A Note on Heteroskedasticity Issues

by
Phoebus J. Dhrymes, Columbia University

October 1992

Discussion Paper Series No. 629

A Note on Heteroskedasticity Issues*

Phoebus J. Dhrymes

October 14, 1992

1 Introduction

The purpose of this paper is to clarify certain issues related to the incidence of heteroskadisticity in the General Linear Model (GLM); to provide simpler and more accessible proofs for a number of propositions, and to allow the results to stand **under conditions considerably less stringent that those hitherto available in the literature**. A systematic treatment of the heteroskedasticity problem in the context of the GLM is given in the very interesting paper by White (1980) and the follow-up paepr by Nicholls and Pagan (1983). White's context consists of stochastic regressors that are uncorrelated with the error process, but **no lagged dependent variables**; the context in Nicholl and Pagan consists **only** of lagged dependent variables, and the error process is assumed to be a **martingale difference**. If the underlying difference equation is stable, their dependent variable is also a **martingale difference**. Although this may appear to be a generalization of White's result, in fact it is not, since the admission of "exogenous" (contemporaneous) regressors will destroy the martingale difference property of the y-process.

In our paper the context is one of stochastic exogenous and lagged dependent explanatory variables, and an error process as in White. Thus, our context is a **generalization of the White context**, but it is at once more

* This is a preliminary draft. Not to be quoted without the author's permission. Comments are solicited.

I would like to thank H. White for an interesting discussion on the subject of this paper.

general and more restrictive than the one in Nicholls and Pagan, in that ours can accommodate lagged dependent and current exogenous variables, but the error process is one of independent non-identically distributed random variables, while in Nicholls and Pagan the error process is one of martingale differences.

Although a number of the results presented therein had been available in the statistics literature for some time past, e.g., Eicker (1963), (1967), Rao (1970), (1973), White's paper is quite important in having introduced them in the econometrics literature and having employed them in addressing certain important problems arising in empirical applications. On the other hand, the presentation is overly complex and constitutes a barrier to the wider understanding and appreciation of the material contained therein.

2 Formulation of the Problem

Consider the GLM

$$y_t = x_{t\cdot}\beta_0 + u_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

or more compactly, $y = X\beta_0 + u$, where β_0 is the **true** parameter vector, $x_{t\cdot}$ is an $(n+1)$-element **row** vector containing the explanatory variables, y, u and X are the corresponding observation vectors and matrix, respectively, etc. The problem is: (a) to establish the strong consistency of the OLS estimator of β and (b) to establish the limiting distribution of the OLS estimator of β, when the error process is one of zero mean independent random variables with

$$\mathrm{Cov}(u) = \Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2), \tag{2}$$

and (c) to produce a (strongly) consistent estimator of the covariance matrix of the limiting distribution noted above.

Following White (1980), we assume

(A.1) The sequence $\{z_{t\cdot} = (x_{t\cdot}, u_t) : t \geq 1\}$ is one of independent (but not necessarily identically distributed) random variables, such that

$$M_{xx} = \lim_{T\to\infty} \frac{1}{T} E(X'X), \quad E(x_{t\cdot}' u_t) = 0,$$

for all t,[1] and M_{xx} is a nonsingular matrix;

(A.2). the sequence $\{z_{t\cdot} = (x_{t\cdot}, u_t) : t \geq 1\}$ possesses **fourth order moments obeying**

$$E(|z_{ti}z_{tj}z_{tk}z_{tr}|) \leq ct^{\alpha} \quad \alpha \in [0,1).$$

3 Consistency and Asymptotic Normality of OLS Estimators

3.1 Strong Consistency

Proposition 1. The OLS estimator of β_0 of the GLM of the previous section, under (A.1) and (A.2), obeys

$$\hat{\beta} = (X'X)^{-1}X'y, \quad \hat{\beta} \stackrel{\text{a.c}}{\to} \beta_0. \tag{3}$$

Proof: By Corollary 3, Ch. 3 in Dhrymes (1989), it will suffice to show that $(X'X/T)$ and $(X'u/T)$, converge to their respective limits a.c. (almost certainly). This is so since $\hat{\beta}$ is a continuous functions of the entities above, and the limit of the expectation of the matrix is nonsingular by assumption (A.1). By the same result it will suffice to consider, for arbitrary conformable γ,

$$\frac{1}{T}\gamma'X'X\gamma = \frac{1}{T}\sum_{t=1}^{T}(x_{t\cdot}\gamma)^2 = \frac{S_T}{T}. \tag{4}$$

