

DIAGNOSTIC TESTS FOR THE PARTIAL
ADJUSTMENT AND ADAPTIVE EXPECTATIONS MODELS

H.E. Doran*

No. 17 - February 1985

*Senior Lecturer, Department of Econometrics, University of New England,
Armidale, Australia.

ISSN 0157-0188
ISBN 0 85834 5870

1. Introduction

Over the last few years, theoretical econometricians have shown great interest in the question of model validation. This has arisen because so many models which appear to be satisfactorily estimated, have in fact performed rather poorly. Pagan (1984) comments on the 'weakness of research strategy' in the following terms: 'Almost no attempt was made to discover weaknesses in the model selected, or at least to present them to an observer. Once a model agreeing with *a priori* conceptions of 'signs' and 'significance' was obtained, research terminated.'

Two broad approaches to model validation have appeared. The first of these, initiated by early work of Cox (1961, 1962) and Atkinson (1970), provided a framework for testing 'non-nested' or 'separate' models, that is, models of which 'neither may be obtained from the other by appropriate parametric restrictions' (Fisher and McAleer (1981)). We cite just two examples of this approach - Pesaran (1974) and Davidson and MacKinnon (1981) - together with a recent survey by McAleer (1984). This approach is not followed in this article.

The second group of procedures, sometimes called 'classical' methods, are used when the model of interest can be 'nested' or 'embedded' in an *economically meaningful* comprehensive model. The model of interest can be obtained from the comprehensive model by applying parameter restrictions. Diagnostic tests may then be obtained by using the usual F- and t-tests, or the asymptotically equivalent Likelihood Ratio (LR), Wald (W) and Lagrange Multiplier (LM) principles. Particular interest has recently centred on the derivation of LM diagnostic tests - see, for example, Breusch (1978), Breusch and Pagan (1980) and Engle (1982).

The purpose of this paper is twofold. Firstly, it is intended to provide diagnostic tests for two models which are still often used in

applied supply response and inventory investment studies. These models are the Partial Adjustment (PA) model, originally proposed by Nerlove (1958), and the Adaptive Expectations (AE) model, suggested originally by Cagan (1956). Secondly, the paper provides further examples for applied workers on how the LR, W and LM principles can be used to derive practical diagnostic tests of maintained models.

2. The Partial Adjustment-Adaptive Expectations (PAAE) Model

A description of this model and its consistent estimation can be found in Doran and Griffiths (1978). The main features are summarized here for convenience.

Let Y_t^* be the desired level of a dependent variable (e.g., supply or inventory) and X_t^* the expected value of an exogenous variable (e.g., price or sales). Both Y_t^* and X_t^* are unobservable, and it is postulated that they are linearly related. That is,

$$Y_t^* = \alpha + \beta X_t^*. \quad (2.1)$$

The mechanisms which generate Y_t^* and X_t^* are given by the partial adjustment and adaptive expectations hypotheses, respectively. That is,

$$Y_t - Y_{t-1} = (1-\gamma)(Y_t^* - Y_{t-1}) + v_t \quad (2.2)$$

and

$$X_t^* - X_{t-1}^* = (1-\delta)(X_{t-1} - X_{t-1}^*), \quad (2.3)$$

where $0 \leq \gamma < 1$, $0 \leq \delta < 1$ ⁽¹⁾ and v_t is i.i.d. $N(0, \sigma^2)$. Using (2.2) and (2.3) to eliminate the unobservable variables in (2.1), we obtain

(1) We have replaced γ and δ in the usual formulation by $1-\gamma$ and $1-\delta$. This makes for neater algebra subsequently.

3.

$$Y_t = \alpha(1-\gamma)(1-\delta) + \beta(1-\gamma)(1-\delta)X_{t-1} + (\gamma+\delta)Y_{t-1} - \gamma\delta Y_{t-2} + u_t, \quad (2.4)$$

where

$$u_t = v_t - \delta v_{t-1}. \quad (2.5)$$

The model (2.4) and (2.5) is a comprehensive model of which the PA and AE models are special cases.

The PA model occurs when $\delta=0$. That is, 'naive' expectations, given by $X_t^*=X_{t-1}$, hold. There is only one adjustment process, given by (2.2), which accounts for habit persistence or technological rigidities. The PA model is thus given by

$$Y_t = \alpha(1-\gamma) + \beta(1-\gamma)X_{t-1} + \gamma Y_{t-1} + v_t, \quad (2.6)$$

where it should be noted that the disturbance is i.i.d. $N(0, \sigma^2)$.

On the other hand, the AE model occurs when $\gamma=0$. Here it is assumed that actual Y_t equals desired Y_t , apart from an uncontrollable error v_t which accounts, for example, for the effects of weather. The only unobservable variable is X_t^* , expected price or sales, which is formed from corresponding actual values by (2.3). The AE model is given by

$$Y_t = \alpha(1-\delta) + \beta(1-\delta)X_{t-1} + \delta Y_{t-1} + u_t, \quad (2.7)$$

where u_t is given by (2.5). Here we note that the disturbance u_t has a first-order moving average structure.

From this discussion it can be seen that diagnostic tests of the PA and AE models may be constructed by testing the hypotheses $\delta=0$ (for PA model) and $\gamma=0$ (for AE model) on the comprehensive PAAE model (2.4) and (2.5). If the

PAAE model is reparametrized, in the form

$$Y_t = \beta_1 + \beta_2 X_{t-1} + \beta_3 Y_{t-1} + \beta_4 Y_{t-2} + u_t, \quad (2.8)$$

where

$$\beta_4 = -\delta(\beta_3 - \delta), \quad (2.9)$$

and

$$u_t = v_t - \delta v_{t-1}, \quad (2.10)$$

the two main features of this model are clearly displayed. They are

- (i) there is a moving average disturbance term; and
- (ii) there is a non-linear restriction among the β coefficients which involves the parameter δ of the MA disturbance.

These two characteristics should be taken into account in deriving efficient tests.

If we let $\theta = (\alpha, \beta, \gamma, \delta, \sigma^2)'$ and $v = (v_1, v_2, \dots, v_N)'$ then because v_t is i.i.d. $N(0, \sigma^2)$, the log-likelihood function $L(\theta)$ is given by

$$L(\theta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} v'v, \quad (2.11)$$

and the maximum-likelihood (ML) estimate of θ is found when $v'v$ is minimised and $\hat{\sigma}^2 = N^{-1} \min(v'v)$.

Dhrymes (1971, p.98) has shown how the ML estimates can be computed.

