TWO-STEP, INSTRUMENTAL VARIABLE AND MAXIMUM LIKELIHOOD ESTIMATION OF MULTIVARIATE RATIONAL EXPECTATIONS MODELS

by

M. Hashem Pesaran

Trinity College, Cambridge
and
University of California, Los Angeles

January 1986

Revised: July 1988

UCLA Working Paper No. 493

ABSTRACT

This paper extends the results obtained by Pagan (1984), and Turkington (1985) for single equation rational expectation (RE) models to the multivariate RE models and shows that the errors-in-variables method and the substitution method discussed in Wickens (1982) lead exactly to the same likelihood function. Therefore, as far as the maximum likelihood estimation of RE models is concerned, the paper argues that there is little to choose between the substitution and the errors-in-variables methods. The paper also considers multivariate RE models with unanticipated variables, and demonstrates that in the simple case where only current values of the unanticipated variables are included in the RE model, the IV and the two-step estimators of the coefficients of the included unanticipated variables will be asymptotically efficient irrespective of whether the predetermined variables of the RE model are included amongst variables of the expectations formation model. This result does not, however, carry over to models with lagged values of unanticipated variables, unless it is assumed that the expectations errors are distributed independently of the future as well as the current and the past values of the explanatory variables of the RE model: an assumption which is not warranted under the REH.

1. Introduction

This paper extends the results obtained by Pagan (1984) and Turkington (1985) and reviewed in Pesaran (1987, Ch. 7) for single equation rational expectations (RE) models containing unobserved expectations of a single variable to the multivariate RE models. Apart from providing a generalization of Pagan-Turkington results, the paper also shows that the errors-in-variables method (EVM) and the substitution method (SM) discussed in Wickens (1982) for estimation of RE models lead exactly to the same likelihood function. It will be therefore shown that as far as the maximum likelihood (ML) estimation of the RE model is concerned, the two estimation methods are equivalent.

The plan of the paper is as follows: Section 2 deals with the basic RE model which is a multivariate version of Pagan's model 2. In the case of this model we obtain the asymptotic covariance matrices of the ML, the two-step, and the instrumental variable (IV) estimators and show that in general the two-step and the IV estimators are asymptotically less efficient than the ML estimators. The condition for the full asymptotic efficiency of the two-step and the IV estimators turns out to be the same as that already derived in the literature for the univariate case: namely that all the predetermined variables of the RE model should also be included amongst the explanatory variables of the expectations formation model. The exact equivalence of the likelihood functions of the EVM and the SM applied to the multivariate RE model of Section 2 will be demonstrated in Section 3. This result also establishes the full asymptotic efficiency of the three-stage least squares (3SLS) estimation method when applied to the errors-in-variable version of the RE model. In Section 4, we turn our attention to multivariate RE models with current and lagged unanticipated variables. We show that contrary to what is stated in the literature, it is possible to employ the EVM to obtain consistent estimates of the parameters of the unanticipated variables. We also demonstrate that in the simple case where only current values of the unanticipated variables are included in the RE model, the IV and the two-step estimators of the coefficients of the included unanticipated variables will be asymptotically efficient irrespective of whether the predetermined variables of the RE model are included amongst the variables of the expectations formation model. This result does not, however, carry over to models with lagged values of unanticipated variables, unless it is assumed that the expectations errors are distributed independently of the future as well as the current and the past values of the explanatory variables of the RE model; an assumption which is not warranted under the REH.

2. The Basic Multivariate RE Model

Consider the model

$$\underset{\sim}{y}_t = A_1' \underset{\sim}{x}_t^* + A_2' \underset{\sim}{w}_t + \underset{\sim}{u}_t, \tag{1}$$

$$\underset{\sim}{x}_t = B' \underset{\sim}{z}_t + \underset{\sim}{v}_t, \tag{2}$$

$$\underset{\sim}{x}_t^* = E(\underset{\sim}{x}_t | \Omega_{t-1}), \quad t = 1,2,\ldots,n, \tag{3}$$

where $\underset{\sim}{w}_t$ and $\underset{\sim}{z}_t$ are column vectors of predetermined variables of orders p and q respectively, and $\underset{\sim}{y}_t$ and $\underset{\sim}{x}_t$ are column vectors of endogenous variables of orders m and k respectively. The $k \times 1$ vector $\underset{\sim}{x}_t^*$ stands for the unobserved values of expectations of $\underset{\sim}{x}_t$ formed at time $t-1$. Under the REH, $\underset{\sim}{x}_t^*$ is given by (3) where $E(\underset{\sim}{x}_t | \Omega_{t-1})$ represents the mathematical expectations of $\underset{\sim}{x}_t$ conditional on the information set Ω_{t-1}, which is assumed to include current observations on at least the variables

$\underset{\sim}{w}_t$, $\underset{\sim}{z}_t$ and observations on the past values of $\underset{\sim}{y}_t$, $\underset{\sim}{x}_t$, $\underset{\sim}{w}_t$, $\underset{\sim}{z}_t$, namely $\Omega_{t-1} = (\underset{\sim}{w}_t, \underset{\sim}{z}_t; \underset{\sim}{y}_{t-1}, \underset{\sim}{x}_{t-1}, \underset{\sim}{w}_{t-1}, \underset{\sim}{z}_{t-1}; \underset{\sim}{y}_{t-2}, \underset{\sim}{x}_{t-2}, \underset{\sim}{w}_{t-2}, \underset{\sim}{z}_{t-2}, \ldots)$. The above model characterizes a general multivariate version of model 2 analyzed by Pagan (1984), and more recently by Turkington (1985). The Pagan-Turkington model corresponds to the simple case where $m = k = 1$. The unobservable variables models discussed in the literature by Zellner (1970), Goldberger (1972), and Aigner (1974) can also be viewed as special cases of the above basic model.

For the purpose of comparing the asymptotic efficiency of the alternative estimators of the parameters of the RE model (1), namely $\mathbf{A}_1$ and $\mathbf{A}_2$, we make the following standard assumptions.

<u>Assumption 1</u>: Conditional on the information set Ω_{t-1}, the disturbances $\underset{\sim}{\xi}_t = (\underset{\sim}{u}_t', \underset{\sim}{v}_t')'$ are normally distributed with zero means and the non-singular variance matrix Σ of order $m+k$.

