_j \eta(s)$ $\forall s \geq 0$, where $\{\eta(s)\}$ are the NED numbers of $\{X_{ni}\}$. In consequence, Assumption A1(ii) and Lemma 2 with $Z_{ni} = h_j(X_{ni})$ and $a = b = [s/2]$ give

$$\begin{aligned} \gamma_j(s) &\leq \sup_{\substack{i\leq n, n\geq 1 \\ v\geq 1}} \|h_v(X_{ni})\|_2 \left[B_j \eta([s/2]) + 6 \sup_{\substack{i\leq n, n\geq 1 \\ v\geq 1}} \|h_v(X_{ni})\|_r \alpha([s/2])^{(r-2)/(2r)}\right] , \\ \gamma_j &= \sum_{s=-\infty}^{\infty} \gamma_j(s) \leq B_j D_1 \sum_{s=-\infty}^{\infty} \eta([s/2]) + D_2 \sum_{s=-\infty}^{\infty} \alpha([s/2])^{(r-2)/(2r)} < \infty \end{aligned} \tag{2.15}$$

$\forall j \geq 1$ for some finite constants D_1 and D_2 , using the fact that $\{\eta(s)\}$ and $\{\alpha(s)\}$ are of size -1 and $-2r/(r-2)$ respectively. Thus, $\sum_{j=1}^{\infty} a_j \gamma_j < \infty$ is implied by $\sum_{j=1}^{\infty} a_j B_j < \infty$ and $\sum_{j=1}^{\infty} a_j < \infty$. This result and Assumption A1(iii) implies Assumption A(iii). □

2.4. An Empirical Process CLT

The results of Theorems 1–3 above can be combined to yield a CLT for $\{\nu_n(\cdot)\}$. Before stating the result, however, we define weak convergence. The definition we use is due to Dudley (1985) and Hoffman-Jorgensen (1985) (with a slight, inessential but convenient, modification introduced by Pollard (1988)). Let $L^\infty(\mathcal{T})$ denote the space of bounded real functions on $\mathcal{T}$. Endow $L^\infty(\mathcal{T})$ with the uniform metric d .

DEFINITION. *If* $\{\nu_n(\cdot)\}$ *are (not necessarily Borel measurable) maps from* Ω *into the metric space* $(L^\infty(\mathcal{T}), d)$ *and if* $\nu(\cdot)$ *is an* $L^\infty(\mathcal{T})$*-valued Borel measurable rv (not necessarily defined on* $(\Omega, \mathcal{B}, P)$*), then*

(2.16) $$\nu_n(\cdot) \overset{W}{\to} \nu(\cdot) \text{ means } E^*f(\nu_n(\cdot)) \to Ef(\nu(\cdot)) \text{ as } n \to \infty$$

for all bounded uniformly continuous real functions f *on* $L^\infty(\mathcal{T})$, *where* E^* *denotes outer expectation.*

Define

(2.17) $$\mathcal{U}_\rho(\mathcal{T}) = \{y \in L^\infty(\mathcal{T}) : y \text{ is uniformly continuous with respect to } \rho \text{ on } \mathcal{T}\}\ .$$

COROLLARY 1. *For* $\nu_n(\cdot)$ *and* ρ *as defined in (2.1) and (2.3), Assumption A or A1 coupled with Assumption B implies* $\nu_n(\cdot) \overset{W}{\to} \nu(\cdot)$, *where* $\nu(\cdot)$ *is a mean zero Gaussian process with covariance function* $S(\cdot,\cdot)$ *whose sample paths lie in* $\mathcal{U}_\rho(\mathcal{T})$ *with probability one.*

PROOF OF COROLLARY 1. Theorems 1–3 and a result of Pollard (1988, Weak Convergence Theorem) give the desired result. □

3. DIFFERENTIABLE FUNCTIONS OF R^k–VALUED RANDOM VARIABLES

In this section we apply the results of Section 2 to a class of differentiable functions $\{m(\cdot,\tau) : \tau \in \mathcal{T}\}$ that are defined on some open bounded subset $\mathcal{X}$ of R^k whose boundary is *minimally smooth.*[3] (See Stein (1970, pp. 181, 189) or Edmunds and Moscatelli (1977, p. 8) for the definition of minimally smooth.) Examples of sets in R^k with minimally smooth boundaries include open bounded sets that are convex or whose boundaries are C^1–embedded in R^k. Finite unions of disjoint sets of the aforementioned type also have minimally smooth boundaries.