The PAAE model can be written in the form

$$Y_t = \alpha(1-\gamma) + \gamma Y_{t-1} + \beta(1-\gamma)X_t(\delta) + \beta(1-\gamma)(1-\delta)\delta^{t-1}\bar{X}_0^* + v_t, \quad (2.12)$$

where

$$\begin{aligned}\bar{X}_t(\delta) &= 0 \quad \text{when } t=1 \\ &= (1-\delta) \sum_{i=0}^{t-2} \delta^i X_{t-i-1} \quad \text{when } t>1,\end{aligned}\tag{2.13}$$

and

$$\bar{X}_0^* = \sum_{i=0}^{\infty} \delta^i X_{-i}.\tag{2.14}$$

Putting equation (2.12) in the usual vector form, we have

$$Y = X^* \beta^* + v,\tag{2.15}$$

where

$$Y = [Y_1, Y_2, \dots, Y_N]'$$

$$X^* = \begin{bmatrix} 1 & Y_0 & 0 & \delta^0 \\ 1 & Y_1 & (1-\delta)\delta^0 X_1 & \delta^1 \\ . & . & (1-\delta)(\delta^0 X_2 + \delta^1 X_1) & \delta^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & Y_{N-1} & (1-\delta)(\delta^0 X_{N-1} + \delta^1 X_{N-2} + \dots + \delta^{N-2} X_1) & \delta^{N-1} \end{bmatrix}$$

and

$$\beta^* = (\alpha(1-\gamma), \gamma, \beta(1-\gamma), \beta(1-\gamma)(1-\delta)\bar{X}_0^*)'$$

If δ were known, the observation matrix X^* could be constructed. The maximum-likelihood estimates are obtained by applying OLS to (2.15) for different values of δ in the range $0 \leq \delta < 1$, and choosing the estimates that yield the minimum residual sum of squares (RSS).

At this point we introduce notation which will be used throughout the remainder of the article. The symbols ' $\hat{\cdot}$ ' and ' $\tilde{\cdot}$ ' will denote quantities evaluated at the unrestricted ML estimates and those evaluated at estimates restricted by some hypothesis, respectively. Thus, for example, $\hat{\theta}$ would denote unrestricted ML estimates of the parameter θ , and $\hat{v}$ and $\partial \hat{v} / \partial \theta$ would be the residual vector and the differential $\partial v / \partial \theta$ when $\theta = \hat{\theta}$.

3. Principles of Hypothesis Testing

There are three principles which are in common use for testing nested hypotheses. These three principles are equivalent in the sense that asymptotically they are most powerful and, under the null hypothesis, the statistics which emerge have the same distribution. The choice of which principle to use in a given context will be made largely on questions of mathematical tractability and computational convenience.

Let us suppose that the model under consideration and sample data can be described by a log-likelihood function $L(\theta)$, where θ is a p -vector of unknown parameters. Furthermore, the vector θ is to be partitioned as

$$\theta = [\theta_1', \theta_2']$$

in order to test a hypothesis of the form $\theta_1 = \theta_1^*$, where θ^* is a known vector of dimension p_1 .

(a) *Likelihood-ratio principle*

Let $L(\hat{\theta})$ and $L(\tilde{\theta})$ be the unrestricted maximum value of the likelihood function and the maximum value under the null hypothesis $\theta_1 = \theta_1^*$, respectively. Then, if the null hypothesis is true

$$LR \equiv -2[\ln L(\tilde{\theta}) - \ln L(\hat{\theta})] \overset{a}{\sim} \chi_{p_1}^2, \quad (3.1)$$

where the symbol ' $\overset{a}{\sim}$ ' means 'is asymptotically distributed as'. The null hypothesis is rejected if LR is 'too large' relative to the $\chi_{p_1}^2$ distribution.

Fundamental to this method is that both the unrestricted and restricted maximums of the likelihood function have to be found.

(b) *Wald principle*

A property of the maximum-likelihood estimator $\hat{\theta}$ is that

$$N^{\frac{1}{2}}(\hat{\theta} - \theta) \overset{a}{\rightsquigarrow} N(0, I^{-1}), \quad (3.2)$$

where I is called the asymptotic information matrix, and is defined by

$$I = -\text{plim } N^{-1} \frac{\partial^2 L}{\partial \theta \partial \theta'}, \quad (3.3)$$

If I is partitioned conformably with the partitioning of θ , as

$$I = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix},$$

where I_{11} is $p_1 \times p_1$, I_{12} is $p_1 \times p_2$, etc. it follows from (3.2), using well-known results for the inverse of a partitioned matrix (see Graybill (1969), p.165) that under the null hypothesis

$$N^{\frac{1}{2}}(\hat{\theta}_1 - \theta_1^*) \overset{a}{\rightsquigarrow} N(0, (I_{11} - I_{12} I_{22}^{-1} I_{21})^{-1}). \quad (3.4)$$

A scalar statistic is then given by

$$W \equiv N(\hat{\theta}_1 - \theta_1^*)' (\hat{I}_{11} - \hat{I}_{12} \hat{I}_{22}^{-1} \hat{I}_{21}) (\hat{\theta}_1 - \theta_1^*) \quad (3.5)$$

and if the null is true,

$$W \overset{a}{\rightsquigarrow} \chi_{p_1}^2 \quad (3.6)$$

In contrast to the likelihood-ratio principle, application of a Wald test requires only *unrestricted* maximization of the likelihood function.

However, the information matrix has to be obtained and evaluated at $\theta = \hat{\theta}$.

The Wald principle may be regarded as a 'natural' hypothesis testing procedure in that it focuses on the difference between the estimated value $\hat{\theta}_1$ and the hypothesised value θ_1^* of the parameter θ_1 . A large difference, and hence a large value for W , naturally leads to the rejection of the null hypothesis.

(c) *Lagrange multiplier principle*

This principle of hypothesis testing focuses on the rate of change of the likelihood function $\partial L / \partial \theta$, often called the 'efficient score'. It can be shown that the score statistic has an asymptotic distribution given by

$$N^{-\frac{1}{2}} \frac{\partial L}{\partial \theta} \overset{d}{\sim} N(0, I). \quad (3.7)$$

The restricted maximum-likelihood estimator, subject to the hypothesis $\theta_1 = \theta_1^*$, satisfies the vector equation $\frac{\partial L}{\partial \theta_2} = 0$.

Thus the distribution of $\frac{\partial L}{\partial \theta_1}$, under the restriction may be regarded as the distribution of $\frac{\partial L}{\partial \theta_1}$ conditional on $\frac{\partial L}{\partial \theta_2} = 0$. From standard theory of the multivariate normal distribution (see, for example, Graybill (1961, p.62-64)) we have that under the null hypothesis

$$N^{-\frac{1}{2}} \frac{\partial L}{\partial \theta_1} \overset{d}{\sim} N(0, I_{11} - I_{12} I_{22}^{-1} I_{21}). \quad (3.8)$$

A scalar statistic is then given by

$$LM \equiv N^{-1} \left(\frac{\partial L}{\partial \theta_1} \right)' \left(\tilde{I}_{11} - \tilde{I}_{12} \tilde{I}_{22}^{-1} \tilde{I}_{21} \right)^{-1} \left(\frac{\partial L}{\partial \theta_1} \right) \overset{d}{\sim} \chi_{p_1}^2. \quad (3.9)$$

In this case both $\frac{\partial L}{\partial \theta_1}$ and I are evaluated at the restricted estimator $\tilde{\theta}$, and so only *restricted* estimation is required to compute the LM statistic.