<u>Assumption 2</u>: For the predetermined variables of the model, $\underset{\sim}{f}_t = (\underset{\sim}{z}_t', \underset{\sim}{w}_t')'$ the following probability limits exist

$$n^{-1} \sum_{t=1}^{n} \underset{\sim}{f}_t \underset{\sim}{f}_t' \xrightarrow{p} \Sigma_{ff} = \begin{pmatrix} \Sigma_{zz} & \Sigma_{zw} \\ \Sigma_{wz} & \Sigma_{ww} \end{pmatrix}, \tag{4}$$

and

$$n^{-1} \sum_{t=1}^{n} \underset{\sim}{f}_t \underset{\sim}{\xi}_t' \xrightarrow{p} 0,$$

where Σ_{ff} is a finite non-negative definite matrix.

<u>Assumption 3</u>: For the explanatory variables of the RE model (1), namely $h_t^* = (\underset{\sim}{z}_t'\mathbf{B}, \underset{\sim}{w}_t')'$, the following probability limit exists and is non-singular.

$$n^{-1} \sum_{t=1}^{n} \underset{\sim}{h}_t^* \underset{\sim}{h}_t^{*\prime} = \begin{bmatrix} \mathbf{B}'\Sigma_{zz}\mathbf{B} & \mathbf{B}'\Sigma_{zw} \\ \Sigma_{wz}\mathbf{B} & \Sigma_{ww} \end{bmatrix} = \Sigma_{h^*h^*}. \tag{5}$$

<u>Assumption 4</u>: The matrix Σ_{zz} is non-singular.

<u>Remarks</u>: The assumption that conditional on Ω_{t-1} the disturbances are normally distributed is only needed for the derivation of the ML estimators. In general the normality assumption can be relaxed without in any way affecting the results obtained in this paper. The assumption that Σ_{ff} is a non-negative definite matrix allows for the possibility of linear dependencies between the elements of $\underset{\sim}{z}_t$ and $\underset{\sim}{w}_t$. As will be shown below, this is an important consideration for the comparison of the asymptotic efficiency of the alternative estimators of $\mathbf{A}_1$ and $\mathbf{A}_2$. Assumption 4 ensures that the parameter matrix $\mathbf{B}$ is identified, while for a given value of $\mathbf{B}$, Assumption 3 ensures the identification of $\mathbf{A}_1$ and $\mathbf{A}_2$. Assumption 3 also implies that $\mathbf{B}'\Sigma_{zz}\mathbf{B}$ and Σ_{ww} are non-singular. A necessary condition for $\mathbf{B}'\Sigma_{zz}\mathbf{B}$ to be non-singular is given by $q \geq k$ which is clearly satisfied in the case of Pagan-Turkington univariate model where $k=1$. The orthogonality condition $E(\underset{\sim}{\xi}_t|\Omega_{t-1}) = 0$ implicit in Assumption 1, also ensures that $\underset{\sim}{\xi}_t$ have zero means (unconditionally) and are uncorrelated with their own past values and all the current and past values of $\underset{\sim}{f}_t$. Notice, however, that under the REH, $\underset{\sim}{\xi}_t$ are not necessarily uncorrelated with the future values of $\underset{\sim}{f}_t$. More specifically, the important implications of the orthogonality condition which will be used in this paper are

$$E(\underset{\sim}{\xi}_t) = 0,$$

$$E(\underset{\sim}{x}_t|\Omega_{t-1}) = \underset{\sim}{x}_t^* = \mathbf{B}'\underset{\sim}{z}_t,$$

$$E(\underset{\sim}{\xi}_t\underset{\sim}{\xi}_{t-i}') = 0, \quad \text{for } i = 1,2,\ldots$$

and $E(\underset{\sim}{\xi}_t f'_{t-i}) = 0, \quad \text{for} \quad i = 0,1,\ldots$

although $\underset{\sim}{\xi}_t$ need not be uncorrelated with the current and the future values of $\underset{\sim}{f}_t$.

2.1 The ML Estimators

In matrix form the equation system (1)-(3), can be written as:

$$\underset{\sim}{y} = (I_m \otimes X_*)\underset{\sim}{a}_1 + (I_m \otimes W)\underset{\sim}{a}_2 + \underset{\sim}{u}, \qquad (1')$$

$$\underset{\sim}{x} = (I_k \otimes Z)\underset{\sim}{b} + \underset{\sim}{v}, \qquad (2')$$

$$X_* = ZB, \qquad (3')$$

where I_m and I_k are the identity matrices of orders m and k respectively, $\underset{\sim}{y}$ is the $nm\times1$ vector formed by stacking $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_n$, and $\underset{\sim}{x}$ is an $nk\times1$ column vector formed by stacking $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_n$. Similarly, $\underset{\sim}{u} = (\underset{\sim}{u}'_1, \underset{\sim}{u}'_2, \ldots, \underset{\sim}{u}'_n)'$, and $\underset{\sim}{v} = (\underset{\sim}{v}'_1, \underset{\sim}{v}'_2, \ldots, \underset{\sim}{v}'_n)'$. The other notations are as follows: $X' = (\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_n)$, $Z' = (\underset{\sim}{z}_1, \underset{\sim}{z}_2, \ldots, \underset{\sim}{z}_n)$, $W' = (\underset{\sim}{w}_1, \underset{\sim}{w}_2, \ldots, \underset{\sim}{w}_n)$, $\underset{\sim}{a}_1 = \text{Vec}(A_1)$, $\underset{\sim}{a}_2 = \text{Vec}(A_2)$, $\underset{\sim}{b} = \text{Vec}(B)$, where $\text{Vec}(A_1)$ stands for the column vector of order mk formed by stacking the columns of A_1. Notice that in this notation $\underset{\sim}{x} = \text{Vec}(X')$.[1]

Under Assumption 1 the log-likelihood function of the model is given by

$$\ell(A_1, A_2, B, \Sigma) \propto \sum_{t=1}^{n} \log\{p(\underset{\sim}{\xi}_t | \Omega_{t-1})\},$$

where $p(\underset{\sim}{\xi}_t | \Omega_{t-1})$ is the conditional probability density function of $\underset{\sim}{\xi}_t$ assumed to have the multivariate normal distribution with zero means and the variance matrix Σ. In terms of the above notations the log-likelihood function can be written as

---

[1]The operation $A \otimes B$ represents the Kronecker product of the matrices A and B.

