# Center for Research on Economic and Social Theory
# Research Seminar in Quantitative Economics

## Discussion Paper

DEPARTMENT OF ECONOMICS

**University of Michigan**

Ann Arbor, Michigan 48109

[redacted]

A CENTRAL LIMIT THEOREM

WITH APPLICATIONS TO ECONOMETRICS*

E. Philip Howrey
Saul H. Hymans
The University of Michigan

December 1981

R-109.81

*The authors wish to thank the members of the Statistics Seminar and the Econometric Theory Seminar, both of the University of Michigan, for valuable contributions made following presentations of the earliest versions of this paper. We are especially grateful to Professors Bruce M. Hill, Christopher Sims, Michael Woodroofe, and an anonymous referee, who read an earlier version of the paper with extreme care. Any remaining errors, of course, are ours. Research on this topic was supported, in part, by a grant from the National Science Foundation.

## ABSTRACT

This paper is concerned primarily with the asymptotic distribution of the least squares estimator in a linear equation with stochastic regressors. We prove a central limit theorem dealing with a sequence of products of random variables. The theorem is then applied to show asymptotic normality of the least squares estimator in a wide variety of cases, including: a) autoregressive regressors, b) moving average regressors, c) lagged dependent variable regressors. The results are generalized to handle Aitken estimation with stochastic regressors, and instrumental variable estimation in simultaneous equation models.

## I. Motivation

This paper is concerned with the asymptotic distribution of the least squares estimator of $\beta$ in the regression model

$$y_t = \beta x_t + \varepsilon_t \qquad (t = 1, 2, \ldots, T)$$

where $\{\varepsilon_t\}$ is a sequence of independent, identically distributed (i.i.d.) random variables and $x_t$ is a scalar stochastic regressor.[1/] In particular, the asymptotic distribution of the stabilized least squares estimator

$$\sqrt{T}\,(\hat{\beta} - \beta) = \frac{1}{T^{-1}\sum x_t^2}\ \sqrt{T}\,T^{-1}\sum x_t \varepsilon_t$$

is derived under alternative assumptions about the stochastic process governing the generation of the regressor $x_t$. Provided that $T^{-1}\sum x_t^2$ has a finite, non-zero probability limit, it follows from the convergence theorem of Cramér [1946; p. 254] that $\sqrt{T}\,(\hat{\beta} - \beta)$ will be asymptotically normally distributed if $\sqrt{T}\,T^{-1}\sum x_t\varepsilon_t$ converges in distribution to normality. In Sections III and IV of the paper we state and prove a central limit theorem dealing with a stochastic sequence of the form $\{x_t\varepsilon_t\}$. Section V of the paper applies the general theorem to the regression model under alternative assumptions about the generation of $x_t$. The paper concludes with Section VI which discusses extensions of the basic results. Before turning to the theorem itself, we present a brief review of the existing literature relating to the central question of the paper.

## II. The Existing Literature

Most econometrics textbooks provide an explicit derivation of the asymptotic distribution of the least squares estimator only for the "fixed regressor" case. It is generally assumed that the regressor (or vector of regressors) is nonstochastic or, if stochastic, fully independent of the disturbance vector $\varepsilon$ in which case the asymptotic distribution of $\sqrt{T}(\hat{\beta} - \beta)$ is obtained conditional on the observed values of the regressor. For example, Theil [1971, pp. 380-1] uses the familiar Lindeberg-Lévy central limit theorem to prove the asymptotic normality of the least squares estimator for the fixed regressor case. Hannan [1961] considers what amounts to a system of seemingly unrelated regressions. Using the Liapunov form of the central limit theorem, he proves asymptotic normality of the least squares estimator conditional on regressors which satisfy a form of strong law convergence.

In connection with autoregressive models which contain lagged values of the dependent variable among the regressors, both Theil [1971, pp. 412-13] and Malinvaud [1966, p. 453], for example, state without proof theorems which assert that the least squares estimator is asymptotically normally distributed. They both cite Mann and Wald [1943] as the original reference for this result. More recent treatments of this problem include Koopmans, Rubin, and Leipnik [1950], Grenander and Rosenblatt [1957], and Durbin [1960]. Koopmans, Rubin and Leipnik were primarily concerned with the extension of the Mann and Wald results to the case where (nonstochastic) exogenous variables are present among the regressors. Moreover, as Durbin notes, the results given in Koopmans, Rubin, and Leipnik depend on a theorem attributed to Rubin [1948] the proof of which was never published. In their proof of the asymptotic distribution of the least squares estimator, Grenander and Rosenblatt refer to Diananda [1953] who in turn uses a result from Mann and Wald. A careful reading of the Durbin paper reveals that at a critical point in his proof, a result from Mann and Wald is again used.

Thus, while there appears at first glance to be several complete discussions of the asymptotic properties of the least squares estimator in the case of stochastic regressors, the fundamental theorem is that of Mann and Wald. The original proof of the theorem by Mann and Wald is inaccessible to many students of econometrics for a number of reasons. First, Mann and Wald maintain a level of generality which renders their notation and derivations cumbersome and difficult to follow. Second, their primary focus on the presence of a lagged dependent variable in the single-equation model makes it somewhat difficult to see the generalization of their result to a stochastic regressor other than a lagged dependent variable. In view of the importance of the stochastic regressor case in econometrics, a uniform treatment which is fairly simple and sufficiently general to include the classic Mann and Wald result as well as other stochastic regressor cases seems to be highly desirable.

## III. Statement of the Theorem

In the statement and proof of the theorem we use notation which translates naturally into the linear regression context in which the theorem is to be applied. Thus we are concerned with the expression $\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t$ which is in turn constructed from the sequences $\{x_t\}$ and $\{\varepsilon_t\}$. The following five assumptions specify the properties of $\{x_t\}$ and $\{\varepsilon_t\}$.

A.1) The stochastic sequence $\{\varepsilon_t\}$, $t \in [-T, T]$, is i.i.d. with mean zero and variance $\sigma^2$.

A.2) The stochastic sequence $\{v_t\}$, $t \in [-T, T]$, is i.i.d. with mean zero and variance $\delta^2$.

A.3) The random variables $\varepsilon_t$ and $v_{t-j}$ are stochastically independent for $j > 0$ and $j < -L$ where L is a finite positive integer.

A.4) The stochastic sequence $\{x_t\}$ is defined by $x_t = \sum_{j=0}^{\infty} \alpha_j v_{t-j}$,

where the $\alpha_j$ (not all zero) are scalar constants which are absolutely and hence square-summable, i.e., $\sum_{j=0}^{\infty} |\alpha_j|$ and $\sum_{j=0}^{\infty} \alpha_j^2$ are finite.

A.5) The stochastic sequences $\{\varepsilon_t\}$ and $\{v_t\}$ satisfy $E(|v_i v_j v_k \varepsilon_\ell \varepsilon_m \varepsilon_n|) \leq H < \infty$.