From the above discussion it is clear that the Wald and Lagrange Multiplier approaches would usually be more convenient than the Likelihood Ratio procedure as only one maximization is necessary. The exception occurs

when the information matrix is difficult to obtain. The choice between the Wald and Lagrange multiplier procedures would depend on whether restricted or unrestricted estimation is easier.

There is one special case which deserves comment. If $p_1=1$, that is, θ_1 is a *scalar*, then the expressions (3.4) and (3.8) can be used directly to form test statistics based on the normal distribution. This has relevance when prior knowledge enables us to specify a one-sided alternative hypothesis, either $\theta_1 > \theta_1^*$ or $\theta_1 < \theta_1^*$. In such cases, a one-sided test based on the normal distribution is more powerful than the usual W and LM tests using the χ_1^2 distribution.

4. Simplifications of the Information Matrix

It has been shown above in (3.5) and (3.9) that both the Wald and the Lagrange Multiplier tests involve the $p_1 \times p_1$ matrix $I_{11} - I_{12} I_{22}^{-1} I_{21}$. There are two common situations which enable this matrix to be simplified.

- (a) It often happens that, when $\theta = (\theta_1', \theta_2', \sigma^2)'$, $\text{plim } N^{-1} \frac{\partial^2 L}{\partial \theta_1 \partial \sigma^2}$ and $\text{plim } N^{-1} \frac{\partial^2 L}{\partial \theta_2 \partial \sigma^2}$

evaluated at either $\theta = \hat{\theta}$ or $\theta = \tilde{\theta}$ are zero. In these cases I is block diagonal, of the form

$$I = -\text{plim } N^{-1} \begin{bmatrix} \frac{\partial^2 L}{\partial \theta_1 \partial \theta_1'} & \frac{\partial^2 L}{\partial \theta_1 \partial \theta_2'} & 0 \\ \frac{\partial^2 L}{\partial \theta_2 \partial \theta_1'} & \frac{\partial^2 L}{\partial \theta_2 \partial \theta_2'} & 0 \\ 0 & 0 & \frac{\partial^2 L}{\partial (\sigma^2)^2} \end{bmatrix}.$$

Then,

$$I_{11} - I_{12} I_{22}^{-1} I_{21} = -\text{plim } N^{-1} \left[\frac{\partial^2 L}{\partial \theta_1 \partial \theta_1'} - \frac{\partial^2 L}{\partial \theta_1 \partial \theta_2'} \left(\frac{\partial^2 L}{\partial \theta_2 \partial \theta_2'} \right)^{-1} \frac{\partial^2 L}{\partial \theta_2 \partial \theta_1'} \right]$$

and differentials with respect to σ^2 do not appear. In such cases the parameter σ^2 may be omitted from the definition of θ .

- (b) From the form of the log-likelihood function (2.11), the information matrix I involves differentiation of the quadratic form $\frac{1}{2} v'v$. By a standard result of vector calculus (see Dhrymes (1978, p.529)),

$$\frac{\partial^2 \left(\frac{1}{2} v'v \right)}{\partial \theta \partial \theta'} = \left(\frac{\partial v}{\partial \theta} \right)' \left(\frac{\partial v}{\partial \theta} \right) + (v' \otimes I) \frac{\partial^2 v}{\partial \theta \partial \theta'}.$$

Often $\text{plim } N^{-1} (v' \otimes I) \frac{\partial^2 v}{\partial \theta \partial \theta'}$ evaluated at $\hat{\theta}$ or $\tilde{\theta}$ is zero in which case, (assuming differentials with respect to σ^2 can be ignored) I has the simpler form

$$I = \text{plim } \frac{N^{-1}}{\sigma^2} \begin{bmatrix} \left(\frac{\partial v}{\partial \theta_1} \right)' \left(\frac{\partial v}{\partial \theta_1} \right) & \left(\frac{\partial v}{\partial \theta_1} \right)' \left(\frac{\partial v}{\partial \theta_2} \right) \\ \left(\frac{\partial v}{\partial \theta_2} \right)' \left(\frac{\partial v}{\partial \theta_1} \right) & \left(\frac{\partial v}{\partial \theta_2} \right)' \left(\frac{\partial v}{\partial \theta_2} \right) \end{bmatrix}.$$

Thus,

$$I_{11} - I_{12} I_{22}^{-1} I_{21} = \text{plim } \frac{N^{-1}}{\sigma^2} \left(\frac{\partial v}{\partial \theta_1} \right)' M_2 \left(\frac{\partial v}{\partial \theta_1} \right), \quad (4.1)$$

where

$$M_2 = 1 - \left(\frac{\partial v}{\partial \theta_2} \right) \left[\left(\frac{\partial v}{\partial \theta_2} \right)' \left(\frac{\partial v}{\partial \theta_2} \right) \right]^{-1} \left(\frac{\partial v}{\partial \theta_2} \right)' . \quad (4.2)$$

In the special case that θ_1 is a *scalar*, we then have that

$$I_{11} - I_{12} I_{22}^{-1} I_{21} = \text{plim } N^{-1} \text{RSS}^* / \sigma^2, \quad (4.3)$$

where RSS^* is the residual sum of squares which would be obtained if

$\frac{\partial v}{\partial \theta_1}$ were regressed on $\frac{\partial v}{\partial \theta_2}$. This result is used often in the following sections.

5. Testing the Adaptive Expectations Model

As we have already seen, the AE model may be validated by constructing a test for the hypothesis $\gamma=0$ in the model (2.4).

We define $\theta = (\theta_1, \theta_2', \sigma^2)'$, where $\theta_1 = \gamma$ and $\theta_2 = (\alpha, \beta, \delta)'$.