$$\ell(\underset{\sim}{\theta}) \propto -\frac{n}{2}\log|\Sigma| - \frac{1}{2}\underset{\sim}{\xi}'(\Sigma^{-1}\otimes I_n)\underset{\sim}{\xi}, \tag{6}$$

where $\underset{\sim}{\theta} = (\underset{\sim}{a}', \underset{\sim}{b}', \underset{\sim}{\sigma}')'$, $\underset{\sim}{a}' = (\underset{\sim}{a}_1', \underset{\sim}{a}_2')$, $\underset{\sim}{\sigma}$ is the $\frac{1}{2}(k+m)(k+m+1) \times 1$ column vector containing the distinct elements of Σ, and

$$\underset{\sim}{\xi} = \begin{bmatrix} \underset{\sim}{y} - (I_m \otimes X_*)\underset{\sim}{a}_1 - (I_m \otimes W)\underset{\sim}{a}_2 \\ \\ \underset{\sim}{x} - (I_k \otimes Z)\underset{\sim}{b} \end{bmatrix}, \tag{7}$$

which in view of (3') can also equivalently be written as

$$\underset{\sim}{\xi} = \begin{bmatrix} \underset{\sim}{y} - (A_1' \otimes Z)b - (I_m \otimes W)\underset{\sim}{a}_2 \\ \\ \underset{\sim}{x} - (I_k \otimes Z)\underset{\sim}{b} \end{bmatrix}. \tag{7'}$$

Let $\hat{\underset{\sim}{\gamma}}_{ML}$ and $\hat{\underset{\sim}{\sigma}}_{ML}$ be the ML estimators of $\underset{\sim}{\gamma} = (\underset{\sim}{a}', \underset{\sim}{b}')'$, and $\underset{\sim}{\sigma}$ respectively. Then it is easily seen that under our assumptions $\hat{\underset{\sim}{\gamma}}_{ML}$ and $\hat{\underset{\sim}{\sigma}}_{ML}$ are asymptotically independently distributed and that $\sqrt{n}(\hat{\underset{\sim}{\gamma}}_{ML} - \underset{\sim}{\gamma})$ has a limiting normal distribution with zero means and the asymptotic covariance matrix given by the Cramer-Rao lower bound $I_{\gamma\gamma}^{-1}$, where

$$I_{\gamma\gamma} = -\underset{n\to\infty}{\text{plim}} \left\{\frac{1}{n}\frac{\partial^2 \ell(\underset{\sim}{\theta})}{\partial\underset{\sim}{\gamma}\partial\underset{\sim}{\gamma}'}\right\} = \underset{n\to\infty}{\text{plim}} \left\{\frac{1}{n}\left(\frac{\partial \ell(\underset{\sim}{\theta})}{\partial\underset{\sim}{\gamma}}\right)\left(\frac{\partial \ell(\underset{\sim}{\theta})}{\partial\underset{\sim}{\gamma}}\right)'\right\}.$$

To obtain $I_{\gamma\gamma}$, using relations (6) and (7) we first note that

$$\frac{\partial \ell(\underset{\sim}{\theta})}{\partial\underset{\sim}{\gamma}} = \begin{bmatrix} I_m \otimes H_*' & 0 \\ \\ A_1 \otimes Z' & I_k \otimes Z' \end{bmatrix} (\Sigma^{-1}\otimes I_n)\underset{\sim}{\xi}, \tag{8}$$

where $H_* = (X_* : W) = (ZB : W)$. Let Σ be partitioned into m and k columns

$$\Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{pmatrix} = \begin{pmatrix} \Sigma^{uu} & \Sigma^{uv} \\ \Sigma^{vu} & \Sigma^{vv} \end{pmatrix}^{-1},$$

then using (8) it is easily established that

$$I_{\gamma\gamma} = \begin{bmatrix} \Sigma^{uu} \otimes \Sigma_{h*h*} & K \otimes \Sigma_{h*z} \\ K' \otimes \Sigma_{zh*} & \Phi \otimes \Sigma_{zz} \end{bmatrix}, \tag{9}$$

where Σ_{h*h*} is already defined in (5), and

$$K = \Sigma^{uu} A_1' + \Sigma^{uv}, \tag{10}$$

$$\Phi = A_1 \Sigma^{uu} A_1' + \Sigma^{vu} A_1' + A_1 \Sigma^{uv} + \Sigma^{vv} > 0, \tag{11}$$

$$\Sigma_{zh*} = (\Sigma_{zz} B, \Sigma_{zw}).$$

The asymptotic variance matrix of $\sqrt{n}(\hat{\underset{\sim}{a}}_{ML} - \underset{\sim}{a})$ can now be obtained from $I_{\gamma\gamma}$ and is given by

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{ML}) = \{\Sigma^{uu} \otimes \Sigma_{h*h*} - K\Phi^{-1}K' \otimes \Sigma_{h*z} \Sigma_{zz}^{-1} \Sigma_{zh*}\}^{-1}.$$

However, using result 2.9 in Rao (1965, p. 29) we have

$$(\Sigma^{uu} - K\Phi^{-1}K')^{-1} = \Psi,$$

where

$$\Psi = A_1' \Sigma_{vv} A_1 - \Sigma_{uv} A_1 - A_1' \Sigma_{vu} + \Sigma_{uu}. \tag{13}$$

Therefore, asymptotically

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{ML}) = [\Sigma^{uu} \otimes \Sigma_{h*h*} + (\Psi^{-1} - \Sigma^{uu}) \otimes \Sigma_{h*z} \Sigma_{zz}^{-1} \Sigma_{zh*}]^{-1}.$$

The first term in the square bracket represents the sampling error of estimating $\underset{\sim}{a}$ for a known value of $\underset{\sim}{b}$, while the second term in the square bracket represents the sampling error associated with the estimation of $\underset{\sim}{b}$.