The functions $\{m(\cdot,\tau)\}$ are taken to be uniformly smooth in the sense of having a uniformly bounded Sobolev norm of some order. By definition, the Sobolev norm of order (q,p) of a real function f defined on a subset $\mathcal{Y}$ of R^k is

$$(3.1) \qquad \|f\|_{q,p,\mathcal{Y}} = \left[\sum_{|\alpha| \leq q} \int_{\mathcal{Y}} |D^\alpha f(x)|^p dx \right]^{1/p},$$

where q is a non-negative integer, $1 \leq p < \infty$, $\alpha = (\alpha_1, \ldots, \alpha_k)' \in R^k$ has non-negative integer-valued elements, $|\alpha| = \sum_{\ell=1}^{k} \alpha_\ell$, and $D^\alpha f(x) = \partial^{|\alpha|} f(x)/(\partial x_1^{\alpha_1} \times \cdots \times \partial x_k^{\alpha_k})$.[4] Below we assume that $\sup_{\tau \in \mathcal{T}} \|m(\cdot,\tau)\|_{q,2,\mathcal{X}} < \infty$ for some $q > (k+1)/2$. That is, the functions $\{m(\cdot,\tau)\}$ must have more derivatives (q) finite (and bounded in L^2) than the dimension (k) of their domain plus one and divided by two.

The index set $\mathcal{T}$ of the functions $\{m(\cdot,\tau)\}$ can be any set that can be used to index functions on $\mathcal{X}$ whose Sobolev norms of order $(q,2)$ are bounded. In particular, one could equate τ with $m(\cdot,\tau)$, in which case $\mathcal{T}$ would be the same space of functions on $\mathcal{X}$ as $\{m(\cdot,\tau)\}$. Below we make $\mathcal{T}$ a metric space by defining a metric ρ on it.

Exponential Fourier functions are used to obtain series expansions of the functions $\{m(\cdot,\tau)\}$. Let $(a,b)^k$ be an open bounded k-dimensional cube that contains the closure of $\mathcal{X}$. Following the approach of Edmunds and Moscatelli (1977, p. 10), which uses Theorem 5 of Stein (1970, p. 181), we extend the function $m(\cdot,\tau)$ on $\mathcal{X}$ to a function $m^*(\cdot,\tau)$ on $(a,b)^k$ such that $m^*(\cdot,\tau)$ is periodic in each of its elements and

$$(3.2) \qquad \|m(\cdot,\tau)\|_{\tilde{q},2,\mathcal{X}} \leq \|m^*(\cdot,\tau)\|_{\tilde{q},2,(a,b)^k} \leq G\|m(\cdot,\tau)\|_{\tilde{q},2,\mathcal{X}}$$

for all non-negative integers $\tilde{q} \leq q$, for some $G < \infty$ that does not depend on τ. The Fourier expansion of the periodic function $m^*(\cdot,\tau)$ gives the desired expansion of $m(\cdot,\tau)$ by restricting the domain of the expansion to $\mathcal{X}$. (Note that the extension operator used above is linear, so that equation (3.2) also holds with $m(\cdot,\tau)$ and $m^*(\cdot,\tau)$ replaced by

$m(\cdot,\tau) - m(\cdot,\gamma)$ and $m^*(\cdot,\tau) - m^*(\cdot,\gamma)$, respectively, as is used below.)

Call a k–vector of integers a *multi–index.* By Theorem 2 of Edmunds and Moscatelli (1977, pp. 11 and 25), a sequence of multi–indexes $\{\kappa(j) : j \geq 1\}$ can be constructed such that

$$\sup_{x \in (a,b)^k} |m^*(x,\tau) - \sum_{j=1}^{J} c_j(\tau)h_j(x)| \leq G^* J^{-q/k+1/2+\epsilon} \|m^*(\cdot,\tau)\|_{q,2,(a,b)^k} \tag{3.3}$$

$\forall \epsilon > 0$, for some $G^* < \infty$, where

$$\begin{aligned} h_j(x) &= (b-a)^{-k/2} e^{2\pi i \kappa(j)'(x - a\underset{\sim}{1})/(b-a)} \text{ and} \\ c_j(\tau) &= \int_{(a,b)^k} m^*(x,\tau)\overline{h_j(x)}dx \, . \end{aligned} \tag{3.4}$$