(a) *Likelihood-ratio test*

The unrestricted maximum value of the log-likelihood function $L(\hat{\theta})$ is given by

$$L(\hat{\theta}) = - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2} \hat{v}' \hat{v},$$

and as $\hat{\sigma}^2 = N^{-1} \hat{v}' \hat{v}$,

$$L(\hat{\theta}) = - \frac{N}{2} (\ln \hat{\sigma}^2 + \ln(2\pi) + 1).$$

Similarly, the restricted log-likelihood function has a maximum value

$$L(\tilde{\theta}) = - \frac{N}{2} (\ln \tilde{\sigma}^2 + \ln(2\pi) + 1)$$

Thus, by (3.1), the likelihood ratio test statistic is

$$LR = N(\ln \hat{\sigma}^2 - \ln \tilde{\sigma}^2)$$

or

$$LR = N \ln(RSS_0/RSS_1), \quad (5.1)$$

where RSS_0 and RSS_1 are obtained from the restricted and unrestricted models, respectively.

(b) *Wald test*

It is easily verified that, regardless of whether we are dealing with the unrestricted model (2.4), or the restricted model (2.7) which has the regressor Y_{t-2} omitted,

$$\text{plim } N^{-1} \frac{\partial^2 L}{\partial \theta_1 \partial \sigma^2} = \text{plim } N^{-1} \frac{\partial^2 L}{\partial \theta_2 \partial \sigma^2} = 0$$

and

$$\text{plim } N^{-1} \frac{v' \partial^2 v}{\partial \theta_1^2} = \text{plim } N^{-1} (v' \mathbf{0} I) \frac{\partial^2 v}{\partial \theta_2 \partial \theta_2'} = 0.$$

Thus, by (4.3),

$$\hat{I}_{11}^{-1} \hat{I}_{12}^{-1} \hat{I}_{22}^{-1} \hat{I}_{21} = \text{plim } N^{-1} RSS_1^* / \hat{\sigma}^2 \quad (5.2)$$

where RSS_1^* is the residual sum of squares obtained by regressing $\frac{\partial v}{\partial \theta_1}$ on $\frac{\partial v}{\partial \theta_2}$.

Here $\frac{\partial v}{\partial \theta_1} = \alpha j - Y_{-1} + \beta \bar{X}(\delta)$

$$\frac{\partial v}{\partial \theta_2} = -(1-\gamma) \left[j, \bar{X}(\delta), \beta \frac{\partial \bar{X}(\delta)}{\partial \delta} \right]$$

where j is a vector of ones and Y_{-1} contains the observations on Y_{t-1} .

Thus, RSS_1^* is the residual sum of squares which is obtained when

Y_{t-1} is regressed on $\bar{X}_t(\hat{\delta})$ and $\frac{\partial \bar{X}_t}{\partial \delta}$ with an intercept.

Using (3.4), we obtain

$$N^{\frac{1}{2}}(\hat{\gamma}-0) \overset{d}{\rightsquigarrow} N(0, \text{plim}(N^{-1} \text{RSS}_1^*/\hat{\sigma}^2)^{-1}),$$

and because $\hat{\sigma}^2 = \text{RSS}_1/N$, we may define the Wald test statistic W by

$$W = \hat{\gamma}[N(\text{RSS}_1^*/\text{RSS}_1)]^{\frac{1}{2}}. \quad (5.3)$$

If H_0 is true, W is asymptotically standardised normal.

(c) *Lagrange multiplier test*

For this test the information matrix is evaluated at the restricted estimates $\tilde{\theta}$. Thus, analogously to (5.2),

$$\tilde{I}_{11} - \tilde{I}_{12} \tilde{I}_{22}^{-1} \tilde{I}_{21} = \text{plim}(N^{-1} \text{RSS}_0^*/\tilde{\sigma}^2), \quad (5.4)$$

where RSS_0^* is obtained by regressing Y_{t-1} on $\tilde{X}_t(\tilde{\delta})$ and $\frac{\partial \tilde{X}_t}{\partial \delta}$ with an intercept.

By (3.8), we have that

$$N^{-\frac{1}{2}} \left(\frac{\partial \tilde{L}}{\partial \theta_1} \right) \overset{d}{\rightsquigarrow} N(0, \text{plim}(N^{-1} \text{RSS}_0^*/\tilde{\sigma}^2)) \quad (5.5)$$

For the adaptive expectations model,

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial \theta_1} &\equiv \frac{\partial \tilde{L}}{\partial \gamma} = - \frac{1}{\tilde{\sigma}^2} \tilde{v}' \left(\frac{\partial \tilde{v}}{\partial \gamma} \right) \\ &= - \frac{1}{\tilde{\sigma}^2} \tilde{v}' (\tilde{\alpha}j + \tilde{\beta} \tilde{X}(\tilde{\delta}) - Y_{-1}). \end{aligned}$$

By the properties of least squares, the residuals $\tilde{v}$ satisfy

$$\tilde{v}'j = \tilde{v}'\tilde{X}(\tilde{\delta}) = 0,$$

and so

$$\begin{aligned}\frac{\partial L}{\partial \theta_1} &= \frac{1}{\tilde{\sigma}^2} \tilde{v}' Y_{-1} = N \frac{\tilde{v}' Y_{-1}}{\tilde{v}' \tilde{v}} \\ &= N \tilde{\eta}, \text{ say}\end{aligned}\tag{5.6}$$

where $\tilde{\eta}$ is the regression coefficient of Y_{t-1} on $\tilde{v}_t$. Combining (5.5) and (5.6), we define the Lagrange Multiplier statistic by

$$LM = \tilde{\eta} [N(RSS_0 / RSS_0^*)]^{\frac{1}{2}}.\tag{5.7}$$

When the null hypothesis is true, LM is asymptotically distributed as a standardised normal variable.

A summary of the operations required and the properties of the three tests is shown in Table 1.

Because the Likelihood Ratio test involves two search procedures and cannot produce one-sided tests, it appears to be inferior and is not recommended.

Of the Wald and Lagrange Multiplier tests, both are able to utilize one-sided procedures and should therefore be more powerful than the LR test. The W test involves one less simple regression so may be slightly preferable.

6. Testing the Partial Adjustment Model

The PA model is validated by testing $\delta=0$ in the model (2.4). The significant feature of the PA model, is that the disturbance is *not* a moving average. This means that maximum-likelihood estimation on the *restricted* (PA) model merely involves OLS regression. The need for a search procedure is avoided. It is clear then that the Lagrange Multiplier test, involving estimation of the restricted model alone, will be far more convenient than the Likelihood Ratio and Wald tests. Nevertheless, for the sake of completeness, these latter tests will be considered briefly.