2.2 The Two Step Estimators

The first step of the two-step estimators employed in the literature involves estimating $\mathbf{B}$ by the OLS regressions of the columns of $\mathbf{X}$ on $\mathbf{Z}$. In the second step the parameters $\mathbf{A}_1$ and $\mathbf{A}_2$ are then estimated by regression of $\mathbf{Y}$ on $\hat{\mathbf{X}}_*$ and $\mathbf{W}$, where $\hat{\mathbf{X}}_* = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X} = \mathbf{P}_z\mathbf{X}$. In terms of our notations the two-step estimators of $\underset{\sim}{a}$ (say $\hat{\underset{\sim}{a}}_{2s}$) is given by

$$\hat{\underset{\sim}{a}}_{2s} = [\mathbf{I}_m \otimes (\hat{\mathbf{H}}_*'\hat{\mathbf{H}}_*)^{-1}\hat{\mathbf{H}}_*']\underset{\sim}{y}, \tag{15}$$

where $\hat{\mathbf{H}}_* = (\mathbf{P}_z\mathbf{X} : \mathbf{W})$. In order to give some idea of the properties of the two-step estimator $\hat{\underset{\sim}{a}}_{2s}$, we first rewrite (1') as

$$\underset{\sim}{y} = (\mathbf{I}_m \otimes \hat{\mathbf{H}}_*)\underset{\sim}{a} + \underset{\sim}{\epsilon}, \tag{16}$$

where

$$\underset{\sim}{\epsilon} = \underset{\sim}{u} - (\mathbf{A}_1' \otimes \mathbf{P}_z)\underset{\sim}{v}.$$

It is now clear that unlike $\underset{\sim}{u}$, the variance matrix of the compound disturbance vector $\underset{\sim}{\epsilon}$ is not spherical. We have

$$\mathbf{\Sigma}_{\epsilon\epsilon} = V(\underset{\sim}{\epsilon}) = \mathbf{\Sigma}_{uu} \otimes \mathbf{I}_n + (\mathbf{\Psi} - \mathbf{\Sigma}_{uu}) \otimes \mathbf{P}_z, \tag{17}$$

where $\mathbf{\Psi}$ is already defined by (13). Seen in this light it is not surprising that in general $\hat{\underset{\sim}{a}}_{2s}$ is not an efficient estimator of $\underset{\sim}{a}$, since it does not allow for the fact that $V(\underset{\sim}{\epsilon})$ in general is non-spherical. A two-step estimator of $\underset{\sim}{a}$ which takes account of the specific structure of $V(\underset{\sim}{\epsilon})$ can be obtained by application of the generalized least squares method to (16). The resultant estimator

$$\hat{\underset{\sim}{a}}_{GLE} = [(\mathbf{I}_m \otimes \hat{\mathbf{H}}_*')\mathbf{\Sigma}_{\epsilon\epsilon}^{-1}(\mathbf{I}_m \otimes \hat{\mathbf{H}}_*)]^{-1}(\mathbf{I}_m \otimes \hat{\mathbf{H}}_*')\mathbf{\Sigma}_{\epsilon\epsilon}^{-1}\underset{\sim}{y}, \tag{18}$$

will not, however, be of much practical value, since its computation is complicated and requires an *a priori* estimate of $\mathbf{\Sigma}$ and $\mathbf{A}_1$. But it may be interesting to note that in the special case where $\mathbf{W}$ is contained in $\mathbf{Z}$ (i.e., $\mathbf{W} \subset \mathbf{Z}$), the two-step estimator $\hat{\underset{\sim}{a}}_{2s}$ and the generalized least squares estimator $\hat{\underset{\sim}{a}}_{GLE}$ will be algebraically <u>identical</u>. This result follows from the fact that when $\mathbf{W} \subset \mathbf{Z}$ then $P_z\hat{H}_* = \hat{H}_*$, and

$$(\Sigma_{\epsilon\epsilon}^{-1} \otimes \hat{H}_*) = \Sigma_{uu}\Psi^{-1} \otimes \hat{H}_* + (I_m - \Sigma_{uu}\Psi^{-1}) \otimes P_z\hat{H}_*,$$

$$= I_m \otimes \hat{H}_*,$$

or $\Sigma_{\epsilon\epsilon}^{-1}(I_m \otimes \hat{H}_*) = \Psi^{-1} \otimes \hat{H}_*$. Using this result in (18) now gives $\hat{\underset{\sim}{a}}_{2s} = \hat{\underset{\sim}{a}}_{GLE}$.

To obtain the asymptotic variance matrix of $\hat{\underset{\sim}{a}}_{2s}$ we note that using (16) in (15) yields

$$\hat{\underset{\sim}{a}}_{2s} - \underset{\sim}{a} = [I_m \otimes (\hat{H}_*'\hat{H}_*)^{-1}\hat{H}_*']\underset{\sim}{\epsilon},$$

from which and with the aid of (17) the following expression for the asymptotic variance matrix of $\hat{\underset{\sim}{a}}_{2s}$ can be obtained[2]

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{2s}) = \Psi \otimes \Sigma_{h*h*}^{-1} + (\Sigma_{uu} - \Psi) \otimes \Sigma_{h*h*}^{-1}\ \Sigma_0\Sigma_{h*h*}^{-1}, \tag{19}$$

where

$$\Sigma_0 = \Sigma_{h*h*} - \Sigma_{h*z}\Sigma_{zz}^{-1}\Sigma_{zh*}. \tag{20}$$

Similarly, in terms of Σ_0 the asymptotic variance matrix of $\hat{\underset{\sim}{a}}_{ML}$, given by (14), becomes

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{ML}) = [\Psi^{-1} \otimes \Sigma_{h*h*} + (\Sigma^{uu} - \Psi^{-1}) \otimes \Sigma_0]^{-1}. \tag{21}$$

2.3 <u>The IV Estimators</u>

Underlying the IV estimators for $\underset{\sim}{a}$ lies the errors-in-variables method of solving and estimating the RE models first suggested by McCallum (1976) and more recently advocated by Wickens (1982). Since under the REH

---

[2]A general derivation of the asymptotic variance matrix of two-step estimators in the case of a univariate model can be found in Newey (1984) and Murphy and Topel (1985).