Here, $\underset{\sim}{1}$ denotes a k–vector of ones and $\overline{h_j(x)}$ denotes the complex conjugate of $h_j(x)$. By (3.3) and the assumption $q > (k+1)/2$, we get the following pointwise convergent series expansion of $m(\cdot,\tau)$:

$$m(x,\tau) = \sum_{j=1}^{\infty} c_j(\tau)h_j(x) \;\; \forall x \in \mathcal{X} , \;\; \forall \tau \in \mathcal{T} . \tag{3.5}$$

Next we define the metric ρ on $\mathcal{T}$. For a function f defined on a subset $\mathcal{Y}$ of R^k , define

$$\|f\|_{p,\mathcal{Y}} = \left[\int_{\mathcal{Y}} |f(x)|^p dx\right]^{1/p} \text{ for } 1 \leq p < \infty . \tag{3.6}$$

Since the functions $\{h_j(\cdot) : j \geq 1\}$ are orthonormal on $(a,b)^k$ (with respect to Lebesgue measure) and (3.2) holds with $\tilde{q} = 0$ and with $m(\cdot,\tau)$ and $m^*(\cdot,\tau)$ replaced by $m(\cdot,\tau) - m(\cdot,\gamma)$ and $m^*(\cdot,\tau) - m^*(\cdot,\gamma)$, respectively, we get

$$(3.7)\qquad \left[\sum_{j=1}^{\infty}|c_j(\tau)-c_j(\gamma)|^2\right]^{1/2} = \|m^*(\cdot,\tau)-m^*(\cdot,\gamma)\|_{2,(a,b)^k}$$

$$\in [\|m(\cdot,\tau)-m(\cdot,\gamma)\|_{2,\mathcal{X}},\ G\|m(\cdot,\tau)-m(\cdot,\gamma)\|_{2,\mathcal{X}}] \,.$$

Hence, we can define the metric ρ on $\mathcal{T}$ to be

$$(3.8)\qquad \rho(\tau,\gamma) = \|m(\cdot,\tau)-m(\cdot,\gamma)\|_{2,\mathcal{X}}$$

and with this definition ρ is equivalent to (although not identical to) the metric ρ defined in (2.3).

Using Corollary 1 of Section 2, the following assumption is shown to imply that $\{\nu_n(\cdot)\}$ satisfies a CLT when indexed by the family of differentiable functions $\{m(\cdot,\tau)\}$ on $\mathcal{X}$:

ASSUMPTION D: (i) $\mathcal{X}$ *is a bounded open subset of* R^k *with minimally smooth boundary.*

(ii) $\sup_{\tau\in\mathcal{T}} \|m(\cdot,\tau)\|_{q,2,\mathcal{X}} < \infty$ *for some* $q > (k+1)/2$.

(iii) $\{X_{ni}\}$ *is a NED triangular array of size* -1 *on* $\{V_{ni}\}$, *where* $\{V_{ni}\}$ *is some strong mixing double array of size* -2.

(iv) $S(\tau,\gamma) = \lim_{n\to\infty} \mathrm{Cov}(\nu_n(\tau), \nu_n(\gamma))$ *exists* $\forall\tau,\ \gamma\in\mathcal{T}$.

THEOREM 4. *For* $\nu_n(\cdot)$ *and* ρ *as defined in (2.1) and (3.8), respectively, Assumption D implies that* $\nu_n(\cdot) \overset{w}{\to} \nu(\cdot)$, *where* $\nu(\cdot)$ *is a mean zero Gaussian process with covariance function* $S(\cdot,\cdot)$ *whose sample paths lie in* $\mathcal{U}_\rho(\mathcal{T})$ *with probability one.*

Comments: 1. Since strong mixing triangular arrays are NED on themselves, Theorem 4 covers the case where $\{X_{ni}\}$ is strong mixing of size -2. In fact, in this case, the condition $q > (k+1)/2$ can be replaced by $q > k/2$. (This result is obtained by verifying Assumption A, rather than A1, in the proof of Theorem 4.)