Theorem. Assumptions (A.1) - (A.5) imply that as $T \to \infty$, $\sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t \varepsilon_t$ converges in distribution to the Normal distribution with mean zero and variance $\sigma^2 \delta^2 A$, where $A = \sum_{j=0}^{\infty} \alpha_j^2$.

Before proving the theorem, we note that in the proof it will be shown that the sequence $\{x_t \varepsilon_t\}$ is uncorrelated, though not independent. It may be thought that uncorrelatedness (orthogonality) would be sufficient to establish the theorem. Unfortunately, this is not the case; there exists no general central limit theorem for uncorrelated random variales.[2]

A theorem similar to ours was proved by Moran [1947]. Our proof, like the proofs of Moran and Mann and Wald, relies on a form of the Liapunov central limit theorem for a doubly subscripted sequence of random variables. This theorem involves only a modest extension of the standard Lindeberg-Lévy central limit theorem. We first state a lemma that indicates the essential features of this extension.

Lemma. [Chung (1974, p. 199)] Let $\{\theta_{Tt},\ t=1, 2, \ldots, K(T),\ T=1, 2, \ldots\}$ denote a sequence of complex numbers where $K(T) \to \infty$ as $T \to \infty$. If this sequence satisfies the conditions

a) $\lim_{T\to\infty} [\max_{1 \leq t \leq K(T)} |\theta_{Tt}|] = 0$,

b) $\sum_{t=1}^{K(T)} |\theta_{Tt}| \leq M < \infty$, and

c) $\lim_{T\to\infty} [\sum_{t=1}^{K(T)} \theta_{Tt}] = \theta$, where $\theta$ is a finite complex number, then

$$\lim_{T\to\infty} \prod_{t=1}^{T} (1 + \theta_{Tt}) = \exp(\theta).$$

This lemma permits a simple, almost mechanical proof of the following result.

Liapunov Theorem. Let $\{Y_{Tt},\ t=1, 2, \ldots, K(T),\ T=1, 2, \ldots\}$ denote a sequence of random variables where $K(T) \to \infty$ as $T \to \infty$. If

i) $Y_{Ts}$ and $Y_{Tt}$ are independent for $t \neq s$,

ii) $E(Y_{Tt}) = 0$ for all T, t,

iii) $\mathrm{Var}\,(Y_{Tt}) = \sigma_{Tt}$ with $\sum_{t=1}^{K(T)} \sigma_{Tt} \to \sigma^2$ as $T \to \infty$,

iv) $E(|Y_{Tt}|^3) = \gamma_{Tt}$ with $\sum_{t=1}^{K(T)} \gamma_{Tt} \to 0$ as $T \to \infty$,

then

$$Z_T = \sum_{t=1}^{K(T)} Y_{Tt} \overset{D}{\to} N(0, \sigma^2),$$

i.e., $Z_T$ converges in distribution to a normal random variable with mean zero and variance $\sigma^2$.

The proof of this theorem relies on the Taylor series expansion of the characteristic function of $Y_{Tt}$:

$$\phi_{Tt}(s) = 1 - \sigma_{Tt}s^2/2 + \lambda_{Tt}\,\gamma_{Tt}s^3/6 \qquad |\lambda_{Tt}| < 1.$$

The characteristic function of $Z_t$ is thus

$$\phi_T(s) = \prod_{t=1}^{K(T)} \phi_{Tt}(s) = \prod_{t=1}^{K(T)} (1 + \theta_{Tt})$$

where $\theta_{Tt} = -\sigma_{Tt}s^2/2 + \lambda_{Tt}\gamma_{Tt}s^3/6$. It is straight forward to verify that conditions iii) and iv) in the statement of this theorem imply that $\theta_{Tt}$ satisfies the conditions of the previous lemma[3] so that

$$\phi_T(s) \to \exp(-\sigma^2 s^2/2) \text{ as } T \to \infty, \text{ and } Y_T \xrightarrow{D} N(0, \sigma^2).$$

## IV. Proof of the Theorem

The proof of the theorem stated at the beginning of Section III requires the following six results.

R.1) $x_t$ has mean zero and a finite variance. This follows directly from (A.2) and (A.4).

R.2) $x_t$ and $\varepsilon_{t+\ell}$ are stochastically independent for $\ell > 0$. This follows from (A.3) and the definition of $x_t$.

R.3) $E(x_t\varepsilon_t) = 0$ and $Var(x_t\varepsilon_t) = E(x_t\varepsilon_t)^2 = \sigma^2 Var\ x_t$. These follow from (R.2) with $\ell = 0$, (R.1), and (A.1).

R.4) $E(x_t\varepsilon_t x_{t+\ell}\varepsilon_{t+\ell}) = 0$ for $\ell > 0$. To see this, we observe that

$$\begin{aligned} E(x_t\varepsilon_t x_{t+\ell}\varepsilon_{t+\ell}) &= E[(x_t\varepsilon_t x_{t+\ell})E(\varepsilon_{t+\ell}|x_t\varepsilon_t x_{t+\ell})] \\ &= E[(x_t\varepsilon_t x_{t+\ell})(0)] \\ &= 0, \end{aligned}$$

with the conditional expectation of $\varepsilon_{t+\ell}$ being zero by virtue of (A.1) and (R.2). Note that this implies that $\{x_t\varepsilon_t\}$ is an uncorrelated sequence.

Let

$$x_t' = \sum_{j=0}^{q} \alpha_j v_{t-j}$$

and

$$x_t'' = \sum_{j=q+1}^{\infty} \alpha_j v_{t-j}$$

so that

$$x_t = x_t' + x_t''.$$

R.5) It is clear that (R.1) - (R.4) apply with $x_t$ replaced by $x_t'$ or by $x_t''$.

R.6) $\text{Var } x_t = \delta^2 \sum_{j=0}^{\infty} \alpha_j^2 = \delta^2 A$

$$\text{Var } x_t' = \delta^2 \sum_{j=0}^{q} \alpha_j^2 = \delta^2 A_q, \text{ where } A_q = \sum_{j=0}^{q} \alpha_j^2$$

$$\text{Var } x_t'' = \delta^2 \sum_{j=q+1}^{\infty} \alpha_j^2 = \delta^2 (A - A_q).$$

These follow from (A.1) and (A.4). The latter implies that A is finite and positive and, of course, $A > A_q$.

We proceed now to the formal proof of the theorem.