$\underset{\sim}{x}_t^* = B'\underset{\sim}{z}_t = \underset{\sim}{x}_t - \underset{\sim}{v}_t$, then the system of equations (1′) and (2′) can also be written as

$$\underset{\sim}{y} = (I_m \otimes H)\underset{\sim}{a} + \underset{\sim}{e}, \tag{1''}$$

$$\underset{\sim}{x} = (I_k \otimes Z)\underset{\sim}{b} + \underset{\sim}{v}, \tag{2'}$$

where $H = (X : W)$, and

$$\underset{\sim}{e} = \underset{\sim}{u} - (A_1' \otimes I_n)\underset{\sim}{v}. \tag{22}$$

In this characterization of the RE model, H and $\underset{\sim}{e}$ will be correlated irrespective of whether $\Sigma_{uv} = 0$ or not. An obvious estimation method therefore would be to apply the IV procedure to (1") using all the predetermined variables of the model, namely $F = (Z : W)$, as instruments. This gives the following IV estimator of $\underset{\sim}{a}$,

$$\hat{\underset{\sim}{a}}_{IV} = [I_m \otimes (H'P_f H)^{-1} H'P_f]\underset{\sim}{y} \tag{23}$$

where $P_f = F(F'F)^{-1}F'$. Furthermore, it is easily seen that

$$V(\underset{\sim}{e}) = E(\underset{\sim}{e}\underset{\sim}{e}') = \Psi \otimes I_n,$$

where Ψ is already defined by (13). The relevant expression for the asymptotic variance of $\sqrt{n}(\hat{\underset{\sim}{a}}_{IV} - \underset{\sim}{a})$ will therefore be

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{IV}) = \Psi \otimes \underset{n\to\infty}{\text{plim}}\ (H'P_f H/n)^{-1} \tag{24}$$

$$= \Psi \otimes \Sigma_{h^*h^*}^{-1}.$$

2.4 <u>A Comparison of Alternative Estimators</u>

From the above results, it is firstly clear that in general the two-step and the IV estimators are less efficient than the ML estimators. Under our assumptions all the three estimators are equally efficient only in two circumstances: (i) if $A_1 = 0$ and $\Sigma_{uv} = 0$ or (ii) if $W \subset Z$. In the case where $A_1 = 0$ and $\Sigma_{uv} = 0$ the asymptotic variance matrices of all the three estimators collapse to $\Sigma_{uu} \otimes \Sigma_{h*h*}^{-1}$; and when $W \subset Z$ then $\Sigma_0 = 0$ and all the three variance matrices reduce to the expression $\Psi \otimes \Sigma_{h*h*}^{-1}$. In general, however, the two-step and the IV estimators are less efficient than the ML estimators. Furthermore, as shown by Turkington (1985) for the univariate case $(k=m=1)$, ranking of the IV and the two-step estimators by their asymptotic variance matrices, in general, is not possible. But when $\Sigma_{uv} = 0$ it is easily seen that the two-step estimators are asymptotically more efficient than the IV estimators. From (19) and (24) we have

$$V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{IV}) - V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{2s}) = (\Psi - \Sigma_{uu}) \otimes \Sigma_{h*h*}^{-1}\Sigma_0\Sigma_{h*h*}^{-1}.$$

Therefore, the asymptotic efficiency of the two-step estimators relative to the IV estimators depends on the "sign" of $\Psi - \Sigma_{uu} = A_1'\Sigma_{vv}A_1 - A_1'\Sigma_{vu} - \Sigma_{uv}A_1$. In the special case where $\Sigma_{uv} = 0$ the matrix $\Psi - \Sigma_{uu} = A_1'\Sigma_{vv}A_1$ is non-negative definite and $V(\sqrt{n}\ \hat{\underset{\sim}{a}}_{IV}) - V(\sqrt{n}\ \hat{}_{\sim 2s}) \geq 0$.

3. Likelihood Function For The Errors-In-Variables Method

In this section we show that the log likelihood function associated with the errors-in-variables version of the RE model (1)-(3) is algebraically equal to the log likelihood function (6) obtained in the previous section via the substitution method. Thus we demonstrate that insofar as the ML estimation of our basic RE model is concerned there is little to choose between the errors-in-variables and the substitution methods discussed in Wickens (1982). The equivalence of the likelihood functions under these two solution methods, however, establish that the three stage least squares estimators of $\underset{\sim}{a}$ in the equations system (1") and (2') will be asymptotically as efficient as the ML estimators $\hat{\underset{\sim}{a}}_{ML}$ discussed in Section 2.1.

Consider the errors-in-variables representation of the RE model given by (1") and (2'). Let $\underset{\sim}{\eta} = (\underset{\sim}{e}', \underset{\sim}{v}')'$. Then the log likelihood function associated with the EVM will be

$$\ell_{EV}(\underset{\sim}{\theta}) \propto -\frac{1}{2}\log|V(\underset{\sim}{\eta})| - \frac{1}{2}\underset{\sim}{\eta}'[V(\underset{\sim}{\eta})]^{-1}\underset{\sim}{\eta}, \tag{25}$$

where $V(\underset{\sim}{\eta})$ is the variance matrix of $\underset{\sim}{\eta}$. We, however, know from (22) that $\underset{\sim}{e} = \underset{\sim}{u} - (A_1'\otimes I_n)\underset{\sim}{v}$. Therefore we can write

$$\underset{\sim}{\eta} = \begin{bmatrix} I_m\otimes I_n & -A_1'\otimes I_n \\ 0 & I_k\otimes I_n \end{bmatrix}\underset{\sim}{\xi},$$

or $\underset{\sim}{\eta} = (R\otimes I_n)\underset{\sim}{\xi}$, where

$$R = \begin{bmatrix} I_m & -A_1' \\ 0 & I_k \end{bmatrix}.$$

Using (26) it follows that $V(\underset{\sim}{\eta}) = (R\Sigma R'\otimes I_n)$, and since R is a block triangular matrix with all its diagonal elements equal to unity then $|V(\underset{\sim}{\eta})| = |\Sigma|^n$. Substituting these results in (25) now yields

$$\ell_{EV}(\underset{\sim}{\theta}) \propto -\frac{n}{2}\log|\Sigma| - \frac{1}{2}\underset{\sim}{\nu}'(\Sigma^{-1}\otimes I_n)\underset{\sim}{\nu},$$

where $\underset{\sim}{\nu} = (R^{-1}\otimes I_n)\underset{\sim}{\eta}$. But it is easily seen that

$$\underset{\sim}{\nu} = \begin{bmatrix} I_{mn} & A_1'\otimes I_n \\ 0 & I_{kn} \end{bmatrix}\begin{bmatrix} \underset{\sim}{e} \\ \underset{\sim}{v} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{e} + (A_1'\otimes I_n)\underset{\sim}{v} \\ \underset{\sim}{v} \end{bmatrix},$$

which in view of (22) establishes that $\underset{\sim}{\nu} = \underset{\sim}{\xi}$. Therefore, $\ell_{EV}(\underset{\sim}{\theta})$ given by (25) is algebraically equivalent to the log likelihood function associated with the substitution method given by (6).