[1] The assumption that $E(x_{t\cdot}'u_t) = 0$, for all t, is not quite in character with the GLM where, almost invariably, we are interested in $E(y_t \,|\, x_{t\cdot})$, as e.g., in forecasting. The assumption above does not rule out the case $E(u_t \,|\, x_{t\cdot}) = h(x_{t\cdot})$; when this is so, we have $E(y_t \,|\, x_{t\cdot}) = x_{t\cdot}\beta_0 + h(x_{t\cdot})$. But this is not what we are interested in, nor is it what is commonly forecast. Thus it would be more in character, if the assumption of "uncorrelatedness" were framed as

$$\begin{aligned} E(u_t \,|\, x_{t\cdot}) &= 0, \\ \mathrm{Cov}(u_t u_{t'} \,|\, x_{t\cdot}, x_{t'\cdot}) &= \sigma_t^2, \ \text{ if } t = t' \\ &= 0, \ \text{ otherwise.} \end{aligned}$$

From a technical point of view, White's assumption makes the derivation of the limiting distribution of the OLS estimator considerably simpler

By Proposition 22, Ch. 3 in Dhrymes (1989), (especially Remark 11) we have that the Kolmogorov criterion reduces to

$$c_1 \sum_{t=1}^{T} \frac{t^\alpha}{t^2} = c_1 \sum_{t=1}^{T} \frac{1}{t^{2-\alpha}} < \infty,$$

since $2-\alpha > 1$. Hence, by Proposition 22, $(X'X/T) \stackrel{\text{a.c.}}{\to} M_{xx}$. An entirely similar argument will show that $(X'u/T) \stackrel{\text{a.c.}}{\to} 0$.

q.e.d.

Corollary 1. If we strengthen the assumption (A.1) by requiring

$$E(|z_{ti}z_{tj}z_{tk}z_{tr}|) \leq ct^\alpha \quad \alpha \in [0, \frac{1}{2}),$$

the result of the Proposition holds for sequences $\{z_{t\cdot} = (x_{t\cdot}, u_t) : t \geq 1\}$, which are **merely** uncorrelated nonidentically distributed.

Proof: See Proposition 26, Ch. 3 in Dhrymes (1989).

3.2 Asymptotic Normality

To show the asymptotic normality of the estimator $\hat{\beta}$ we assume, in addition to (A.1) and (A.2),

(A.3). For every real conformable vector, γ, the sequence $\{\zeta_t = x_{t\cdot}\gamma u_t : t \geq 1\}$ obeys

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \int_{|\zeta| > \frac{\sqrt{T}}{r}} \zeta^2 dF_t(\zeta) \; = \; 0,$$

for abritrary integer r, where F_t is the distribution function of ζ_t.

(A.4). The matrix Φ is positive definite (nonsingular), where

$$\Phi = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E(u_t^2 x_{t\cdot}' x_{t\cdot}).$$

Remark 1. Our assumptions (A.1), (A.2) are considerably weaker than the assumptions A.2 through A.4 in White, which involve, *inter alia*, the

uniform boundedness of $E|\epsilon_t^2 x_{ti} x_{tj}|^{1+\delta}$, i.e., the uniform boundedness of $(4+4\delta)^{\text{th}}$ moments.

Moreover, this (White's) assumption ensures that the sufficient conditions for the Liapounov CLT hold, i.e., they imply the Liapounov conditions. Our assumption (A.3) is also weaker than White's assumptions, since the Liapounov conditions imply our assumption (A.3), but our assumption does not imply the Liapounov conditions. Finally, our assumption A.4 is precisely assumption A.3 (b) in White.