<u>Proof</u>:

1) In the definition of $x_t'$ and $x_t''$ choose

$$q = T^{\theta}, \qquad 0 < \theta < 1 \text{ [4/]}$$

It follows that

$$\sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t \varepsilon_t = \sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t' \varepsilon_t + \sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t'' \varepsilon_t$$

where

i) $E(\sqrt{T}\, T^{-1} \sum x_t'' \varepsilon_t) = 0$ (R.3, R.5)[5/]

and

ii) $\text{Var}(\sqrt{T}\, T^{-1} \sum x_t'' \varepsilon_t)$

$$= \frac{1}{T} \text{Var} \sum x_t'' \varepsilon_t$$

$$= \frac{1}{T} \sum \text{Var}(x_t'' \varepsilon_t) \qquad \text{(R.4, R.5)}$$

$$= \frac{1}{T} \sum \sigma^2 \text{ Var } x_t'' \qquad \text{(R.3, R.5)}$$

$$= \frac{1}{T} \sigma^2 \sum_{t=1}^{T} \delta^2 (A - A_q) \qquad \text{(R.6)}$$

$$= \sigma^2 \delta^2 (A - A_q) .$$

Since

$$A_q = \sum_{j=0}^{q} \alpha_j^2 \qquad \text{and} \qquad q = T^\theta ,$$

$$\lim_{T\to\infty} A_q = \lim_{T\to\infty} \sum_{j=0}^{T^\theta} \alpha_j^2 = \sum_{j=0}^{\infty} \alpha_j^2 = A .$$

Hence,

$$\lim_{T\to\infty} \text{Var } (\sqrt{T}\, T^{-1} \sum x_t'' \varepsilon_t) = \lim_{T\to\infty} \sigma^2 \delta^2 (A - A_q) = 0 .$$

Thus,

$$\text{Plim } \sqrt{T}\, T^{-1} \sum x_t'' \varepsilon_t = 0$$

and the asymptotic distribution of $\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t$ is the same as that of $\sqrt{T}\, T^{-1} \sum x_t' \varepsilon_t$, which follows from the Convergence Theorem of Cramér [1946, p. 254]. We shall write

$$\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t \overset{D}{\to} \sqrt{T}\, T^{-1} \sum x_t' \varepsilon_t$$

to indicate that the lefthand term has the same asymptotic distribution as the righthand term.

2) Now choose M such that[6/]

$$M = T^{\mu}, \qquad \theta < \mu < 1$$

and define K as

$K = [T/M]$, where the notation $[T/M]$ signifies the largest integer less than or equal to $T/M$.

Obviously, the product KM is (an integer) always less than or equal to T and we have

$$KM + P = T, \qquad 0 \leq p < M.$$

Thus the T elements of the sum $\sum_{t=1}^{T} x_t'\varepsilon_t$ can be rewritten as the sum of K partial sums each containing M products of the form $x_t'\varepsilon_t$ and a remainder sum containing P products of the form $x_t'\varepsilon_t$:

$$\sum_{t=1}^{T} x_t'\varepsilon_t = (x_1'\varepsilon_1 + x_2'\varepsilon_2 + \dots + x_M'\varepsilon_M)$$

$$+ (x_{M+1}'\varepsilon_{M+1} + \dots + x_{2M}'\varepsilon_{2M})$$

$$+ (x_{2M+1}'\varepsilon_{2M+1} + \dots + x_{3M}'\varepsilon_{3M})$$

$$\vdots$$

$$+ (x_{(K-1)M+1}'\varepsilon_{(K-1)M+1} + \dots + x_{KM}'\varepsilon_{KM})$$

$$+ (x_{KM+1}'\varepsilon_{KM+1} + \dots + x_{KM+P}'\varepsilon_{KM+P}$$

$$= \sum_{k=1}^{K} \sum_{m=1}^{M} x_{(k-1)M+m}'\varepsilon_{(k-1)M+m} + \sum_{p=1}^{P} x_{KM+p}'\varepsilon_{KM+p}.$$

Now note that

i) $E(\sqrt{T}\, T^{-1} \sum_{p} x'_{KM+p}\varepsilon_{KM+p}) = 0$ (R.3, R.5)

and

ii) $\mathrm{Var}(\sqrt{T}\, T^{-1} \sum_{p} x'_{KM+p}\varepsilon_{KM+p})$

$$= \frac{1}{T} \sum_{p} \mathrm{Var}(x'_{KM+p}\varepsilon_{KM+p}) \qquad \text{(R.4, R.5)}$$

$$= \frac{1}{T} \sum_{p} \sigma^2 \mathrm{Var}\ x'_{KM+p} \qquad \text{(R.3, R.5)}$$

$$= \frac{1}{T} \sigma^2 \sum_{p} \delta^2 A_q \qquad \text{(R.6)}$$

$$= \frac{P}{T} \sigma^2 \delta^2 A_q .$$

But the definitions of P, K, and M imply

$$0 \leq P/T \leq (M-1)/T = (T^{\mu}-1)/T$$

so that

$$0 \leq \lim_{T\to\infty} P/T \leq \lim_{T\to\infty} (T^{\mu}-1)/T = 0$$

Hence,

$$\lim_{T\to\infty} \mathrm{Var}(\sqrt{T}\, T^{-1} \sum_{p} x'_{KM+p}\varepsilon_{KM+p}) = \lim_{T\to\infty} \frac{P}{T} \sigma^2 \delta^2 A_q = 0 .$$

Thus,

$$\mathrm{Plim}\ \sqrt{T}\, T^{-1} \sum_{p} x'_{KM+p}\varepsilon_{KM+p} = 0,$$

and

$$\sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t' \varepsilon_t \overset{D}{\to} \sqrt{T}\, T^{-1} \sum_{k=1}^{K} \left( \sum_{m=1}^{M} x'_{(k-1)M+m} \varepsilon_{(k-1)M+m} \right).$$

3) Now consider the term

$$\sum_{m=1}^{M} x'_{(k-1)M+m} \varepsilon_{(K-1)M+m} .$$

By the definitions of M and q,

$$\frac{M}{q} = \frac{T^{\mu}}{T^{\theta}} = T^{\mu-\theta} \to \infty$$

so that for sufficiently large T, $M > q + L = r$ where L is as specified in A.3. Hence the M terms in the above sum can be rewritten as the sum of the first $(M-r)$ terms and the remaining r terms;

$$\sum_{m=1}^{M} x'_{(k-1)M+m} \varepsilon_{(k-1)M+m}$$

$$= [x'_{(k-1)M+1} \varepsilon_{(k-1)M+1} + \dots + x'_{(k-1)M+(M-r)} \varepsilon_{(k-1)M+(M-r)}]$$

$$+ [x'_{(k-1)M+(M-r+1)} \varepsilon_{(k-1)M+(M-r+1)} + \dots + x'_{kM} \varepsilon_{kM}] .$$