4. Some Extensions Of The Basic Model

4.1 Model With Current Unanticipated Effects

Consider first the following important extension of the basic model

$$\underset{\sim}{y}_t = A_1'\underset{\sim}{x}_t^* + A_2'\underset{\sim}{w}_t + C'(\underset{\sim}{x}_t - \underset{\sim}{x}_t^*) + \underset{\sim}{u}_t, \tag{27}$$

where as before $\underset{\sim}{x}_t$ is jointly determined by (2′). This is a multivariate generalization of Pagan's model 4 and allows for the possible effects of the "unanticipated" components of $\underset{\sim}{x}_t$ on $\underset{\sim}{y}_t$. Univariate versions of (27) have figured prominently in the literature on the empirical analysis of efficient market hypotheses and the natural rate-rational expectations hypothesis.

The present model can be estimated as before. But the parameter matrice C cannot be identified unless one is prepared to impose the *a priori* restriction that $\Sigma_{uv} = 0$. To see why this is so note that under the REH, $\underset{\sim}{x}_t - \underset{\sim}{x}_t^* = \underset{\sim}{v}_t$ and the errors-in-variables version of (27) becomes

$$\underset{\sim}{y}_t = A_1'\underset{\sim}{x}_t + A_2'\underset{\sim}{w}_t + \underset{\sim}{e}_t, \tag{28}$$

where $\underset{\sim}{e}_t$ is now given by

$$\underset{\sim}{e}_t = \underset{\sim}{u}_t - (A_1 - C)'\underset{\sim}{v}_t. \tag{29}$$

It is clear that the simultaneous system (28) and (2′) which is a reparameterization of (27) and (2′) will in general be "observationally equivalent" to the basic model; unless it is assumed that $\Sigma_{uv} = 0$. Under the restriction $\Sigma_{uv} = 0$, using (29) we have

$$\Sigma_{ev} = -(A_1 - C)'\Sigma_{vv}, \tag{30}$$

where $\Sigma_{ev} = E(\underset{\sim}{e}_t\underset{\sim}{v}_t')$. The parameter matrices A_1, A_2, B, Σ_{ev} and Σ_{vv} are all identifiable (irrespective of whether $\Sigma_{uv} = 0$ or not), and C can be identified from (30). Thus contrary to what is stated in the literature, the EVM can be employed to obtain consistent estimates of the parameters of the RE model (27), in spite of the fact that it contains realized and expected values of the same variables (Wickens, 1982, pp. 56-57). The estimation of (27) by the EVM can be carried out in two stages: First the IV method can be applied to (28) to obtain the IV estimates of $\mathbf{A}_1$ and $\mathbf{A}_2$ using, as in the case of the basic model, all the predetermined variables of the model, $\mathbf{F} = (\mathbf{Z} : \mathbf{W})$ as instruments. Then $\mathbf{C}$ can be consistently estimated by means of (30) noting that

$$\hat{\Sigma}_{vv} = n^{-1}(\mathbf{X}'\mathbf{M}_z\mathbf{X}),$$

and

$$\hat{\Sigma}_{ev} = n^{-1}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{A}}_{1,IV} - \mathbf{W}\hat{\mathbf{A}}_{2,IV})'\mathbf{M}_z\mathbf{X},$$

where $\mathbf{M}_z = \mathbf{I}_n - \mathbf{P}_z$.[3] That is

$$\hat{\mathbf{C}}_{IV} = (\mathbf{X}'\mathbf{M}_z\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}_z(\mathbf{Y} - \mathbf{W}\hat{\mathbf{A}}_{2,IV}). \tag{31}$$

The asymptotic variance matrix of $\sqrt{n}(\hat{\underset{\sim}{a}}_{IV}-\underset{\sim}{a})$ is still given by (24) but with this difference that $\mathbf{\Psi}$ is now equal to

$$\Sigma_{ee} = \Sigma_{uu} + (\mathbf{A}_1 - \mathbf{C})'\Sigma_{vv}(\mathbf{A}_1 - \mathbf{C}). \tag{32}$$

To obtain the asymptotic variance matrix of $\sqrt{n}(\hat{\mathbf{C}}_{IV}-\mathbf{C})$ we first note that by utilizing the matrix forms of (28), (29) and (2′) in (31) we can arrive at

$$\hat{\mathbf{C}}_{IV} - \mathbf{C} = (\mathbf{V}'\mathbf{M}_z\mathbf{V})^{-1}(\mathbf{V}'\mathbf{M}_z\mathbf{U}) - (\mathbf{V}'\mathbf{M}_z\mathbf{V})^{-1}(\mathbf{V}'\mathbf{M}_z\mathbf{W})(\hat{\mathbf{A}}_{2,IV} - \mathbf{A}_2),$$

where $\mathbf{V}' = (\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_n)$ and $\mathbf{U}' = (\underset{\sim}{u}_1,\underset{\sim}{u}_2,\ldots,\underset{\sim}{u}_n)$. But under Assumptions 1 and 2 and the identification restriction $\Sigma_{uv} = 0$, we have

$$n^{-1}(\mathbf{V}'\mathbf{M}_z\mathbf{V}) \xrightarrow{p} \Sigma_{vv}, \quad n^{-1}(\mathbf{V}'\mathbf{M}_z\mathbf{U}) \xrightarrow{p} 0, \quad n^{-1}(\mathbf{V}'\mathbf{M}_z\mathbf{W}) \xrightarrow{p} 0.$$

---

[3]Recall that $\hat{\underset{\sim}{a}}_{1,IV} = \mathrm{Vec}(\hat{\mathbf{A}}_{1,IV})$, $\hat{\underset{\sim}{a}}_{2,IV} = \mathrm{Vec}(\hat{\mathbf{A}}_{2,IV})$, $\hat{\underset{\sim}{a}}_{IV} = (\hat{\underset{\sim}{a}}'_{1,IV},\hat{\underset{\sim}{a}}'_{2,IV})'$, and $\hat{\underset{\sim}{a}}_{IV}$ is already given by (23).