2. The smoothness condition $q > (k+1)/2$ of Assumption D(ii) (or $q > k/2$ in the strong mixing case) can be compared with other conditions used in the literature to obtain CLTs for empirical processes indexed by smooth functions. In particular, if Assumptions D(i) and (ii) hold but with q unspecified, if $\{X_{ni}\}$ are iid, and if one verifies Ossiander's (1987, Thm. 3.1, p. 904) metric entropy with bracketing condition using Kolmogorov and Tihomirov's (1961, Thm. XIII, p. 308) calculation of the (sup–norm) metric entropy numbers of $\{m(\cdot,\tau)\}$, then one needs the condition $q > k/2$. This is exactly the same condition as used here in the strong mixing $\{X_{ni}\}$ case, and is only slightly weaker than the condition used in the NED case.[5]

Actually, Kolmogorov and Tihomirov's (1961) metric entropy calculations allow q to be real–valued—they assume the functions $\{m(\cdot,\tau)\}$ have bounded Sobolev norm of order $([q], \infty)$ and the [q]–th partial derivatives of $\{m(\cdot,\tau)\}$ are in $\mathrm{Lip}(\alpha, C, \mathcal{X})$ $\forall \tau \in \mathcal{T}$ where $\alpha = q - [q]$ —whereas Assumption D(ii) requires q to be integer–valued. Thus, Ossiander's (1987) metric entropy condition can be verified under slightly weaker conditions than those of Assumption D even in the case where $\{X_{ni}\}$ is strong mixing.

3. For the case where $\mathcal{X}$ is a bounded interval (a,b) and $\{X_{ni}\}$ is strong mixing of size –2, the smoothness Assumption D(ii) can be relaxed. In particular, it can be replaced by:

D(ii)* $m(\cdot,\tau) \in \mathrm{Lip}(\alpha, C, (a,b))$ $\forall \tau \in \mathcal{T}$ for some $\alpha \in (1/2, 1]$ and $C < \infty$.

(This result is proved by verifying Assumptions A and B using trigonometric series expansions and results of Zygmund (1955, p. 135) and Dehling (1983, Lemma 10.1, p. 428).)

The smoothness condition D(ii)* is exactly the same as the smoothness condition needed to verify Ossiander's (1987, Thm. 3.1, p. 904) metric entropy with bracketing condition using Kolmogorov and Tihomirov's (1961, Thm. XIII, p. 308) computation of the (sup norm) metric entropy numbers for Lipschitz functions on a bounded interval. Ossiander's bracketing condition yields an empirical process CLT for iid sequences of rv's. In addition, the condition $\alpha > 1/2$ is needed by Dehling (1983, Thm. 5, p. 399) to apply his CLT for Banach space–valued rv's to Lipschitz functions on a bounded interval. Dehling's CLT applies to sequences of weakly stationary strong mixing rv's of size –4.

4. If the support of the rv's $\{X_i\}$ is not an open set, Theorem 4 still applies provided the functions $\{m(\cdot,\tau) : \tau \in \mathcal{T}\}$ are defined on an open bounded set $\mathcal{X}$ that contains the support of $\{X_i\}$. On the other hand, if the functions $\{m(\cdot,\tau) : \tau \in \mathcal{T}\}$ are only defined on the support of $\{X_i\}$ and certain elements of X_i only take on a finite number of values, then $\mathcal{X}$ cannot be an open set. The latter problem can be circumvented in this case by writing $\nu_n(\cdot)$ as the sum of several empirical processes based on X_i vectors of lower dimension with the discrete elements of X_i eliminated. For example, if the last element of X_i only takes values 0 and 1, write $\nu_n(\cdot) = \nu_{n0}(\cdot) + \nu_{n1}(\cdot)$, where $\nu_{n0}(\cdot)$ is the empirical process based on $m((\tilde{X}_i, 0), \cdot)1(X_{ki} = 0)$ for $\tilde{X}_i = (X_{1i}, \ldots, X_{k-1,i})$ and likewise for $\nu_{n1}(\cdot)$. Then, if Assumption D holds for the $k-1$ dimensional spaces $\mathcal{X}_0$ and $\mathcal{X}_1$ (which denote the support of $\tilde{X}_i$ conditional on $X_{ki} = 0$ and $X_{ki} = 1$, respectively), $\{\nu_{n0}(\cdot)\}$ and $\{\nu_{n1}(\cdot)\}$ are stochastically equicontinuous. In turn, this implies $\{\nu_n(\cdot)\}$ is stochastically equicontinuous and satisfies a CLT, as desired.