Letting $W_k$ and $S_k$ denote the first and second sums, respectively, on the righthand side of the preceding equation,

$$\sum_{m=1}^{M} x'_{(k-1)M+m} \varepsilon_{(k-1)M+m} = W_k + S_k,$$

and

$$\sqrt{T}\, T^{-1} \sum x_t' \varepsilon_t \overset{D}{\to} \sqrt{T}\, T^{-1} \left( \sum^{K} W_k + \sum^{K} S_k \right).$$

4) We now show that Plim $\sqrt{T}\, T^{-1} \sum_{k=1}^{K} S_k = 0$.

i) $E(\sqrt{T}\, T^{-1} \sum_{k=1}^{K} S_k) = 0$

since $S_k$ is a sum of $x_t'\varepsilon_t$ products each of which has a zero mean.

ii) $\text{Var}\ (\sqrt{T}\, T^{-1} \sum_{k=1}^{K} S_k)$

$$= \frac{1}{T} \text{Var} \sum_{k=1}^{K} S_k .$$

But the $S_k$'s are mutually uncorrelated because they are sums of distinct $x_t'\varepsilon_t$ products which are uncorrelated. And each $S_k$ has variance r Var $x_t'\varepsilon_t$ since there are r uncorrelated terms in each $S_k$. Hence

$$\text{Var}\ (\sqrt{T}\, T^{-1} \sum_{k=1}^{K} S_k)$$

$$= \frac{1}{T} \sum_{k=1}^{K} \text{Var}\ S_k$$

$$= \frac{1}{T} \sum_{k=1}^{K} r\ \text{Var}\ x_t'\varepsilon_t$$

$$= \frac{1}{T} \sum_{k=1}^{K} r\ \sigma^2\ \text{Var}\ x_t' \qquad \text{(R.3, R.5)}$$

$$= \frac{1}{T}\, r\ \sigma^2 \sum_{k=1}^{K} \delta^2 A_q$$

$$= \frac{1}{T}\, r\ \sigma^2\ K\ \delta^2 A_q = \frac{rK}{T}\, \sigma^2 \delta^2 A_q .$$

But

$$\frac{rK}{T} = \frac{(T^{\theta} + L)[T/M]}{T} < \frac{(T^{\theta} + L)(T/M)}{T} = \frac{(T^{\theta} + L)T^{1-\mu}}{T} = T^{\theta-\mu} + LT^{-\mu}$$

so that

$$\lim_{T\to\infty} \frac{rK}{T} = 0 .$$

Thus

$$\lim_{T\to\infty} \mathrm{Var}(\sqrt{T}\, T^{-1} \sum_k S_k) = \lim_{T\to\infty} \frac{rK}{T}\sigma^2\delta^2 A_q = 0$$

and

$$\mathrm{Plim}\, (\sqrt{T}\, T^{-1} \sum_k S_k) = 0$$

which implies

$$\sqrt{T}\, T^{-1} \sum_t x_t' \varepsilon_t \overset{D}{\to} \sqrt{T}\, T^{-1} \sum_{k=1}^{K} W_k .$$

5) Now consider two successive W's, say $W_1$ and $W_2$.

$$W_1 = x_1'\varepsilon_1 + x_2'\varepsilon_2 + \dots + x_{M-r}'\varepsilon_{M-r}$$

$$W_2 = x_{M+1}'\varepsilon_{M+1} + \dots + x_{2M-r}'\varepsilon_{2M-r} .$$

The last term in $W_1$ involves $x_{M-r}'$, while the first term in $W_2$ involves $x_{M+1}'$. But

$$x_{M-r}' = \sum_{j=0}^{q} \alpha_j v_{M-r-j}$$

$$= \alpha_0 v_{M-r} + \alpha_1 v_{M-r-1} + \dots + \alpha_q v_{M-r-q}$$

and

$$x'_{M+1} = \sum_{j=0}^{q} \alpha_j v_{M+1-j}$$

$$= \alpha_0 v_{M+1} + \alpha_1 v_M + \ldots + \alpha_q v_{M+1-q} .$$

By construction, the $W_k$ contain distinct $v_t$'s; i.e., no two $W_k$'s contain any of the same $v_t$'s. And, of course, no two $W_k$'s contain any of the same $\varepsilon_t$'s. Moreover, the time separation between an $\varepsilon_t$ in $W_k$ and a $v_{t+s}$ in $W_{k'}$ is at least L + 1. Hence, while each $W_k$ itself is a sum of uncorrelated but dependent $x'_t\varepsilon_t$ products, the sequence $\{W_k\}$ k = 1, 2, ..., K is a sequence of <u>independent</u> random variables; in fact, for a given value of T, an i.i.d. sequence, since the $W_k$ are identically constructed across k.

6) We now consider the doubly subscripted sequence of random variables defined by

$$Z_{Tk} = W_k/\sqrt{T}, \qquad k = 1, 2, \ldots, K(T), \quad T = 1, 2, \ldots$$

so that $\sqrt{T}\, T^{-1} \sum_t x'_t\varepsilon_t \overset{D}{\rightarrow} \sum_k Z_{Tk}$. Note further that

i) $Z_{Tk}$ and $Z_{Ts}$ are independent if $k \neq s$,

ii) $E(Z_{Tk}) = 0$,

and

iii) $\sigma_{Tk} = \text{Var}(Z_{Tk})$

$$= \text{Var}(Z_{T1})$$

$$= \text{Var}\Big(\sum_{m=1}^{M-r} x'_m\varepsilon_m/\sqrt{T}\Big)$$

$$= \sum_{m=1}^{M-r} T^{-1}\ \text{Var}(x'_m\varepsilon_m) \qquad \text{(R.4, R.5)}$$

$$= \sum_{m=1}^{M-r} T^{-1}\sigma^2 Var(x_m') \qquad (R.3,\ R.5)$$

$$= \sigma^2 \sum_{m=1}^{M-r} \delta^2 A_q/T \qquad (R.6)$$

$$= (M-r)\sigma^2\delta^2 A_q/T.$$

It follows that

$$\sum_{k=1}^{K} \sigma_{Tk} = K(M-r)\sigma^2\delta^2 A_q/T.$$

In view of the definitions of K, M, and r, we have

$$\lim_{T\to\infty} K(M-r)/T = \lim_{T\to\infty} KM/T - \lim_{T\to\infty} rK/T$$

$$= \lim_{T\to\infty} KM/T = \lim_{T\to\infty} (1-R/T) = 1;$$

thus

$$\lim_{T\to\infty} \sum_{k=1}^{k} \sigma_{Tk} = \sigma^2\delta^2 A.$$

We conclude that $\{Z_{Tk}\}$ satisfies the first three conditions of the Liapunov Central Limit Theorem.