We also know that $\sqrt{n}(\hat{A}_{2,IV} - A_2)$ is asymptotically normally distributed with a finite variance. Hence it follows that $\hat{C}_{IV}$ is a consistent estimator of C and $\sqrt{n}(\hat{C}_{IV}-C)$ has the same asymptotic distribution as $\Sigma_{vv}^{-1}(V'U/\sqrt{n})$. Therefore, in stacked form

$$\sqrt{n}(\hat{\underset{\sim}{c}}_{IV}-\underset{\sim}{c}) \overset{d}{\rightarrow} n^{-\frac{1}{2}}(U'\otimes\Sigma_{VV}^{-1})\underset{\sim}{v},$$

and since by assumption $\underset{\sim}{u}_t$ and $\underset{\sim}{v}_t$ are independently distributed, then asymptotically

$$V(\sqrt{n}\hat{\underset{\sim}{c}}_{IV}) = \underset{n\rightarrow\infty}{\text{plim}}\ n^{-1}\{(U'\otimes\Sigma_{vv}^{-1})(\underset{\sim}{v}\underset{\sim}{v}')(U\otimes\Sigma_{vv}^{-1})\},$$

$$= \underset{n\rightarrow\infty}{\text{plim}}\ \{(\frac{U'U}{n}) \otimes \Sigma_{vv}^{-1}\},$$

$$= \Sigma_{uu} \otimes \Sigma_{vv}^{-1},$$

and

$$\sqrt{n}(\hat{\underset{\sim}{c}}_{IV}-\underset{\sim}{c}) \overset{d}{\rightarrow} N(0,\ \Sigma_{uu}\otimes\Sigma_{vv}^{-1}).$$

The above results also show that $\hat{\underset{\sim}{a}}_{IV}$ and $\hat{\underset{\sim}{c}}_{IV}$ are asymptotically independently distributed. This is hardly surprising considering that under the REH, the term $\underset{\sim}{v}_t = \underset{\sim}{x}_t - \underset{\sim}{x}_t^*$ in (26) is distributed independently of the other explanatory variables included in the model.

To examine the asymptotic efficiency of the above IV estimators, let $\hat{\underset{\sim}{a}}_{ML}$ and $\hat{\underset{\sim}{c}}_{ML}$ by the ML estimators of $\underset{\sim}{a}$ and $\underset{\sim}{c}$ in (27) respectively. Then along the lines similar to that in Section (2.1) we obtain

$$V(\sqrt{n}\hat{\underset{\sim}{a}}_{ML}) = [\Sigma_{ee}^{-1}\otimes\Sigma_{h*h*} + (\Sigma_{uu}^{-1} - \Sigma_{ee}^{-1}) \otimes \Sigma_0]^{-1},$$

and

$$V(\sqrt{n}\hat{\underset{\sim}{c}}_{ML}) = \Sigma_{uu} \otimes \Sigma_{vv}^{-1},$$

where Σ_{h*h*} and Σ_0 are given as before by (5) and (20) and Σ_{ee} is defined above by (32). These results firstly establish that the IV estimator of $\underset{\sim}{c}$ given by (31) is asymptotically efficient irrespective of whether W is contained in Z or not. But as in the case of the basic model, the asymptotic efficiency of $\hat{\underset{\sim}{a}}_{IV}$ depends on whether $W \subset Z$. The same is also true of the two-step estimators of $\underset{\sim}{a}$ and $\underset{\sim}{c}$. Asymptotically nothing can be gained by the ML estimation of $\underset{\sim}{c}$. As far as the estimation of $\underset{\sim}{a}$ is concerned the two-step estimator, $\hat{\underset{\sim}{a}}_{2s}$ is less efficient than the ML estimator, $\hat{\underset{\sim}{a}}_{ML}$ if W is not contained in Z, but nevertheless is more efficient than the IV estimator, $\hat{\underset{\sim}{a}}_{IV}$. This is because under the identifying restrictions $\Sigma_{uv} = 0$, asymptotically we have (see Section 2.4)

$$V(\sqrt{n}\hat{\underset{\sim}{a}}_{IV}) - V(\sqrt{n}\hat{\underset{\sim}{a}}_{2s}) = (A_1 - C)'\Sigma_{vv}(A_1 - C) \otimes \Sigma^{-1}_{h*h*}\Sigma_0\Sigma^{-1}_{h*h*},$$

which is a non-negative definite matrix for the values of the unknown parameters. In general, therefore, the two-step estimator of $\underset{\sim}{a}$ in (27) is to be preferred to the IV estimators, assuming of course that the processes generating $\underset{\sim}{x}_t$ are correctly specified.

4.2 Models With Current and Lagged Unanticipated Effects

Consider now the following general distributed lag version of (27) where lagged values of $\underset{\sim}{x}^*_t$ and $\underset{\sim}{x}_t - \underset{\sim}{x}^*_t$ as well as their current values appear amongst the explanatory variables:

$$\underset{\sim}{y}_t = \sum_{i=1}^{s} A'_i\, \underset{\sim}{x}^*_{t-i} + A'_{s+1}\, \underset{\sim}{w}_t + \sum_{i=0}^{s} C'_i(\underset{\sim}{x}_{t-i} - \underset{\sim}{x}^*_{t-i}) + \underset{\sim}{u}_t. \tag{33}$$

The univariate version of this model used in empirical work, for example, by Barro (1977) and Mishkin (1982), has been extensively analyzed in the econometric literature, notably by Abel and Mishkin (1983), Leiderman (1980), and Attfield et al. (1981). The presence of lagged values of the unanticipated variables amongst the explanatory variables has important consequences both for the asymptotic distribution of the ML estimators and for the efficiency of the two-step estimators. This is primarily due to the fact that although under the REH $\underset{\sim}{v}_t = \underset{\sim}{x}_t - \underset{\sim}{x}_t^*$ is uncorrelated with $\underset{\sim}{x}_t$, $\underset{\sim}{w}_t$, and their lagged values, the same is not necessarily true of $\underset{\sim}{v}_t$ and the future values of $\underset{\sim}{x}_t^*$ and $\underset{\sim}{w}_t$. As a result the asymptotic distribution of the various estimators of $\underset{\sim}{c} = (\underset{\sim}{c}_0', \underset{\sim}{c}_1', \ldots, \underset{\sim}{c}_s')'$ will not be independent of the asymptotic distribution of the estimators of $\underset{\sim}{a} = (\underset{\sim}{a}_0', \underset{\sim}{a}_1', \ldots, \underset{\sim}{a}_{s+1}')'$. One important consequence of this is that the two-step estimator of $\underset{\sim}{c}$ in the general distributed lag model (33) will no longer be asymptotically efficient, unless it is assumed that

$$n^{-1} \sum_{t=1}^{n} \underset{\sim}{f}_{t+i} \, \underset{\sim}{v}_t' \overset{p}{\rightarrow} 0, \quad \text{for} \quad i = 1,2,\ldots$$

which is not warranted under the REH.