In view of the preceding we obtain

Proposition 2. Consider the model in Eq. (1.1), together with assumptions A.1 through A.4. Then

$$\sqrt{T}(\hat{\beta} - \beta_0) \stackrel{\mathrm{d}}{\to} N(0, M_{xx}^{-1} \Phi M_{xx}^{-1}).$$

Proof: We note that

$$\sqrt{T}(\hat{\beta} - \beta_0) = \left(\frac{X'X}{T}\right)^{-1} \frac{X'u}{\sqrt{T}}.$$

By Proposition 26, Corollary 6 in Ch. 4, Dhrymes (1989)

$$\sqrt{T}(\hat{\beta} - \beta_0) \sim M_{xx}^{-1} \xi,$$

where $(X'u/\sqrt{T}) \stackrel{\mathrm{d}}{\to} \xi$. By Proposition 34 in Ch. 4, Dhrymes (1989)

$$\frac{X'u}{\sqrt{T}} \stackrel{\mathrm{d}}{\to} \xi, \quad \text{if and only if } \frac{\gamma' X' u}{\sqrt{T}} \stackrel{\mathrm{d}}{\to} \gamma' \xi,$$

where γ is an arbitrary (real) conformable vector. By assumption (A.3) the sequence $\{\zeta_t = x_{t\cdot}\gamma u_t : t \geq 1\}$ obeys the **Lindeberg** condition; consequently, by Proposition 45 in Ch. 4, Dhrymes (1989)

$$\frac{\gamma' X' u}{\sqrt{T}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \zeta_t \stackrel{\mathrm{d}}{\to} N(0, \gamma' \Phi \gamma).$$

Thus, we conclude $\sqrt{T}(\hat{\beta} - \beta_0) \sim N(0, M_{xx}^{-1} \Phi M_{xx}^{-1})$.

q.e.d.

The next problem is to produce (at least) a consistent estimator of for Φ.

4 Convergence of the Covariance Estimator

The proposed estimator for Φ is

$$\hat{\Phi} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 x_{t\cdot}' x_{t\cdot}, \tag{5}$$

where $\hat{u}_t = y_t - x_{t\cdot}\hat{\beta} = u_t - x_{t\cdot}(\hat{\beta} - \beta_0)$. Thus,

$$\hat{u}_t^2 = u_t^2 - 2x_{t\cdot}(\hat{\beta} - \beta_0)u_t + (\hat{\beta} - \beta_0)'(x_{t\cdot}'x_{t\cdot})(\hat{\beta} - \beta_0) = u_t^2 - 2\phi_{tT1} + \phi_{tT2}, \tag{6}$$

and, in the obvious notation, $\hat{\Phi} = \hat{\Phi}_1 - 2\hat{\Phi}_2 + \hat{\Phi}_3$. We have,

Proposition 3. Under assumptions A.1 through A.4, and in addition,[2]

$$\text{(A.5)} \quad \mathrm{Var}[(x_{t\cdot}\gamma)^2 u_t^2] \leq c\, t^{\alpha}, \quad t \geq 1, \quad \text{for} \quad \alpha \in [0, 1),$$

$$\hat{\Phi} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 x_{t\cdot}' x_{t\cdot} \stackrel{\text{a.c.}}{\to} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(u^2 x_{t\cdot}' x_{t\cdot}).$$

Proof: We note that

$$\hat{\Phi}_3 \leq T[(\hat{\beta} - \beta_0)'(\hat{\beta} - \beta_0)] \left(\frac{\max_{t\leq T} x_{t\cdot}' x_{t\cdot}}{T}\right) \frac{1}{T}\sum_{t=1}^{T} x_{t\cdot}' x_{t\cdot},$$

and by Propositions 1, 2, we conclude that the right member above converges a.c. to the zero matrix. We further note that

$$\hat{\Phi}_2 \leq \left(T[(\hat{\beta} - \beta_0)'(\hat{\beta} - \beta_0)]\right)^{(1/2)} \left(\frac{\max_{t\leq T} x_{t\cdot} x_{t\cdot}'}{T}\right)^{(1/2)} \hat{\Phi}_1.$$

The proof of the proposition will be concluded if we show that $\hat{\Phi}_1$ converges a.c. to its limit, Φ. But this is ensured, in view of assumption (A.5), by Kolmogorov's criterion as cited above.

q.e.d.

This concludes the derivation of results in White (1980), under assumptions equal to or less stringent than those employed by that author.