7) To complete our proof, we must examine

$$\sum_{k=1}^{K} \gamma_{Tk} = K\,\gamma_{T1}$$

$$= KE(|Z_{T1}|^3)$$

$$= KE(|\sum_{m=1}^{M-r} x_m'\varepsilon_m/\sqrt{T}|^3)$$

$$< KE([\sum_{m=1}^{M-r} |x_m' \varepsilon_m| / \sqrt{T}]^3)$$

$$< K(M/\sqrt{T})^3 H_T$$

where $H_T = \max E(|x_h' x_i' x_j' \varepsilon_\ell \varepsilon_m \varepsilon_n|)$. In view of the definition of $x_i'$, we have

$$H_T = \max E(|\sum_{r=1}^{q} \sum_{s=1}^{q} \sum_{t=1}^{q} \alpha_r \alpha_s \alpha_t v_{h-r} v_{i-s} v_{j-t} \varepsilon_\ell \varepsilon_m \varepsilon_n|)$$

$$< \sum_r \sum_s \sum_t \alpha_r \alpha_s \alpha_t H \qquad \text{(A.5)}$$

$$< q^3 \alpha H$$

where $\alpha = \max |\alpha_j^3|$. Thus

$$\sum_{k=1}^{K} \gamma_{Tk} < K(M/\sqrt{T})^3 q^3 \alpha H$$

$$= \alpha H[T/M](M/\sqrt{T})^3 q^3$$

$$< \alpha H(T/M) M^3 T^{3\theta} / T^{3/2}$$

$$= \alpha H(T) M^2 T^{3\theta} / T^{3/2}$$

$$< \alpha H(T) T^{2\mu} T^{3\theta} / T^{3/2}$$

$$= \alpha H T^{2\mu + 3\theta - 1/2}.$$

For $\theta < \mu < 1/10$, $2\mu + 3\theta - 1/2 < 0$, so that $\sum \gamma_{Tk} \to 0$ as $T \to \infty$. Hence the fourth condition of the Liapunov Central Limit Theorem is satisfied and we conclude that

$$\sum_{k=1}^{K} Z_{Tk} \overset{D}{\to} N(0, \sigma^2 \delta^2 A)$$

and since

$$\sqrt{T}\, T^{-1} \sum_{t=1}^{T} x_t \varepsilon_t \overset{D}{\rightarrow} \sum_{k=1}^{K} Z_{Tk},$$

it follows, finally, that

$$\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t \overset{D}{\rightarrow} N(0, \sigma^2 \delta^2 A).$$

## V. Econometric Applications

The central limit theorem of the preceding section is directly applicable to a number of specific models that are commonly encountered in econometrics. This section is devoted to a discussion of the following special cases: 1) the regressor $x_t$ is generated by an autoregressive process, 2) $x_t$ is generated by a finite moving average process, 3) $x_t$ is an i.i.d. sequence, 4) $x_t$ is a lagged dependent variable, 5) $x_t$ is an endogenous variable in a Wold recursive system, 6) $x_t$ is an exogenous variable to be used as an instrument in a simultaneous equations model.

In each of the cases considered below the estimator to be examined is of the form

$$\sqrt{T}\,(\hat{\beta} - \beta) = D_T \sqrt{T}\, T^{-1} \sum x_t \varepsilon_t .$$

The asymptotic normality of $\sqrt{T}\,(\hat{\beta} - \beta)$ is obtained by applying the central limit theorem to $\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t$ after observing that $D_T$ converges in probability to a finite non-zero constant. In the first three cases that are considered, $D_T$ is given by

$$D_T^{-1} = T^{-1} \sum x_t^2 .$$

With $x_t$ defined by (A.2) and (A.4),[7] it is not difficult to verify that

$$\text{plim } T^{-1} \sum x_t^2 = \delta^2 A$$

provided $v_t$ has a finite fourth moment.[8] This means that the sample second moment is a consistent estimator of the variance of $x_t$ — an assumption commonly made in the econometric literature. If $\{x_t\}$ is an i.i.d. sequence as in case 3, second moment consistency follows immediately from the weak law of large numbers. If $\{x_t\}$ is a correlated sequence, second moment consistency is not so obvious. However, it is true that the sample variance is a consistent estimator of the population variance if a finite fourth moment is assumed.

1) Autoregressive $x_t$. In this case the model is written as

$$y_t = \beta x_t + \varepsilon_t$$

where

i) $x_t = \rho x_{t-1} + v_t, \ |\rho|<1$

ii) A.1, A.2, A.3, and A.5 are satisfied.[9]

From i) it follows that the moving average representation of $x_t$ is

$$x_t = \sum_{j=0}^{\infty} \rho^j v_{t-j}$$

so that $\alpha_j = \rho^j$ in (A.4). Since $|\rho|<1$, the $\alpha_j$ are square-summable and $A = \sum \alpha_j^2 = 1/(1 - \rho^2)$. Thus the assumptions of the theorem are satisfied and we conclude that

$$\sqrt{T}\, T^{-1} \sum \ x_t \varepsilon_t \overset{D}{\rightarrow} N(0, \sigma^2 \delta^2 A)$$

and, since

$$\text{plim } D_T = \delta^2/(1 - \rho^2),$$

we have

$$\sqrt{T}\,(\hat{\beta} - \beta) \overset{D}{\rightarrow} N[0,\ (1 - \rho^2)\sigma^2/\delta^2).$$

We merely note that a similar result holds if $x_t$ is generated by a stable autoregressive process of any finite order. The moving-average representation as well as the expression for the variance of $x_t$ ($\delta^2 A$) are more complicated but no further difficulties are involved in the consideration of higher order autoregressive processes.

2) Finite Moving Average $x_t$. With a finite moving average regressor the model is written as

$$y_t = \beta x_t + \varepsilon_t$$

where

i) $x_t = \sum_{j=0}^{q} \alpha_j v_{t-j}$

ii) A.1, A.2, A.3, and A.5 are satisfied.

Since $x_t$ is already in moving average form, the central limit theorem applies.[10/] Hence $\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t \overset{D}{\rightarrow} N(0,\ \sigma^2\delta^2 A)$ where $A = \sum_{j=0}^{q} \alpha_j^2$. In addition, plim $D_T = \delta^2 A$ so that $\sqrt{T}\,(\hat{\beta} - \beta) \overset{D}{\rightarrow} N[0,\ \sigma^2/(\delta^2 A)]$.

3) I.i.d. $x_t$. This is a special case of moving average $x_t$ where $q = 0$ and $\alpha_0 = 1$. Hence we conclude immediately that $\sqrt{T}\,(\hat{\beta} - \beta) \overset{D}{\rightarrow} N(0,\ \sigma^2/\delta^2)$.

4) $x_t = y_{t-1}$. In the lagged dependent variable case, the model is

$$y_t = \beta x_t + \varepsilon_t \qquad |\beta|<1$$

with

i) $x_t = y_{t-1} = \sum_{j=1}^{\infty} \beta^{j-1} \varepsilon_{t-j} = \sum_{j=0}^{\infty} \beta^j v_{t-j}$

ii) $v_t = \varepsilon_{t-1} \sim \text{i.i.d.}\ (0, \sigma^2)$.