The errors-in-variables method can still be applied to the general distributed lag model (33), but will be much more cumbersome to implement than the two-step estimators. The errors-in-variables version of (33) can be written as

$$y_t = \sum_{i=0}^{s} \mathbf{A}_i' \, \underset{\sim}{x}_{t-i} + \underset{\sim}{\mathbf{A}}_{s+1}' \, \underset{\sim}{w}_t + \underset{\sim}{e}_t,$$

where

$$\underset{\sim}{e}_t = \underset{\sim}{u}_t - \sum_{i=0}^{s} (\mathbf{A}_i - \mathbf{C}_i)' \underset{\sim}{v}_{t-i}.$$

Now even if the REH holds, the orthogonality conditions $E(\underset{\sim}{e}_t | \underset{\sim}{f}_{t-j}) = 0$ need not hold for $j = 0,1,2,\ldots,s-1$. The REH only ensures that $E(\underset{\sim}{e}_t | \underset{\sim}{f}_{t-j}) = 0$, for $j = s, s+1, \ldots,$ and as a result the use of f_t or its lagged values $f_{t-1}, f_{t-2}, \ldots, f_{t-s+1}$ as instruments for the IV estimation of (34) need not be valid. The appropriate set of instruments in this general case would include $f_{t-s}, f_{t-s-1}, \ldots$ and $x_{t-s-1}, x_{t-s-2}, \ldots.$

To obtain the IV estimates of C_i we can again use the following moment relations

$$\Sigma_{ei} = (C_i - A_i)'\Sigma_{vv}, \quad \text{for} \quad i = 0,1,\ldots,s$$

where $\Sigma_{ei} = E(e_t v'_{t-i})$.

The two-step estimation of the parameters of the general model (33) does not, however, pose any special problems. Rewriting the general model in terms of the first step estimators $\hat{B} = (Z'Z)^{-1}Z'X$, we have

$$y_t = \sum_{i=0}^{s} A'_i \hat{x}^*_{t-i} + A'_{s+1} w_t + \sum_{i=0}^{s} C'_i \hat{v}_{t-i} + \epsilon_t, \tag{35}$$

where $\hat{v}_t = x_t - \hat{x}^*_t$, $\hat{x}^*_t = \hat{B}' z_t$ and

$$\epsilon_t = u_t - \sum_{i=0}^{s} (A_i - C_i)'(\hat{B} - B)' z_{t-i}.$$

Therefore, so long as $\hat{B}$ converges in probability to B, the two-step estimators will be consistent and the addition of lagged values of v_t to the model does not create new problems, at least as far as the consistent estimation of the parameters are concerned. By contrast the IV method, when appropriately implemented, will lead to consistent estimates of $A_0, A_1, \ldots$ even if the expectations formation model (2′) happens to be misspecified.

5. Concluding Remarks

Most tests of the rational expectations hypothesis, whether in the area of market efficiency or the policy neutrality, utilize univariate models; although given the interactions between markets and nominal and real variables in the economy a multivariate framework may often be more appropriate. In this paper we have considered a number of multivariate RE models that should be of interest in analysis of market interactions under the REH. We show that most of the results obtained for the univariate RE models can be generalized to the multivariate case. We also derive a number of new results not noted in the literature. We show the equivalence of the errors-in-variables and the substitution methods when there are no lagged anticipated variables in the model, and therefore establish the asymptotic efficiency of the 3SLS method when applied to the structural and the predictions equations treated as a joint system of equations. In the case when the RE model contains lagged anticipated (or unanticipated) variables the econometric analysis of the RE models is much more complicated and requires special care. Some of these difficulties are also emphasized in the paper.

References

Abel, A.B., and F.S. Mishkin (1983), "An Integrated View of Tests of Rationality, Market Efficiency and the Short-Run Neutrality of Monetary Policy," Journal of Monetary Economics, 11, 3-24.

Aigner, D. (1974), "An Appropriate Econometric Framework for Estimating a Labor Supply Function From the SEO File," International Economic Review, 15, 59-68.

Attfield, C.L.F., D. Demery and N.W. Duck (1981), "A Quarterly Model of Unanticipated Monetary Growth, Output and the Price Level in the U.K. 1963-1978," Journal of Monetary Economics, 8, 331-50.

Barro, R.J. (1977), "Unanticipated Money Growth and Unemployment in the United States," American Economic Review, 67, 101-15.

Goldberger, A.S. (1972), "Maximum-Likelihood Estimation of Regressions Containing Unobservable Independent Variables," International Economic Review, 13, 1-15.

Leiderman, L. (1980), "Macroeconometric Testing of the Rational Expectations and Structural Neutrality Hypotheses for the United States," Journal of Monetary Economics, 6, 69-82.

McCallum, B.T. (1976), "Rational Expectations and the Estimation of Econometric Models: An Alternative Procedure," International Economic Review, 17, 484-90.

Mishkin, F. (1982), "Does Anticipated Monetary Policy Matter? An Econometric Investigation," Journal of Political Economy, 84, 207-37.

Murphy, K.M. and R.H. Topel (1985), "Estimation and Inference in Two-Step Econometric Models," Journal of Business and Economic Statistics, 3, 370-79.

Newey, W.K. (1984), "A Method of Moments Interpretation of Sequential Estimators," Economic Letters, 14, 201-06.

Pagan, A. (1984), "Econometric Issues in the Analysis of Regressions With Generated Regressors," International Economic Review, 25, 221-47.

Pesaran, M.H. (1987), The Limits to Rational Expectations, Basil Blackwell, Oxford.

Rao, C.R. (1965), Linear Statistical Inference and Its Application, John Wiley, New York.

Turkington, D.A. (1985), "A Note on Two-Stage Least Squares, Three-Stage Least Squares and Maximum Likelihood Estimation in an Expectations Model," International Economic Review, XLIX, 55-67.

Wickens, M.R. (1982), "The Efficient Estimation of Econometric Models With Rational Expectations," Review of Economic Studies, XLIX, 55-67.

Zellner, A. (1970), "Estimation of Regression Relationships Containing Unobservable Independent Variables," International Economic Review, 11, 441-54.