TABLE 1 - Tests for the Adaptive Expectations Model

Test	Operations Required	Statistic	Distribution of Statistic under H_0
Likelihood	Unrestricted ML estimation $\rightarrow$ RSS_1	$LR = N \ln(RSS_0/RSS_1)$	χ^2_1
Ratio	Restricted ML estimation $\rightarrow$ RSS_0		
Wald	Unrestricted ML estimation $\rightarrow$ RSS_1 , $\hat{\gamma}$, $\hat{\delta}$.	$W = \hat{\gamma} [N(RSS_1^*/RSS_1)]^{\frac{1}{2}}$	$N(0,1^2)$
	OLS of Y_{t-1} on $\bar{X}_t(\hat{\delta})$, $\frac{\partial \bar{X}_t}{\partial \delta}$ (with intercept) $\rightarrow$ RSS_1^*		
Lagrange	Restricted ML estimation $\rightarrow$ RSS_0 , $\tilde{v}$.	$LM = \tilde{\eta} [N(RSS_0/RSS_0^*)]^{\frac{1}{2}}$	$N(0,1^2)$
Multiplier	OLS of Y_{t-1} on $\bar{X}_t(\tilde{\delta})$, $\frac{\partial \bar{X}_t}{\partial \delta}$ (with intercept) $\rightarrow$ RSS_0^*		
	OLS of Y_{t-1} on $\tilde{v}_t$ (no intercept) $\rightarrow$ $\tilde{\eta}$		

(a) *Likelihood-ratio test*

As in (5.1), the test statistic is given by

$$LR = N \ln(RSS_0/RSS_1). \quad (6.1)$$

Here RSS_1 has the same meaning as previously, but RSS_0 is the residual sum of squares obtained when the PA model (2.6) is estimated by OLS.

(b) *Wald test*

In this case $\theta_1 = \delta$, giving

$$\frac{\partial \hat{v}}{\partial \theta_1} = -\hat{\beta}(1-\hat{\gamma}) \left(\frac{\partial \hat{X}}{\partial \theta} \right)$$

$$\frac{\partial \hat{v}}{\partial \theta_1} = -[(1-\hat{\gamma})j, (1-\hat{\gamma})\bar{X}(\hat{\delta}), Y_{-1}].$$

Thus, RSS_1^* is obtained by regressing

$$\hat{\beta}(1-\hat{\gamma}) \frac{\partial \hat{X}}{\partial \theta} \text{ on } j, \bar{X}(\hat{\delta}) \text{ and } Y_{-1}.$$

With this meaning for RSS_1^* , the Wald statistic is then given by

$$W = \hat{\delta} [N(RSS_1^*/RSS_1)]^{\frac{1}{2}} \quad (6.2)$$

(c) *Lagrange multiplier test*

We have seen in Section 2 that the PAAE model may be regarded as a dynamic linear model, having a moving average disturbance and a non-linear restriction between the coefficients and the parameter of the disturbance. Breusch (1978) has shown that the well-known Durbin h-statistic is in fact a Lagrange Multiplier test of the hypothesis $\delta=0$. However, the h-test takes no account of the restriction, and would therefore be expected

to be less powerful than a LM test which does take this into account.

In order to show clearly the relationship between the LM test and the h-test, the LM test will be derived using the dynamic linear model (2.8) rather than the form (2.12) which has been used so far.

We define the matrix $M(\delta)$ by

$$M(\delta) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -\delta & 1 & 0 & \dots & 0 & 0 \\ 0 & -\delta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\delta & 1 \end{bmatrix} \quad (6.3)$$

Then

$$v = M^{-1}(\delta)u. \quad (6.4)$$

The likelihood function (2.11) can now be written in terms of u as

$$L(\theta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} u' S^{-1}(\delta) u \quad (6.5)$$

where

$$S(\delta) = M(\delta)M'(\delta). \quad (6.6)$$

Four properties of $S(\delta)$ which are easily established from (6.3) and (6.6) may be noted

$$(i) \quad S^{-1}(0) = I_N, \quad (6.7)$$

$$(ii) \quad \left(\frac{S^{-1}}{\partial \delta} \right)_{\delta=0} \equiv V = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad (6.8)$$

$$(iii) \quad \frac{1}{2} \left(\frac{\partial^2 S}{\partial \delta^2} \right)_{\delta=0} \equiv W = \begin{bmatrix} 0 & 0 \\ 0 & I_{N-1} \end{bmatrix}, \quad (6.9)$$

$$(iv) \quad \frac{1}{2} \left(\frac{\partial^2 S^{-1}}{\partial \delta^2} \right)_{\delta=0} = V^2 - W. \quad (6.10)$$

We define $\theta = (\delta, \beta^*)'$, where $\beta^* = (\beta_1, \beta_2, \beta_3)'$. The parameter σ^2 has been omitted because it can be established that $\text{plim } N^{-1} \frac{\partial^2 L}{\partial \delta \partial \sigma^2}$ and $\text{plim } N^{-1} \frac{\partial^2 L}{\partial \beta \partial \sigma^2}$ are both zero at the restricted estimate $\tilde{\theta}$.

From (6.5),

$$\frac{\partial L}{\partial \delta} = - \frac{1}{\sigma^2} \left[\frac{1}{2} u' \frac{\partial S^{-1}}{\partial \delta} u + u' S^{-1} \frac{\partial u}{\partial \delta} \right],$$

giving

$$\frac{\tilde{\partial} L}{\partial \delta} = - \frac{1}{\tilde{\sigma}^2} \left[\frac{1}{2} \tilde{u}' V \tilde{u} + \tilde{u}' \frac{\partial \tilde{u}}{\partial \delta} \right]. \quad (6.11)$$

It can now be shown using standard algebra that the information matrix $\tilde{I}$ is given by

$$\tilde{I} = \text{plim} \frac{1}{\tilde{\sigma}^2} \begin{bmatrix} \tilde{h}' \tilde{h} - \tilde{u}' W \tilde{u} & \tilde{h}' \tilde{Z} \\ \tilde{Z}' \tilde{h} & \tilde{Z}' \tilde{Z} \end{bmatrix},$$

where

$$\tilde{h} = V \tilde{u} + \frac{\partial \tilde{u}}{\partial \delta}, \quad (6.12)$$

and

$$\tilde{Z} = \frac{\partial \tilde{u}}{\partial \beta^*} = -(j, X, Y_{-1}). \quad (6.13)$$

Defining

$$\tilde{M} = I - \tilde{Z}(\tilde{Z}' \tilde{Z})^{-1} \tilde{Z}',$$

it follows that

$$\begin{aligned}\tilde{I}_{11}^{-1}\tilde{I}_{12}\tilde{I}_{22}^{-1}\tilde{I}_{21} &= \text{plim} \frac{1}{N\tilde{\sigma}^2}[\tilde{h}'\tilde{M}\tilde{h} - \tilde{u}'\tilde{W}\tilde{u}] \\ &= \text{plim} (RSS_0^*/RSS_0) - 1,\end{aligned}\tag{6.14}$$

where we have used

- (i) $\text{plim} \frac{1}{N\tilde{\sigma}^2} \tilde{u}'\tilde{W}\tilde{u} = 1$ (from (6.9)),
- (ii) $\tilde{\sigma}^2 = RSS_0/N$,
- (iii) RSS_0^* is the residual sum of squares obtained when $\tilde{h}$ is regressed on $\tilde{Z}$. Using (6.12) to evaluate $\tilde{h}$, it can be seen that RSS_0^* is obtained by regressing $\tilde{u}_{+1} + \tilde{u}_{-1} + \tilde{\beta}_3 Y_{-2}$ on j, X, Y_{-1} ,

$$\tag{6.15}$$

where u_{+1} is the vector containing observations on $\tilde{u}_{t+1}$.