[2] The corresponding assumption in White is that $E|(x_{t\cdot}\gamma)^2 u_t^2|^{1+\delta}$ is **uniformly bounded**. In our case we require the existence of the second moment but allow the latter to grow at the rate t^{α}, so that we have basically similar kinds of restrictive assumptions. Which set one chooses to work with is a matter of taste. In this paper one can take White's assumptions and prove the same result by relying on a certain SLLN, see for example Chung (1968), Corollary (ii.) p. 119, as also cited by White.

5 Extension to Dynamic Models

In this section we reexamine the model in Eq. (1.1) with two important changes: first, among the explanatory variables we take y_{t-1}, and second, we adopt the framework of footnote 1. Thus, the model is

$$y_t = x_{t\cdot}\beta_0 + y_{t-1}\lambda_0 + u_t, \; t = 1, 2 \ldots, T, \tag{7}$$

so that it is the same model as in Eq. (1.1), except that it is dynamic. Since here we need greater detail in our exposition we adopt a more formal framework. Let $(\Omega, \mathcal{A}, \mathcal{P})$ be the probability space and let $z_{t\cdot} = (x_{t\cdot}, y_{t-1})$.

Assumptions made in previous sections on the vector $x_{t\cdot}$, remain in effect, except as modified. Moreover, put

$$\mathcal{C}_t = \sigma(x_{s\cdot} : s \leq t), \quad \mathcal{D}_t = \sigma(u_s : s \leq t), \quad t \geq 1, \quad \mathcal{C}_0 = (\emptyset, \Omega), \quad \mathcal{D}_0 = (\emptyset, \Omega).$$

define

$$\mathcal{A}_t = \sigma(\mathcal{C}_{t+1} \cup \mathcal{D}_t), \quad t \geq 1, \quad \mathcal{A}_0 = \sigma[\mathcal{C}_1 \cup \mathcal{D}_0], \tag{8}$$

and note that $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \mathcal{A}_2 \ldots$.

In this segment of the paper we have the following framework:

(B.1) The sequence $\{(x_{t\cdot}, u_t) : t \geq 1\}$ is one of independent (but not necessarily identically distributed) random variables, such that

$$M_{xx} = \lim_{T\to\infty} \frac{1}{T} E(X'X), \quad E(u_t|\mathcal{C}_t) = 0, \quad E(u_t^2|\mathcal{C}_t) = \sigma_t^2,$$

for all t, where M_{xx} is a well defined nonsingular (positive definite) matrix;

(B.2) the sequence $\{x_{t\cdot} : t \geq 1\}$ possesses **fourth order moments obeying**

$$E(|x_{ti}x_{tj}x_{tk}x_{tr}|) \leq ct^{\alpha} \quad \alpha \in [0, 1);$$

(B.3) the sequence $\{u_t^2 : t \geq 1\}$ is **uniformly integrable**;

(B.4) $E[(u_t^2 - \sigma_t^2)^2 v_t^4] \leq ct^{\alpha}$, $\alpha \in [0, \frac{1}{2})$, $v_t = \sum_{t=0}^{t-2} \lambda_0^j u_{t-1-j}$;

(B.5) the matrix Φ is positive definite (nonsingular), where

$$\Phi = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sigma_t^2 E(z_{t\cdot}' z_{t\cdot}), \quad \text{and} \quad \lambda_0 \in (-1, 1).$$

To investigate the problem and at the same time maintain maximal correspondence between the argument here and that in the earlier part of the paper, we adopt the notation $y_t = z_{t\cdot}\theta_0 + u_t$, where, evidently, $z_{t\cdot} = (x_{t\cdot}, y_{t-1})$, $\theta_0 = (\beta_0', \lambda_0)'$.

5.1 Consistency

Proposition 4. The OLS estimator of θ_0 of the GLM of the previous section, under (B.1) and (B.2), obeys

$$\hat{\theta} = (Z'Z)^{-1}Z'y, \quad \hat{\theta} \overset{\text{a.c}}{\to} \theta_0, \tag{9}$$

where $Z = (X, y_{-1})$.