It is clear from the definition of $v_t$ that $\varepsilon_t$ and $v_{t-j}$ are independent for all $j > 0$ and $j < -1$. Under the restriction that $|\beta|<1$, it follows that $\text{Var}\ x_t = \sigma^2/(1 - \beta^2)$ and this case is thus equivalent to case 1 with $L = 1$.[11/] Hence we conclude that

$$\sqrt{T}\ (\hat{\beta} - \beta) \overset{D}{\rightarrow} N[0,\ (1 - \beta^2)].$$

Note that the model $y_t = \beta x_t + \varepsilon_t$ with $x_t = y_{t-2}$ can be handled in exactly the same fashion and would correspond to the particular case $L = 2$.

5) Wold Recursive System. Suppose that $x_t$ is an endogenous variable in the recursive system

$$x_t = \gamma z_t + \eta_t$$

$$y_t = \beta x_t + \varepsilon_t$$

where

i) $z_t \sim \text{i.i.d.}\ (0, \sigma_z^2)$

ii) $\eta_t \sim \text{i.i.d.}\ (0, \sigma_\eta^2)$ and independent of $z_{t-j}$ for all $j$

iii) $\varepsilon_t \sim \text{i.i.d.}\ (0, \sigma^2)$

iv) $\varepsilon_t$ is independent of $\eta_{t-j}$ and $z_{t-j}$ for all $j > 0$ and $j < -L$.

It follows that $v_t = \gamma z_t + \eta_t$ is i.i.d. $(0, \sigma_v^2)$ where $\sigma_v^2 = \gamma^2\sigma_z^2 + \sigma_\eta^2$ and $\varepsilon_t$ is independent of $v_{t-j}$ for all $j > 0$. This case is therefore equivalent to Case 3 and we conclude that

$$\sqrt{T}\,(\hat{\beta} - \beta) \xrightarrow{D} N(0,\ \sigma^2/\sigma_v^2).$$

We note in passing that if assumption i) is relaxed to allow the exogenous variable $z_t$ to be generated by either an autoregressive or a moving average process, $x_t$ is no longer of the form postulated in (A.4). For example, suppose $z_t$ is generated by

$$z_t = \rho z_{t-1} + \xi_t$$

where $\xi_t \sim$ i.i.d. $(0, \sigma_\xi^2)$ and $\varepsilon_t$ and $\xi_{t-j}$ are independent for all $j > 0$ and $j < -L$. Then $x_t$ becomes

$$x_t = \eta_t + \gamma \sum_{j=0}^{\infty} \rho^j\, \xi_{t-j}$$

which is not directly of the form $\sum \alpha_j\, v_{t-j}$. It would not be difficult, however, to modify our theorem to accomodate such a case.

6) Simultaneous Equations Model. Consider a single equation

$$y_t = \beta y_t^* + \varepsilon_t$$

embedded in a simultaneous system where $y_t^*$ is also an endogenous variable. If $x_t$ is an exogenous variable, an instrumental variable estimator of $\beta$ is

$$\hat{\beta} = \beta + (\sum x_t y_t^*)^{-1} \sum x_t \varepsilon_t$$

and

$$\sqrt{T}\,(\hat{\beta} - \beta) = D_T \sqrt{T}\, T^{-1} \sum x_t \varepsilon_t$$

where

$$D_T^{-1} = T^{-1} \sum x_t y_t^*.$$

If $\varepsilon_t \sim$ i.i.d. $(0, \sigma^2)$ and $x_t$ is a) i.i.d., b) autoregressive, or c) moving average, the conditions of the central limit theorem will be satisfied. Therefore

$$\sqrt{T}\, T^{-1} \sum x_t \varepsilon_t \overset{D}{\to} N[0, \sigma^2 \operatorname{Var}(x_t)].$$

Provided that $D_T$ converges in probability to a finite positive constant, say Q, we conclude that

$$\sqrt{T}\, (\hat{\beta} - \beta) \overset{D}{\to} N[0, \sigma^2 \operatorname{Var}(x_t)/Q^2].$$

## VI. Extensions and Conclusions

The central limit theorem of Section III is readily applied to establish asymptotic normality of the Aitken estimator corresponding to a regression equation with an autoregressive error term. We present the result for the case of first-order autoregression; the generalization to any finite order stable autoregressive process is immediately apparent. Suppose that

$$y_t^* = \beta x_t^* + u_t$$

where

i) $u_t = \rho u_{t-1} + \varepsilon_t \qquad |\rho| < 1$

ii) A.1, A.2, and A.3 are satisfied[12/]

iii) the stochastic sequence $\{x_t^*\}$ is defined by $x_t^* = \sum_{j=0}^{\infty} \alpha_j^* v_{t-j}$

where the $\alpha_j^*$ (not all zero) are scalar constants and absolutely summable, and $\{\varepsilon_t\}$ and $\{v_t\}$ satisfy (A.5).

Let

$$y_t = y_t^* - \rho y_{t-1}^*$$

$$x_t = x_t^* - \rho x_{t-1}^*$$

so that the original equation may be transformed to yield:

$$y_t = \beta x_t + \varepsilon_t.$$

Now observe that

$$x_t = \sum_{j=0}^{\infty} \alpha_j^* v_{t-j} - \rho \sum_{j=1}^{\infty} \alpha_{j-1}^* v_{t-j}$$

$$= \sum_{j=0}^{\infty} \alpha_j v_{t-j}$$

where

$$\alpha_0 = \alpha_0^*$$

$$\alpha_j = \alpha_j^* - \rho\alpha_{j-1}^*, \text{ for } j > 1.$$

Further,

$$\alpha_j^2 = \begin{cases} \alpha_0^{*2}, \text{ for } j = 0 \\ \alpha_j^{*2} + \rho^2\alpha_{j-1}^{*\,2} - 2\rho\alpha_{j-1}^*\alpha_j^*, \text{ for } j > 1 \end{cases}$$

so that

$$\sum_{j=0}^{\infty} \alpha_j^2 = \sum_{j=0}^{\infty} \alpha_j^{*2} + \rho^2 \sum_{j=1}^{\infty} \alpha_{j-1}^{*\,2} - 2\rho \sum_{j=1}^{\infty} \alpha_{j-1}^*\alpha_j^*.$$