From the conditional distribution of the score statistic (see (3.8)), we have

$$N^{-\frac{1}{2}} \frac{\partial \tilde{L}}{\partial \delta} \overset{\sim}{\sim} N \left(0, \text{plim} \frac{RSS_0^*}{RSS_0} - 1 \right)$$

leading to the definition of the LM statistic by

$$LM = \frac{\partial \tilde{L}}{\partial \delta} / \left[N \left(\frac{RSS_0^*}{RSS_0} - 1 \right) \right]^{\frac{1}{2}}.\tag{6.16}$$

This expression can be made more specific by evaluating $\partial \tilde{L}/\partial \delta$ from (6.11).

From the definition of V ,

$$\frac{1}{2} \tilde{u}'V\tilde{u} = \sum_{t=1}^{N-1} \tilde{u}_t' \tilde{u}_{t+1} = N \tilde{\sigma}^2 \tilde{r}_1 = N \tilde{\sigma}^2 (1 - \frac{1}{2} d)$$

where $\tilde{r}_1$ is the sample first-order autocorrelation coefficient of the residuals $\tilde{u}$ and d the Durbin-Watson Statistic. Also, $\partial \tilde{u} / \partial \delta = \tilde{\beta}_3 Y_{-2}$ because of the non-linear restriction.

Thus,

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial \delta} &= -\frac{1}{\tilde{\sigma}^2} [N \tilde{\sigma}^2 (1 - \frac{1}{2}d) + \tilde{\beta}_3 \tilde{u}' Y_{-2}] \\ &= -N [(1 - \frac{1}{2}d) + \tilde{\beta}_3 \tilde{u}' Y_{-2} / \tilde{u}' \tilde{u}] \\ &= -N [(1 - \frac{1}{2}d) + \tilde{\beta}_3 \tilde{\eta}], \end{aligned} \quad (6.17)$$

where $\tilde{\eta}$ is the regression coefficient of Y_{-2} on $\tilde{u}$.

Finally then

$$LM = N^{\frac{1}{2}} [(1 - \frac{1}{2}d) + \tilde{\beta}_3 \tilde{\eta}] / [(RSS_0^* / RSS_0) - 1]^{\frac{1}{2}}, \quad (6.18)$$

and when the null hypothesis $\delta=0$ is true, LM is asymptotically distributed as a standardised normal random variable.

The connection between LM and h statistics

If the non-linear constraint (2.9) is ignored it would follow that

$$\frac{\partial \tilde{u}}{\partial \delta} = 0$$

and hence that

$$\frac{\partial \tilde{L}}{\partial \delta} = -N(1 - \frac{1}{2}d)$$

and

$$h = Vu.$$

Thus,

$$\tilde{I}_{11} - \tilde{I}_{12} \tilde{I}_{22}^{-1} \tilde{I}_{21} = \text{plim} \frac{1}{N\sigma^2} [\tilde{u}' V^2 \tilde{u} - \tilde{u}' V Z (Z' Z)^{-1} Z' V \tilde{u}] - 1 \quad (6.19)$$

From (6.8), it can be seen that

$$\text{plim} \frac{\tilde{u}' V^2 \tilde{u}}{N\sigma^2} = 2, \quad (6.20)$$

and

$$\text{plim} \frac{\tilde{u}' V Z}{N\sigma^2} = [0, 0, 1].$$

Thus,

$$\begin{aligned} & \text{plim} \frac{1}{N\sigma^2} \tilde{u}' V Z (Z' Z)^{-1} Z' V \tilde{u} \\ &= \text{plim} \frac{\tilde{u}' V Z}{N\sigma^2} \quad \text{plim} N\sigma^2 (Z' Z)^{-1} \quad \text{plim} \frac{Z' V \tilde{u}}{N\sigma^2} \\ &= \text{plim} N\sigma^2 [0, 0, 1] (Z' Z)^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \text{plim} N S^2(\tilde{\beta}_3), \end{aligned} \quad (6.21)$$

where $S^2(\tilde{\beta}_3)$ is the estimated variance of $\tilde{\beta}_3$.

Substituting (6.20) and (6.21) into (6.19),

$$\tilde{I}_{11} - \tilde{I}_{12} \tilde{I}_{22}^{-1} \tilde{I}_{21} = 1 - \text{plim} N S^2(\tilde{\beta}_3),$$

giving

$$LM = (1 - \frac{1}{2}d) [N / (1 - N S^2(\tilde{\beta}_3))]^{\frac{1}{2}}. \quad (6.22)$$

which is precisely Durbin's h-statistic.

A summary of the operations required and the distributions of the test statistics under the null hypothesis are shown in Table 2. Because the Likelihood Ratio and Wald tests involve *unrestricted* estimation and hence, a search procedure, they are never likely to be preferred over the Lagrange Multiplier and h tests which require only OLS estimation. Of these latter two tests, the h test is easier to compute but will be less powerful than the LM test.

7. Empirical Power Curves

The collection of tests derived in the earlier sections and summarized in Tables 1 and 2, were all based on *asymptotic* results of maximum-likelihood estimation. In order to gain some experience of how these tests may perform with small samples, a very limited set of simulation experiments were undertaken and power curves estimated.

The model used for the experiments is characterised as follows:

- (i) Y_t ($t=1,2,\dots,N$) was generated according to equations (2.4) and (2.5), with $\alpha=0$, $\beta=5$ and $\sigma_v^2=1$.
- (ii) The values of X_t , fixed in repeated samples, were observations from

$$X_t = \rho X_{t-1} + \varepsilon_t,$$

where $\rho=0.99$, $\sigma_\varepsilon^2=1.99$.

This choice of σ_v^2 , ρ and σ_ε^2 meant that the ratio $\sigma_X^2/\sigma_v^2=100$.

Power curves were estimated for samples of size $N=20$ and $N=40$. The number of replications was 600 in the case of tests of the AE model and 1000 for the PA model, and the nominal size of the tests was 0.05.