Proof: By the reasons given in the proof of Proposition 1, it will suffice to show that $(Z'Z/T)$ and $(Z'u/T)$, converge to their respective limits a.c. . Since

$$\frac{1}{T}Z'Z = \frac{1}{T}\begin{bmatrix} X'X & X'y_{-1} \\ y_{-1}'X & y_{-1}'y_{-1} \end{bmatrix},$$

we need only be concerned about $X'y_{-1}$ and $y_{-1}'y_{-1}$, in view of Proposition 1. Define the matrix

$$B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda & 1 & 0 & \dots & 0 \\ \lambda^2 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda^{T-1} & \lambda^{T-2} & \lambda^{T-3} & \dots & 1 \end{bmatrix}$$

and note that, solving the difference equation of the model we have

$$y = BX\beta_0 + Bu = \bar{y} + v. \tag{10}$$

Since

$$\frac{1}{T}\bar{y}'\bar{y} \leq \frac{c}{(1-\lambda_0)^2}\frac{1}{T}\beta_0'X'X\beta_0, \quad \frac{1}{T}v'v \leq \frac{c}{(1-\lambda_0)^2}\frac{1}{T}u_{-1}'u_{-1},$$

etc., the existence of the least squares estimator is proved, i.e.

$$\frac{Z'Z}{T} \overset{\text{a.c.}}{\to} \lim_{T\to\infty}\frac{E(Z'Z)}{T} > 0;$$

as for its convergence a.c. we note that proving

$$\frac{Z'u}{T} = \frac{(X, \bar{y}_{-1})'u}{T} + \frac{(0, v)'u}{T} \stackrel{\text{a.c.}}{\to} 0$$

is exactly the same problem dealt with in Proposition 1, except for the term $v'u/T$, which has mean zero. Since

$$\psi_T = \left(\frac{v'u}{T}\right) = \frac{1}{T}\sum_{t=2}^{T} v_t u_t$$

is a sequence of zero mean uncorrelated random variables whose variance is certainly bounded by (B.4), it follows that ψ_T converges a.c. to zero, and the strong consistency of the least squares estimator is proved.

q.e.d.

5.2 Asymptotic Normality

We note that

$$\sqrt{T}(\hat{\theta} - \theta_0) \sim M_{zz}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_{t\cdot}' u_t, \tag{11}$$

and it will suffice to show that, for arbitrary (confomrable) vector γ the sequence

$$\mathcal{Z}_T = \sum_{t=1}^{T} \zeta_{tT}, \quad \zeta_{tT} = \frac{1}{\sqrt{T}} z_{t\cdot} \gamma u_t, \tag{12}$$

converges in distribution. We have

Propostion 5. Under assumptions (B.1) through (B.4),

$$\sqrt{T}(\hat{\theta} - \theta_0) \sim N(0, M_{zz}^{-1} \Phi M_{zz}^{-1}).$$

Proof: By the discussion just above, it will suffice to show that the sequence $\mathcal{Z}_T$ converges in distribution. We note that $\mathcal{A}_t$, or modified to take account of division by $\sqrt{T}$ as $\mathcal{A}_{tT}$, $t \le T$, $T \ge 1$ is a stochastic basis, i.e. $\mathcal{A}_{tT} \subset \mathcal{A}_{t+1,T} \subset \mathcal{A}_{t+1,T+1}$, etc. Moreover, $(\zeta_t, \mathcal{A}_t)$ is a **martingale difference** stochastic sequence, since

$$E(|\zeta_t|) < \infty, \quad \text{and} \quad E(\zeta_t | \mathcal{A}_{t-1}) = 0.$$

Moreover, the sequence $\mathcal{Z}_T$ satisfies a **Lindeberg condition**, as follows. Let, for arbitrary integer r,

$$A_{tT} = \{\omega : |\zeta_{tT}| > \frac{1}{r}\}, \quad B_{tT} = \{\omega : |u_t| > \frac{1}{\alpha_T r}\},$$
$$\alpha_T = \left(\frac{\max_{t\leq T} \gamma' z'_{t\cdot} z_{t\cdot} \gamma}{T}\right)^{1/2},$$

and note that $A_{tT} \subset B_{tT}$, $\alpha_T \stackrel{\text{a.c.}}{\to} 0$ by Propositions 1 and 4. Letting $I(A)$ denote the indicator function of the set A, we have

$$\begin{aligned}
\mathcal{L}_T &= \sum_{t=1}^{T} E[\zeta_t^2 I(A_{tT})|\mathcal{A}_{t-1,T}] \\
&\leq \frac{1}{T}\sum_{t=1}^{T} E[\zeta_t^2 I(B_{tT})|\mathcal{A}_{t-1}] \\
&\leq \frac{1}{T}\sum_{t=1}^{T} (z_{t\cdot}\gamma)^2 E[u_t^2 I(B_{tT})|\mathcal{A}_{t-1}] \\
&\leq \max_{t\leq T} E[u_t^2 I(B_{tT})|\mathcal{A}_{t-1}] \, \frac{1}{T}\sum_{t=1}^{T}(z_{t\cdot}\gamma)^2 \stackrel{\text{a.c.}}{\to} 0.
\end{aligned} \tag{13}$$