5. Boundedness of X_{ni} can be a restrictive assumption in some applications of Theorem 4. It can be relaxed somewhat, however, in a way that is useful in some applications. Suppose $m(X_{ni}, \tau)$ is of the form $m^1(X_{1ni}, \tau)X_{2ni}$, where $X_{ni} = (X_{1ni}', X_{2ni}')'$ and $X_{1ni} \in R^{k_1}$. Let $\mathcal{X}_1$ be the space of possible values of X_{1ni}. Replace $\mathcal{X}$, $m(\cdot,\tau)$, and $k+1$ in Assumption D and (3.8) by $\mathcal{X}_1$, $m^1(\cdot,\tau)$, and k_1. Assume that the revised Assumption D holds, $\sup_{i\leq n, n\geq 1} E\|X_{2ni}\|^r < \infty$ for some $r > 2$, and $\{X_{ni}\}$ is a strong mixing triangular array of size $-2r/(r-2)$. Then, the conclusion of Theorem 4 still holds even though X_{ni} and $m(\cdot,\tau)$ may be unbounded.

To prove this, take the series expansion of $m(\cdot,\tau)$ to be given by X_{2ni} times a Fourier series expansion of $m^1(\cdot,\tau)$. Then, verify Assumption A rather than Assumption

A1, since the series functions $h_j(\cdot)$ do not satisfy Assumption A1(ii) in this case. Next, note that the smoothness assumption B(iv) on $m(\cdot,\tau)$, which is violated in this case, can be eliminated and the result of Theorem 2 still holds, provided $\{X_{ni}\}$ satisfies the additional conditions introduced above.

PROOF OF THEOREM 4: We show that $D \Rightarrow A1$ and $D \Rightarrow B$. Corollary 1 then gives the desired result.

Under Assumptions D(i) and (ii), Assumption A1(i) holds by the argument given in equations (3.3)–(3.5) above. Next we show Assumption A1(ii) holds. Since $(a,b)^k$ is open and convex, the mean value theorem gives: $\forall x$, $y \in (a,b)^k$

$$(3.9)\quad |h_j(x) - h_j(y)| \le \sup_{x^* \in (a,b)^k} \left\| \frac{\partial}{\partial x} h_j(x^*) \right\| \cdot \|x-y\| \le 2\pi(b-a)^{-1-k/2} \sum_{\ell=1}^{k} |\kappa_\ell(j)| \, \|x-y\| .$$

Since the multi–indexes $\{\kappa(j)\}$ considered by Edmunds and Moscatelli (1977, p. 11) and used in (3.3)–(3.5) above satisfy

$$(3.10)\qquad |\kappa_\ell(j)| \le Dj^{1/k} \;\; \forall \ell = 1, \ldots, k \text{ for some } D < \infty ,$$

$h_j(\cdot)$ is in $\mathrm{Lip}(1, B_j, (a,b)^k)$ with $B_j = 2\pi(b-a)^{-1-k/2} kDj^{1/k}$. In addition, $|h_j(x)|$ is bounded by $(b-a)^{-k/2}$ for all $x \in \mathcal{X}$ and $j \ge 1$. Thus, Assumption A1(ii) holds for all $r < \infty$.

To establish Assumption A1(iii), let $a_j = j^{-2q/k+\epsilon}$ for some $\epsilon \in (0, -1 + (2q-1)/k)$. Then, we have

$$(3.11)\qquad \begin{aligned} \sup_{\tau \in \mathcal{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2 / a_j &= \sup_{\tau \in \mathcal{T}} \sum_{j=J}^{\infty} j^{2q/k-\epsilon} |c_j(\tau)|^2 \\ &\le J^{-\epsilon} \sup_{\tau \in \mathcal{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 = o(1) \text{ as } J \to \infty , \end{aligned}$$

since $\sup_{\tau \in \mathcal{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 < \infty$ by (3.12) below. For $q > (k+1)/2$ and ϵ as defined above, $\sum_{j=1}^{\infty} a_j B_j \propto \sum_{j=1}^{\infty} j^{-(2q-1)/k+\epsilon} < \infty$, and hence, Assumption A1(iii) holds.