By assumption, the first two sums on the righthand side of the preceding equation are finite. With respect to the third sum, it follows from Schwarz's Inequality that

$$\left|\sum_{j=1}^{\infty} \alpha_{j-1}^* \alpha_j^*\right| < \left(\sum_j \alpha_{j-1}^{*\,2} \sum_j \alpha_j^{*2}\right)^{1/2}$$

so that absolute and hence square-summability of the $\alpha_j^*$ implies finiteness of $\sum \alpha_{j-1}^*\alpha_j^*$. Thus, the $\alpha_j$ are square-summable. It follows that the transformed equation

$$y_t = \beta x_t + \varepsilon_t$$

satisfies all of the conditions for the central limit theorem of Section III to be applied and

$$\sqrt{T}\,(\hat{\beta} - \beta) = \frac{1}{T^{-1}\sum x_t^2} \sqrt{T}\, T^{-1} \sum x_t \varepsilon_t \xrightarrow{D} N[0,\ \sigma^2/(\delta^2 A)]$$

where

$$E\ \varepsilon_t^2 = \sigma^2,\ E\ v_t^2 = \delta^2,\ \text{and}\ \sum_j \alpha_j^2 = A.$$

But when expressed in terms of the original variables,

$$\sqrt{T}\,(\hat{\beta} - \beta) = \frac{1}{T^{-1}\sum (x_t^* - \rho x_{t-1}^*)^2} \quad \sqrt{T}\, T^{-1} \sum (x_t^* - \rho x_{t-1}^*)(u_t - \rho u_{t-1})$$

so that $\hat{\beta}$ is the Aitken estimator of $\beta$ [13/]

Up to now we have restricted attention to the case of a single explanatory variable. Our results, however, can be extended to the multiple regression model without difficulty. We first state the multivariate analogue of the central limit theorem of Section III and then show how the result would be used in practice.

The assumptions that underlie the multivariate central limit theorem are as follows.

A'.1) $\varepsilon_t \sim$ i.i.d. $(0, \sigma^2)$.

A'.2) $V_t \sim$ i.i.d. $(0, \Delta)$ where $V_t$ is a $(P \times 1)$ vector.

A'.3) $\varepsilon_t$ is independent of (each element of) $V_{t-j}$ for $j > 0$ and $j < -L$.

A' .4) The random vector $X_t'$ is defined as

$$X_t' = \sum_{j=0}^{\infty} V_{t-j}' \, D(\underline{\alpha}_j)$$

where $D(\underline{\alpha}_j)$ denotes a diagonal matrix with elements of the vector $\underline{\alpha}_j' = (\alpha_{1j}\ \alpha_{2j}\ \ldots\ \alpha_{Pj})$ on the diagonal. The sequence of vectors $\{\underline{\alpha}_j\}$ is assumed to satisfy the condition

$$\sum_{j=0}^{\infty} \underline{\alpha}_j \, \underline{\alpha}_j' = A,$$

where A is a non-null matrix of finite constants.

A' .5) The stochastic sequences $\{\varepsilon_t\}$ and $\{v_{jt}\}$ satisfy $E(|v_{jr}v_{js}v_{jt}\varepsilon_\ell\varepsilon_m\varepsilon_n|) \leq H < \infty$, where $v_{jt}$ is the jth element of $V_t$.

It is readily apparent that the assumptions are generalizations of those in Section III and guarantee that each element of the vector $X_t$ satisfies the conditions which were previously postulated for the scalar $x_t$.

<u>Theorem</u>. Assumptions A' .1-A' .5 imply that as $T \to \infty$, $\sqrt{T}\, T^{-1} \sum X_t'\varepsilon_t$ converges in distribution to the P-variate normal with mean vector 0 and covariance matrix $\sigma^2(\Delta * A)$, where $\Delta * A$ denotes the element by element product of $\Delta$ and A (each of which is P x P).

A proof of this theorem is obtained by going through the steps of Section IV for the vector case. Rather than do this here, we simply show how the covariance matrix of $\sqrt{T}\, T^{-1} \sum X_t'\varepsilon_t$ is obtained. Since $E(X_t'\varepsilon_t) = 0$ it follows that the covariance matrix (denoted in general by $\Omega$) of $X_t'\varepsilon_t$ is

$$\Omega_{X_t'\varepsilon_t} = E(\varepsilon_t^2\, X_t X_t') = \sigma^2\, \Omega_{X_t'}.$$

From the definition of $X_t'$, the covariance matrix of $X_t'$ is

$$\Omega_{X_t'} = \sum_{j=0}^{\infty} E\{D(\underline{\alpha}_j)\, V_{t-j}\, V_{t-j}'\, D(\underline{\alpha}_j)\}$$

$$= \sum_{j=0}^{\infty} D(\underline{\alpha}_j)\, \Delta\, D(\underline{\alpha}_j).$$

Further, it can be shown that

$$\sum_{j=0}^{\infty} D(\underline{\alpha}_j)\, \Delta\, D(\underline{\alpha}_j) = [\delta_{i\ell}\; a_{i\ell}] \qquad i, \ell = 1, 2, \ldots, P$$

$$= \Delta * A$$

where $[a_{i\ell}] = A$ as defined in (A' .4) and, as a notational matter, $\Delta * A$ is used to denote the element by element product of $\Delta$ and $A$. We conclude that

$$\Omega_{X_t'\varepsilon_t} = \sigma^2\, (\Delta * A).$$

Further, since successive elements of the sum $\sum_t X_t'\varepsilon_t$ are uncorrelated,

$$\Omega(\sqrt{T}\; T^{-1} \sum X_t'\varepsilon_t) = \sigma^2 (\Delta * A).$$

Thus the assumption in (A' .4) guarantees that $\sqrt{T}\; T^{-1} \sum X_t'\varepsilon_t$ has a finite covariance matrix.

As an illustration of the use of this theorem in practice, consider the multiple regression model

$$y_t = X_t'\, \beta + \varepsilon_t \qquad (t = 1, 2, \ldots, T)$$

where $\beta$ is now a (P x 1) vector. The stabilized least squares estimator is given by

$$\sqrt{T}\, (\hat{\beta} - \beta) = (\frac{1}{T} \sum X_t X_t')^{-1} \quad (\sqrt{T}\; T^{-1} \sum X_t'\varepsilon_t).$$

Provided that i) $X_t'$ and $\varepsilon_t$ satisfy (A' .1) - (A' .5), and ii) $\frac{1}{T} \sum X_t X_t'$

converges in probability to $\Delta^*A$, and iii) $\Delta^*A$ is nonsingular, it follows that

$$\sqrt{T}\,(\hat{\beta} - \beta) \overset{D}{\rightarrow} N[0,\ \sigma^2(\Delta^*A)^{-1}].$$

Clearly, the elements of $X_t'$ can be any mixture of autoregressive, moving average, or lagged dependent variables which satisfy the assumptions (A'.2) - (A'.5).