In addition to the diagnostic tests shown in Tables 1 and 2, an additional test (denoted by OLS) was carried out on both the AE and PA models. This test consisted of estimating (2.4) by OLS and testing the significance of the

TABLE 2 - Tests for the Partial Adjustment Model

Test	Operations Required	Statistic	Distribution of Statistic under H_0
Likelihood	Unrestricted ML estimation $\rightarrow$ RSS_1	$LR = N \ln(RSS_0/RSS_1)$	χ_1^2
Ratio	Restricted ML estimation $\rightarrow$ RSS_0		
Wald	Unrestricted ML estimation $\rightarrow$ $RSS_1, \hat{\beta}, \hat{\gamma}, \hat{\delta}$ OLS of $\hat{\beta}(1-\hat{\gamma})\frac{\partial \bar{X}}{\partial \hat{\delta}}$ on $\bar{X}(\hat{\delta}), Y_{-1}$ (with intercept) $\rightarrow$ RSS_1^*	$W = \hat{\delta}[N(RSS_1^*/RSS_1)]^{\frac{1}{2}}$	$N(0, 1^2)$
Lagrange	OLS on PA model $\rightarrow$ $RSS_0, \tilde{u}, d, \tilde{\beta}_3$	$LM = N^{\frac{1}{2}}[1 - \frac{1}{2}d + \tilde{\beta}_3\tilde{\eta}]/$ $[(RSS_0^*/RSS_0) - 1]^{\frac{1}{2}}$	$N(0, 1^2)$
Multiplier	OLS of $\tilde{u}_{+1} + \tilde{u}_{-1} + \tilde{\beta}_3 Y_{-2}$ on X, Y_{-1} (with intercept) $\rightarrow$ RSS_0^*		
	OLS of Y_{-2} on $\tilde{u}$ (no intercept) $\rightarrow$ $\tilde{\eta}$		
Durbin	OLS on PA model $\rightarrow$ $d, S^2(\tilde{\beta}_3)$	$h = (1 - \frac{1}{2}d) \times$ $[N/(1-N S^2(\tilde{\beta}_3))]^{\frac{1}{2}}$	$N(0, 1^2)$
h-test			

coefficient of Y_{t-2} . It appears that such a test has been used to validate both the AE and the PA models (see Jones (1962), Nerlove (1958, p.248) and Doran and Griffiths (1978, p.140).

The main features emerging from the simulations are as follows:

(a) The AE Model (test of $\gamma=0$ - see Table 3)

- (i) The empirical size of the OLS test was generally lower than the nominal size 0.05. In addition, the power was low indicating that the OLS test would usually falsely support an AE or PA hypothesis.
- (ii) The LR, W and LM tests were usually much more powerful than the OLS test. Also, because W and LM tests are one-sided tests, they are more powerful than the LR test.
- (iii) The tests perform best in the midrange values of δ , that is, $\delta=0.4$ and 0.6 .
- (iv) Even when $N=20$, the W test shows considerable power - particularly in the midrange of δ .
- (v) When $N=40$, the empirical size of the W test is consistent with the nominal size 0.05. The LM test seems to be inferior to the W test.

For the AE model computational convenience matches empirical evidence and the W test, given in (5.3), is strongly recommended. The OLS test, based as it is on an inconsistent estimator should never be used.

TABLE 3 - Empirical Power Curves for the AE Model

N=20 (600 replications)		$\delta=0.2$					$\delta=0.4$					$\delta=0.6$					$\delta=0.8$				
γ		LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS
0.0		.083	.017	.008	.037	.107	.037	.010	.032	.115	.067	.007	.010	.123	.028	.013	.002				
0.2		.113	.088	.102	.138	.290	.475	.218	.210	.343	.562	.273	.136	.198	.285	.197	.012				
0.4		.415	.572	.318	.272	.717	.883	.575	.475	.782	.912	.740	.395	.518	.663	.587	.073				
0.6		.528	.790	.427	.298	.840	.965	.792	.578	.925	.978	.903	.573	.725	.818	.787	.193				
0.8		.450	.653	.427	.143	.713	.827	.722	.270	.790	.892	.850	.267	.613	.832	.612	.085				

N=40 (600 replications)		$\delta=0.2$					$\delta=0.4$					$\delta=0.6$					$\delta=0.8$				
γ		LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS	LR	W	LM	OLS
0.0		.060	.048	.020	.02	.087	.033	.017	.012	.075	.045	.018	.000	.100	.035	.012	.000				
0.2		.253	.175	.260	.183	.502	.633	.463	.212	.560	.752	.558	.100	.405	.547	.398	.008				
0.4		.737	.750	.687	.358	.918	.980	.908	.562	.970	.997	.982	.455	.925	.970	.932	.075				
0.6		.890	.938	.823	.345	.968	.995	.977	.655	.990	1.000	.998	.625	.993	.998	.993	.272				
0.8		.880	.960	.862	.138	.958	.992	.972	.348	.995	1.000	.997	.452	.992	1.000	.998	.357				

(b) The PA Model (test of $\delta=0$ - see Table 4)

Because LR and W tests involve a search procedure, whereas h, LM and OLS tests are carried out using OLS on the restricted model (2.6), only the properties of the latter tests were investigated.

- (i) A disappointing aspect of the results is the very poor performance of the h-test. It tends to incorrectly reject the null $\delta=0$ too often, and has very poor power.
- (ii) The OLS test appears to perform much better than for the AE model. Even for $N=20$ the empirical size of the test is consistent with 0.05, and in the midrange for γ the power is reasonable.
- (iii) The LM test rejects the true null $\delta=0$ more often than it should. However, the gain in power obtained by taking account of the non-linear restriction is very considerable when compared to the performance of the h-test.

For the PA model, the simple OLS test seems to be a viable procedure for small samples. However, the LM test (6.18), even though its empirical size is too large, has considerably more power.

TABLE 4 - Empirical Power Curves for the PA Model

N=20 (1000 replications)		$\gamma=0.2$			$\gamma=0.4$			$\gamma=0.6$			$\gamma=0.8$		
δ		h	LM	OLS	h	LM	OLS	h	LM	OLS	h	LM	OLS
0.0		.079	.142	.058	.088	.154	.064	.094	.122	.044	.098	.135	.037
0.2		.174	.361	.169	.115	.506	.305	.077	.545	.322	.104	.338	.147
0.4		.225	.498	.207	.092	.755	.462	.057	.797	.540	.058	.635	.271
0.6		.264	.453	.121	.085	.768	.394	.036	.849	.531	.029	.685	.295
0.8		.347	.275	.024	.143	.544	.101	.061	.579	.160	.046	.430	.097

N=40 (1000 replications)		$\gamma=0.2$			$\gamma=0.4$			$\gamma=0.6$			$\gamma=0.8$		
δ		h	LM	OLS	h	LM	OLS	h	LM	OLS	h	LM	OLS
0.0		.069	.080	.050	.081	.111	.051	.081	.096	.039	.087	.101	.038
0.2		.239	.461	.188	.154	.595	.337	.098	.620	.370	.078	.448	.173
0.4		.431	.621	.229	.163	.871	.538	.043	.902	.665	.043	.783	.397
0.6		.568	.640	.103	.191	.877	.432	.044	.960	.618	.019	.889	.448
0.8		.692	.404	.005	.310	.730	.065	.064	.863	.261	.013	.735	.304

8. Summary and Conclusions

Diagnostic tests for the PA and AE models have been derived by nesting these models in the more general PAAE model and applying the Likelihood Ratio (LR), Wald (W) and Lagrange Multiplier (LM) principles. The resulting tests are summarized in Tables 1 and 2. Computational convenience and the results of a simulation experiment lead to the W test (equation (5.3)) being recommended in the case of the AE model and the LM test (equation (6.18)) for the PA model.