To show $\sup_{\tau \in \mathcal{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 < \infty$, we use Assumption D(ii) and (3.2). Let (f,g) denote $\int_{(a,b)^k} f(x)\overline{g(x)}dx$. We have

$$
\begin{aligned}
\infty &> G \sup_{\tau \in \mathcal{T}} \|m(\cdot,\tau)\|_{q,2,\mathcal{X}} \\
&\geq \sup_{\tau \in \mathcal{T}} \|m^*(\cdot,\tau)\|_{q,2,(a,b)^k} \\
&= \sup_{\tau \in \mathcal{T}} \left[\sum_{|\alpha| \leq q} \sum_{j=1}^{\infty} |(D^{\alpha} m^*(\cdot,\tau), h_j(\cdot))|^2 \right]^{1/2} \\
&= \sup_{\tau \in \mathcal{T}} \left[\sum_{|\alpha| \leq q} \sum_{j=1}^{\infty} \left[\frac{2\pi}{b-a}\right]^{2\alpha} \prod_{v=1}^{k} \kappa_v(j)^{2\alpha_v} |(m^*(\cdot,\tau), h_j(\cdot))|^2 \right]^{1/2} \qquad (3.12)\\
&\geq \min\left[\left[\frac{2\pi}{b-a}\right]^q, 1\right] \left[\sup_{\tau \in \mathcal{T}} \sum_{j=1}^{\infty} \max_{1 \leq v \leq k} \kappa_v(j)^{2q} |c_j(\tau)|^2 \right]^{1/2} \\
&\geq B^* \min\left[\left[\frac{2\pi}{b-a}\right]^q, 1\right] \left[\sup_{\tau \in \mathcal{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 \right]^{1/2}
\end{aligned}
$$

for some constant $B^* < \infty$. The last inequality holds because $|\kappa(j)| = \sum_{v=1}^{k} |\kappa_v(j)| \approx j^{1/k}$ (see Edmunds and Moscatelli (1977, p. 11)) implies that $\max_{1 \leq v \leq k} |\kappa_v(j)| \approx j^{1/k}$ (where $f(j) \approx g(j)$ means that $0 < \underline{\lim}_{j \to \infty} |f(j)/g(j)| \leq \overline{\lim}_{j \to \infty} |f(j)/g(j)| < \infty$).

Assumption A1(iv) holds for some r sufficiently large by D(iii).

Next we show $D \Rightarrow B$. Assumptions B(i) and D(iv) are equivalent. Assumption B(ii) holds for all $r < \infty$ since $\mathcal{X}$ is bounded. Assumption B(iii) holds for some r sufficiently large by Assumption D(iii). To show Assumption B(iv), it suffices to show $m^*(\cdot,\tau) \in \mathrm{Lip}(1, C, (a,b)^k)$ $\forall \tau \in \mathcal{T}$ for some $C < \infty$. This follows by the same argument as in (3.9) using the uniform bound on the first derivatives of $m^*(\cdot,\tau)$ given by Assumption D(ii) and (3.2). □

4. EXAMPLE

This section briefly illustrates the use of the results given above in establishing the asymptotic normality and efficiency of a semiparametric estimator. We consider a weighted least squares (LS) estimator of a dynamic nonlinear regression model with weights that adapt to the form of heteroskedasticity present. Carroll (1982) and Robinson (1987) have considered this estimator in non–dynamic linear regression models. The model is

$$(4.1) \qquad Y_i = f(Y_{i-1}, \ldots, Y_{i-p}, Z_i, \theta_0) + U_i \text{ for } i = 1, \ldots, n,$$

where $Y_i, U_i \in R$, $Z_i \in R^{\xi}$, and $\theta_0 \in \Theta \subset R^{\gamma}$. The errors $\{U_i\}$ are independent with the conditional mean of U_i given $(Y_{i-1}, \ldots, Y_{i-p}, Z_i)$ equal to 0 and the conditional variance function $\tau_0(\cdot)$ of U_i defined by

$$(4.2) \qquad \tau_0(Z_i) = E(U_i^2 | Z_i) = E(U_i^2 | Y_{i-1}, \ldots, Y_{i-p}, Z_i).$$

The regressors $\{Z_i\}$ may be fixed or random.