As a final illustration we re-cast the preceding example in the matrix notation most used in the econometric literature. The matrix $\Delta^*A$ is, of course, the population covariance matrix associated with the vector of regressors $X_t'$. The form $\Delta^*A$ emphasizes the functional dependence of $X_t'$ on the sequence $\{V_{t-j}\}$. Ignore this dependence and denote the matrix $\Delta^*A$ by $M_x$. Write the multiple regression model in matrix form as

$$Y = X\beta + \varepsilon$$

where Y is (T x 1), X is (T x P), $\beta$ is (P x 1), and $\varepsilon$ is (T x 1). Assume

$$\text{Plim}\,(T^{-1}X'X) = M_x \text{ and } M_x \text{ non-singular.}$$

Then if the rows of the matrix X satisfy the assumptions of the central limit theorem, it follows that

$$\sqrt{T}\,(\hat{\beta} - \beta) \overset{D}{\rightarrow} N(0,\ \sigma^2 M_x^{-1}).$$

The extension of the multiple regression result to allow for Aitken estimation of the vector $\beta$ when $\varepsilon_t$ is a stable autoregressive process is entirely analogous to the extension already presented for the simple regression case.

In effect, this paper shows that under fairly general conditions, it is valid to assume asymptotic normality of the least squares (or Aitken) estimator in a multivariate, stochastic regressor, linear model -- just as most of us have done all along.

FOOTNOTES

1/ In the final section of the paper we extend our basic result to the case of a vector of regressors and also relax the assumption of an i.i.d. error structure. For expositional purposes, however, the bulk of the paper focuses on the simple regression model with i.i.d. errors.

2/ See Grenander and Rosenblatt (1957, pp. 180-1) for several examples of uncorrelated random variables whose stabilized means are not asymptotically Normal.

3/ In particular, it follows from Liapunov's inequality $\sigma_{Tt}^{3/2} < \gamma_{Tt}$ and assumption iv) that

$$\max_t \sigma_{Tt}^{3/2} < \max_t \gamma_{Tt} < \sum_t \gamma_{Tt} \to 0 \qquad \text{as } T \to \infty.$$

This verifies that condition a) of the Lemma is satisfied. Conditions b) and c) of the Lemma are also satisfied because, as $T \to \infty$,

$$\sum |\theta_{Tt}| < \sum \sigma_{Tt} s^2/2 + \sum \gamma_{Tt} s^3 \to \sigma^2 s^2/2$$

and

$$\sum \theta_{Tt} = -\sum \sigma_{Tt} s^2/2 + \sum \lambda_{Tt}\gamma_{Tt} s^3/6 \to -\sigma^2 s^2/2.$$

4/ Strictly speaking, q must be an integer and $q = T^\theta$ must be thought of as defining the largest integer less than or equal to $T^\theta$. For any choice of $\theta \in (0, 1)$, this causes no difficulty as long as T exceeds some finite value $T(\theta)$. Since we shall be letting T increase without limit, we choose to let the proof proceed a bit more clearly using $q = T^\theta$.

5/ This notation is used to indicate which of the previous results are used to obtain the current result.

6/ Again, $M = T^\mu$ should be thought of as defining the greatest integer less than or equal to $T^\mu$.

7/ The assumption that $x_t$ is generated by A.4 means that there is an infinite amount of pre-sample history. There are ways to get around this assumption by conditioning on initial values or by moving the origin of the sample as T increases. Such complications hardly seem worthwhile since most econometric applications would not likely involve quantitatively large $\alpha_j$ for large values of j.

8/ It is immediately apparent from (R.6) that $E(T^{-1} \sum x_t^2) = \delta^2 A$. A sufficient condition for $T^{-1} \sum x_t^2$ to converge in probability to $\delta^2 A$ is that $\lim_{T\to\infty} Var(T^{-1} \sum x_t^2) = 0$ or equivalently, for the case at hand, that $\lim_{T\to\infty} E[(T^{-1} \sum x_t^2)^2] = \delta^4 A^2$. If $(T^{-1} \sum x_t^2)^2$ is written in terms of the generating process $x_t = \sum \alpha_j v_{t-j}$, an examination of the expectation of the resulting expression indicates that the limiting variance of $T^{-1} \sum {x_t}^2$ is zero if $v_t$ has a finite fourth moment. A proof of this assertion is given by Fuller (1976), pp. 239-240.

9/ If $\varepsilon_t$ and $v_s$ are independent for all t and s, assumption A.5 will be satisfied if both $\varepsilon_t$ and $v_s$ have a finite third absolute moment.

10/ In the proof given in Section IV, q was represented as $q = T^\theta$. The purpose of this was to render the term $\sqrt{T}\, T^{-1} \sum x_t''$ negligible as $T \to \infty$. With $x_t$ defined as a finite moving average to begin with, there is no $x_t''$ term; i.e., $x_t \equiv x_t'$ and the first step of the proof can be eliminated.

11/ Assumption A.5 will be satisfied in this case if $\varepsilon_t$ has a finite sixth absolute moment.

12/ Assumptions A.1 and A.3 refer to $\{\varepsilon_t\}$, not $\{u_t\}$.

13/ Obviously, there is no need to be concerned with the so-called "first-observation problem" in this asymptotic context.

REFERENCES

Chung, K.L. (1974), A Course in Probability Theory, Second Edition, New York: Academic Press.

Cramér, Harold (1946), Mathematical Methods of Statistics. Princeton: Princeton University Press.

Diananda, P.H. (1953), "Some Probability Limit Theorems with Statistical Applications," Proceedings of the Cambridge Philosophical Society, 49, 239-246.

Durbin, J. (1960), "Estimation of Parameters in Time-Series Regression Models," Journal of the Royal Statistical Society, Series B, 22, 139-153.

Feller, William (1966), An Introduction to Probability Theory and its Applications, Volume II, New York: Wiley.

Fuller, Wayne A. (1976), Introduction to Statistical Time Series, New York: Wiley.

Grenander, Ulf and Murray Rosenblatt (1957), Statistical Analysis of Stationary Time Series. New York: Wiley.

Hannan, E.J. (1961), "A Central Limit Theorem for Systems of Regressions," Proceedings of the Cambridge Philosophical Society, 57, 583-588.

Koopmans, T.C., H. Rubin, and R.B. Leipnik (1950), "Measuring the Equation Systems of Dynamic Economics." Chapter 2 of Statistical Inference in Dynamic Economic Models, edited by T.C. Koopmans. New York: Wiley.

Malinvaud, E. (1966), Statistical Methods of Econometrics. Chicago: Rand McNally.

Mann, H.B. and A. Wald (1943), "On the Statistical Treatment of Linear Stochastic Difference Equations," Econometrica, 11, 173-220.

Moran, P.A.P. (1947), "Some Theorems on Time Series I," Biometrika, 34, 281-291.

Rubin, Herman (1948), "Some Results on the Asymptotic Distribution of Maximum - and Quasi-Maximum - Likelihood Estimates," Annals of Mathematical Statistics, 20, 598.

Theil, Henri (1971), Principles of Econometrics. New York: Wiley.