The convergence a.c. to zero is valid by (B.3) and Proposition 4. Next we observe that

$$\frac{1}{T}\sum_{t=1}^{T} E(\zeta_t^2|\mathcal{A}_{t-1}) = \frac{1}{T}\sum_{t=1}^{T} \sigma_t^2 (z_{t\cdot}\gamma)^2 \stackrel{\text{a.c.}}{\to} \gamma'\Phi\gamma;$$

consequently, by Proposition 21 (i.) in Ch. 5, Dhrymes (1989) we conclude

$$\mathcal{Z}_T \stackrel{\text{a.c}}{\to} N(0, \gamma'\Phi\gamma), \quad \text{and} \quad \sqrt{T}(\hat{\theta} - \theta_0) \sim N(0, M_{zz}^{-1}\Phi M_{zz}^{-1}).$$

q.e.d.

5.3 Convergence of the Covariance Estimator in the Dynamic Model

This is by far the most complex issue in the extension. Since we have already shown that $(Z'u/T) \stackrel{\text{a.c.}}{\to} 0$, the arguments given in Proposition 3 apply to the three components of the estimator

$$\hat{\Phi} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 z'_{t\cdot} z_{t\cdot} = \hat{\Phi}_1 - 2\hat{\Phi}_2 + \hat{\Phi}_3, \tag{14}$$

the three components above corresponding to the three components of $\hat{u}_t^2$ in the equation below,

$$\hat{u}_t^2 = u_t^2 - 2z_{t\cdot}(\hat{\theta} - \theta_0)u_t + (\hat{\theta} - \theta_0)'(z_{t\cdot}'z_{t\cdot})(\hat{\theta} - \theta_0) = u_t^2 - 2\phi_{tT1} + \phi_{tT2}. \quad (15)$$

By an argument identical to the one we made in the previous section, $\hat{\Phi}_3$ converges a.c. to zero and $\hat{\Phi}_2$ converges a.c. to zero if $\hat{\Phi}_1$ converges. Thus, the (strong) consistency of the estimator of the covariance matrix in Eq. (14) will be proved, if we prove that $\hat{\Phi}_1$ converges a.c. to its limit.

Proposition 5. Under assumptions (B.1) through (B.5),

$$\hat{\Phi} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 z_{t\cdot}' z_{t\cdot} \overset{\text{a.c.}}{\to} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sigma_t^2 E(z_{t\cdot}' z_{t\cdot}).$$

Proof: We first recall that $z_{t\cdot} = \bar{z}_{t\cdot} + (0, v_t)$, and consequently,

$$u_t^2 z_{t\cdot}' z_{t\cdot} = u_t^2 \bar{z}_{t\cdot}' \bar{z}_{t\cdot} + u_t^2 \bar{z}_{t\cdot}'(0, v_t) + u_t^2 (0, v_t)' \bar{z}_{t\cdot} + \begin{pmatrix} 0 & 0 \\ 0 & v_t^2 \end{pmatrix} u_t^2.$$

The first term evidently converges to its limit, by the arguments of Proposition 3; since the mean of the second and third terms is zero these terms will converge to zero a.c., if the last term is shown to converge a.c. To do so, note that we are dealing only with the scalar sequence

$$S_T = \sum_{t=2}^{T} v_t^2 u_t^2, \quad E(S_T) = \sum_{t=2}^{T} \sigma_t^2 \beta(t-2), \quad \beta(t-2) = \sum_{j=0}^{t-2} \lambda_0^{2j} \sigma_{t-1-j}^2, \quad (16)$$

and that $S_T - E(S_T)$ is a sequence of zero mean **dependent variables**. Unfortunately, there is no general theorem to which we can appeal for its a.c. convergence; thus, we resort to first principles. Write

$$S_T = \sum_{t=2}^{T} \left[v^2(u_t^2 - \sigma_t^2) + \sigma_t^2[v_t^2 - \beta(t-2)] + \sigma_t^2 \beta(t-2) \right] = S_T^{(1)} + S_T^{(2)} + S_T^{(3)}.$$