The unknown parameter θ_0 is estimated using a preliminary estimator $\hat{\tau}(\cdot)$ of $\tau_0(\cdot)$:

$$(4.3) \qquad \hat{\theta} \text{ minimizes } \sum_{i=1}^{n} (Y_i - f_i(\theta))^2 / \hat{\tau}(Z_i)$$

over $\theta \in \Theta$, where $f_i(\theta)$ denotes $f(Y_{i-1}, \ldots, Y_{i-p}, Z_i, \theta)$. Under suitable conditions, this estimator of θ_0 is consistent and satisfies

$$(4.4) \quad \sqrt{n}(\hat{\theta} - \theta_0) = \left[\frac{1}{n}\sum_{i=1}^{n} E\frac{\partial}{\partial\theta} f_i(\theta_0)\frac{\partial}{\partial\theta'} f_i(\theta_0)/\tau_0(Z_i)\right]^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\hat{\tau}(Z_i) + o_p(1),$$

where the inverted matrix is $O(1)$ as $n \to \infty$. The linear approximation (4.4) and the CLT give the asymptotic normality of $\sqrt{n}(\hat{\theta} - \theta_0)$ if it can be shown that

$$(4.5) \qquad \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\hat{\tau}(Z_i) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau_0(Z_i) \xrightarrow{p} \underset{\sim}{0} .$$

If the errors are normally distributed, then (4.4) and (4.5) combine to establish the asymptotic efficiency of $\hat{\theta}$, since $\hat{\theta}$ has the same asymptotic distribution as the weighted LS "estimator" that uses $\tau_0(\cdot)$ to form the weights.

Define an empirical process $\nu_n(\cdot)$ as follows:

$$(4.6) \qquad \nu_n(\tau) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau(Z_i) \quad \text{for } \tau \in \mathcal{T} ,$$

where $\mathcal{T}$ is a class of smooth functions $\tau(\cdot)$ for which $E\left\|U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau(Z_i)\right\|^2 < \infty$. Note that $EU_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau(Z_i) = 0 \ \forall \tau \in \mathcal{T}$, since U_i has conditional mean zero.

The results of this paper (in particular, Theorem 4 or Comment 5 following Theorem 4) can be used to show that $\{\nu_n(\cdot) : n \geq 1\}$ is stochastically equicontinuous. Equation (4.5) then follows provided $\hat{\tau}$ converges in probability to τ_0 with respect to the appropriate metric. For example, if Theorem 4 or Comment 5 following Theorem 4 is used, then $\hat{\tau} \xrightarrow{p} \tau_0$ if

$$(4.7) \qquad \int_{\mathcal{Z}} (\hat{\tau}(z) - \tau_0(z))^2 dz \xrightarrow{p} 0 \quad \text{as } n \to \infty$$

and $\hat{\tau}(\cdot)$ and $\tau_0(\cdot)$ are bounded away from zero, where $\mathcal{Z}$ denotes the (bounded) set of possible realizations of $\{Z_i : i \geq 1\}$.

In sum, the stochastic equicontinuity of an empirical process can be used very effectively in establishing the asymptotic normality and efficiency of semiparametric estimators. See Andrews (1989a,b,c) for the application of empirical process results, such as those of the present paper, to a broad class of semiparametric estimators and to tests based on them.

FOOTNOTES

[1]I thank David Pollard and Yoon Whang for helpful comments. I gratefully acknowledge research support from the Alfred P. Sloan Foundation and the National Science Foundation through a Research Fellowship and grant no. 8618617 respectively.

[2]This is true of the results of this paper only if $\{X_{ni}\}$ is assumed to be strong mixing. If $\{X_{ni}\}$ is only assumed to be NED, then a slightly stronger smoothness condition than Ossiander's is needed, see Section 3.

[3]The reason for requiring $\mathcal{X}$ to have a minimally smooth boundary is to permit the smooth extension of the functions $\{m(\cdot,\tau)\}$ to periodic functions defined on a cube that contains $\mathcal{X}$. Fourier expansions of the latter functions yield the desired series expansions of the functions $\{m(\cdot,\tau)\}$.

[4]The partial derivative $D^{\alpha}f$ of f can be defined in the usual sense or in a weaker sense (see Stein (1970, p. 180)). The latter allows the partial derivatives of order $q-1$ to be Lipschitz (with exponent one) rather than partially differentiable in the usual sense while maintaining $\|f\|_{q,p,\mathcal{Y}} < \infty$.

[5]Theorem 4 also differs slightly from Ossiander's result in that her result uses the $L^2(P)$ metric on $\mathcal{T}$, whereas Theorem 4 uses the $L^2(\mu)$ metric on $\mathcal{T}$.

REFERENCES

Andrews, D. W. K. (1985) A nearly independent, but non–strong mixing, triangular array. *Journal of Applied Probability* **22**, 729–731.