Durbin's h-statistic, used as a test of the PA model, seems to perform very poorly.

The common test of applying OLS to the PAAE model and testing the significance of the coefficient of Y_{t-2} lacked power as a test of the AE model, but was somewhat better for the PA model. The recommended tests were always superior to this test.

References

- Atkinson, A.C. (1970), 'A Method for Discriminating Between Models', *Journal of the Royal Statistical Society B* 32, 323-353.
- Breusch, T.S. (1978), 'Testing for Autocorrelation in Dynamic Linear Models', *Australian Economic Papers* 17, 334-355.
- Breusch, T.S. and A.R. Pagan (1980), 'The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics', *Review of Economic Studies* 47, 239-253.
- Cagan, P. (1956), 'The Monetary Dynamics of Hyper Inflations', in M. Friedman (ed.), *Studies in the Quantity Theory of Money*, University of Chicago Press, Chicago.
- Cox, D.R. (1961), 'Tests of Separate Families of Hypotheses', *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, University of California Press, Berkeley.
- Cox, D.R. (1962), 'Further Results on Tests of Separate Families of Hypotheses', *Journal of the Royal Statistical Society B* 24, 406-424.
- Davidson, R. and J. MacKinnon (1981), 'Several Tests for Model Specification in the Presence of Alternative Hypotheses', *Econometrica* 49, 781-793.
- Dhrymes, P.J. (1971), *Distributed Lags: Problems of Estimation and Formulation*, Holden-Day, San Francisco.
- Dhrymes, P.J. (1978), *Introductory Econometrics*, Springer-Verlag, New York.
- Doran, H.E. and W.E. Griffiths (1978), 'Inconsistency of the OLS Estimator of the Partial Adjustment-Adaptive Expectations Model', *Journal of Econometrics* 7, 133-146.
- Engle, R.F. (1982), 'A General Approach to Lagrange Multiplier Model Diagnostics', *Journal of Econometrics* 20(1), *Annals of Applied Econometrics*, 83-104.

- Fisher, G.R. and M. McAleer (1981), 'Alternative Procedures and Associated Tests of Significance for Non-nested Hypotheses', *Journal of Econometrics* 16, 103-119.
- Graybill, F.A. (1961), *An Introduction to Linear Statistical Models*, Vol.1, McGraw-Hill, New York.
- Graybill, F.A. (1969), *Introduction to Matrices with Applications in Statistics*, Wadsworth, Belmont.
- Jones, G.T. (1962), 'The Response of the Supply of Agricultural Products in the United Kingdom to Price', Part II, *Farm Economist* 10, 1-28.
- McAleer, M. (1984), 'Specification Tests for Separate Models: A Survey', Working Paper in Economics and Econometrics No. 110, Australian National University.
- Nerlove, M. (1958), *The Dynamics of Supply: Estimation of Farmers' Response to Price*, Johns Hopkins, Baltimore
- Pagan, A.R. (1984), 'Model Evaluation by Variable Addition', in D.F. Hendry and K.F. Wallis (eds.), *Econometrics and Quantitative Economics*, Blackwell.
- Pesaran, M.H. (1974), 'On the General Problem of Model Selection', *Review of Economic Studies* 41, 153-171.

WORKING PAPERS IN ECONOMETRICS AND APPLIED STATISTICS

- The Prior Likelihood and Best Linear Unbiased Prediction in Stochastic Coefficient Linear Models.* Lung-Fei Lee and William E. Griffiths, No. 1 - March 1979.
- Stability Conditions in the Use of Fixed Requirement Approach to Manpower Planning Models.* Howard E. Doran and Rozany R. Deen, No. 2 - March 1979.
- A Note on a Bayesian Estimator in an Autocorrelated Error Model.* William Griffiths and Dan Dao, No. 3 - April 1979.
- On R^2 -Statistics for the General Linear Model with Nonscalar Covariance Matrix.* G.E. Battese and W.E. Griffiths, No. 4 - April 1979.
- Construction of Cost-Of-Living Index Numbers - A Unified Approach.* D.S. Prasada Rao, No. 5 - April 1979.
- Omission of the Weighted First Observation in an Autocorrelated Regression Model: A Discussion of Loss of Efficiency.* Howard E. Doran, No. 6 - June 1979.
- Estimation of Household Expenditure Functions: An Application of a Class of Heteroscedastic Regression Models.* George E. Battese and Bruce P. Bonyhady, No. 7 - September 1979.
- The Demand for Sawm Timber: An Application of the Diewert Cost Function.* Howard E. Doran and David F. Williams, No. 8 - September 1979.
- A New System of Log-Change Index Numbers for Multilateral Comparisons,* D.S. Prasada Rao, No. 9 - October 1980.
- A Comparison of Purchasing Power Parity Between the Pound Sterling and the Australian Dollar - 1979.* W.F. Shepherd and D.S. Prasada Rao, No. 10 - October 1980.
- Using Time-Series and Cross-Section Data to Estimate a Production Function with Positive and Negative Marginal Risks.* W.E. Griffiths and J.R. Anderson, No. 11 - December 1980.
- A Lack-Of-Fit Test in the Presence of Heteroscedasticity.* Howard E. Doran and Jan Kmenta, No. 12 - April 1981.
- On the Relative Efficiency of Estimators Which Include the Initial Observations in the Estimation of Seemingly Unrelated Regressions with First Order Autoregressive Disturbances.* H.E. Doran and W.E. Griffiths, No. 13 - June 1981.
- An Analysis of the Linkages Between the Consumer Price Index and the Average Minimum Weekly Wage Rate.* Pauline Beesley, No. 14 - July 1981.
- An Error Components Model for Prediction of County Crop Areas Using Survey and Satellite Data.* George E. Battese and Wayne A. Fuller, No. 15 - February 1982.

2.

*Networking or Transshipment? Optimisation Alternatives for Plant Location
Decisions.* H.I. Toft and P.A. Cassidy, No. 16 - February 1985