Since $S_T^{(1)}$ is a sum of zero mean **uncorrelated random variables**, by (B.3) and Proposition 26, in Chapter 3 Dhrymes (1989), $(S^{(1)}/T) \overset{\text{a.c.}}{\to} 0$. Next note that, it being understood that $u_{-s} = 0$, for $s \geq 0$,

$$S_T^{(2)} \leq \sup_{t\geq 1} \sigma_t^2 \sum_{j=0}^{\infty} \lambda_0^{2j} \left(\sum_{t=j+2}^{T} [u_{t-1-j}^2 - \sigma_{t-1-j}^2] \right) + 2 \sum_{j'<j} \lambda_0^{j+j'} \left(\sum_{t=j+2}^{T} u_{t-1-j} u_{t-1-j'} \right).$$

In both terms we are dealing with sums of uncorrelated random variables; thus, the sums over the index t, when divided by T, converge a.c. to zero. Since the sums over j and/or j' converge absolutely by (B.5), it follows again by (B.4) and Proposition 26, in Chapter 3 Dhrymes (1989) that $(S^{(2)}/T) \stackrel{\text{a.c.}}{\to} 0$. Hence,

$$\frac{S_T}{T} \stackrel{\text{a.c.}}{\to} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sigma_t^2 \left(\sum_{j=0}^{t-2} \lambda_0^{2j} \sigma_{t-1-j}^2\right),$$

which shows that $\hat{\Phi} \stackrel{\text{a.c.}}{\to} \Phi$.

q.e.d.

Remark 2. The results herein may be easily extended to the case where the model errors are a stable autoregressive process, in the following sense. If the model is transformed so that it contains exogenous, lagged exogenous, and lagged dependent explanatory variables, the error process would be one of independent nonidentically distributed random variables, and hence amenable to the analysis of this section. If the model **initially** contains lagged dependent variables, the same remark would apply **if one is not interested in identifying the autoregressive parameters in the error process**. If one is, nonlinear methods would have to be employed and, while one might conjecture that the same results would hold, the problem is not so transparent *ab initio.*

If all one is willing to specify is that the error process is a **martingale difference**, i.e., if all one is willing to specify is that $E|u_t| < \infty$ and $E(u_t|\mathcal{A}_{t-1}) = 0$, one is forced to rely only on **martingale convergence** theorems, as in Nicholls and Pagan. The results one obtains, however, limit the nature of the explanatory variables one may employ. In Nicholls and Pagan, one has a stable AR model with martingale difference errors. This immediately implies that the dependent variable process is also a martingale difference. This simplifies the problem a great deal, but would not admit exogenous variables unless they were **also specified to be a zero mean martingale difference process**. If one did that one would have no difficulty in extending the Nicholls and Pagan results to the case where there are exogenous variables, lagged as well as contemporaneous. It would appear, however, that such a model would not be very useful in empirical applications. At least the Nicholls and Pagan result provides an interesting twist to the VAR model, although it is quite contrary to the stochastic orientation embodied in the latter.

BIBLIOGRPHY

Dhrymes, P. J. (1989) *Topics In Advanced Econometrics: Probability Foundations*, New York, Springer-Verlang

Eicker, F. (1967) "Limit Theorems for Regressions with Unequal and Dependent Errors", in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, Berkeley, University of California Press.

————————————(1963) "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions",*Annals of Mathematic Statistics*, 34, pp. 447-56.

Nicholls, D.F. and A.R. Pagan (1983) "Heteroskedasticity in Models with Lagged Dependent Variables", *Econometrica*, 51, pp.1233-42.

Rao, C. R. (1973) *Linear Statistical Inference and its Applications*, Neww York, Wiley.

————————————-(1970) "Estimation of Heteroskedastic Variances in Linear Models", *Journal of the American Statistical Association*, 65, pp. 161-72.

White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, pp.817-38.