Andrews, D. W. K. (1988a) Chi–square diagnostic tests for econometric models: Introduction and applications. *Journal of Econometrics* **37**, 135–156.

Andrews, D. W. K. (1988b) Chi–square diagnostic tests for econometric models: Theory. *Econometrica* **56**, 1419–1453.

Andrews, D. W. K. (1989a) Asymptotics for semiparametric econometric models: I. Estimation. Cowles Foundation Discussion Paper No. 908, Cowles Foundation, Yale University.

Andrews, D. W. K. (1989b) Asympototics for semiparametric econometric models: II. Stochastic equicontinuity. Cowles Foundation Discussion Paper No. 909, Cowles Foundation, Yale University.

Andrews, D. W. K. (1989c) Asymptotics for semiparametric econometric models: III. Testing and examples. Cowles Foundation Discussion Paper No. 910, Cowles Foundation, Yale University.

Bierens, H. J. (1981) *Robust Methods and Asymptotic Theory in Nonlinear Econometrics.* Lecture Notes in Economics and Mathematical Systems, No. 192. Springer–Verlag, New York.

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

Carroll, R. J. (1982) Adapting for heteroscedasticity in linear models. *Annals of Statistics* **10**, 1224–1233.

Dehling, H. (1983) Limit theorems for sums of weakly dependent Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **63**, 393–432.

Dehling, H. and W. Philipp (1982) Almost sure invariance principles for weakly dependent random vectors. *Annals of Probability* **10**, 689–701.

Dudley, R. M. (1978) Central limit theorems for empirical measures. *Annals of Probability* **6**, 899–929.

Dudley, R. M. (1985) An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. *Springer Lecture Notes in Mathematics* **1153**, 141–178. Springer, New York.

Dudley, R. M. (1987) Universal Donsker classes and metric entropy. *Annals of Probability* **15**, 1306–1326.

Edmunds, D. E. and V. B. Moscatelli. (1977) Fourier approximations and embeddings in Sobolev space. *Dissertationes Mathematicae* **145**, 1–46.

Gallant, A. R. (1987) *Nonlinear Statistical Models.* Wiley, New York.

Gallant, A. R. and H. White (1988) *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* Basil Blackwell, New York.

Hoffmann–Jorgensen, J. (1985) Stochastic processes on Polish spaces. Unpublished manuscript, University of Aarhus, Denmark.

Ibragimov, I. A. (1962) Some limit theorems for stationary processes. *Theory of Probability and its Applications* 7, 349–382.

Kolmogorov, A. N. and V. M. Tihomirov (1961) ϵ–entropy and ϵ–capacity of sets in functional spaces. *American Mathematical Society Translations*, Ser. 2, Vol. 17, 277–364.

McLeish, D. L. (1975a) A maximal inequality and dependent strong laws. *Annals of Probability* **3**, 829–839.

McLeish, D. L. (1975b) Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32**, 165–178.

Ossiander, M. (1987) A central limit theorem under metric entropy with L_2 bracketing. *Annals of Probability* **15**, 897–919.

Pakes, A. and D. Pollard (1986) The asymptotics of simulation estimators. Unpublished manuscript. Department of Statistics, Yale University.

Pollard, D. (1979) General chi–square goodness–of–fit tests with data dependent cells. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **50**, 317–331.

Pollard, D. (1984) *Convergence of Stochastic Processes.* Springer Verlag, New York.

Pollard, D. (1985) New ways to prove central limit theorems. *Econometric Theory* 1, 295–314.

Pollard, D. (1988) *Empirical Processes: Theory and Applications.* SIAM, Philadelphia, to appear.

Pollard, D. (1989) A maximal inequality for sums of independent processes under a bracketing condition. Unpublished manuscript, Department of Statistics, Yale University.

Robinson, P. M. (1987) Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* **55**,875–891.

Shorack, G. R. and J. A. Wellner (1986) *Empirical Processes with Applications to Statistics.* Wiley, New York.

Stein, E. M. (1970) *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, Princeton, NJ.

Withers, C. S. (1981) Central limit theorems for dependent variables I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **57**, 509–534.

Wooldridge, J. (1986) Asymptotic properties of econometric estimators. Ph.D dissertation, Department of Economics, University of California, San Diego.

Zygmund, A. (1955) *Trigonometrical Series.* Dover, New York. First published, 1935.