# RATIONAL FORECASTS FROM NON-RATIONAL MODELS

by

A. Snell

NUMBER 194

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK
COVENTRY

# RATIONAL FORECASTS FROM NON-RATIONAL MODELS

by

A. Snell

NUMBER 194

October 1981

Paper presented at the Econometric Society European Meeting in 1981

I wish to thank Kenneth F. Wallis for numerous invaluable discussions during the writing of this paper and Adrian Pagan, Sean Holly and John Driffell for helpful comments on an earlier draft.

determined schemes are considered because of their overwhelming popularity. An unbiased and optimal (in the sense of minimum forecast error variance) extrapolative predictor is also described and used in the analysis because it provides a useful benchmark as the 'best' extrapolative proxy available. Section 4 examines the implications for estimation of using these three extrapolative proxies in the context of a simple two equation macro model due to Wallis (1980) and section 5 extends the model to include dynamics. The simple two equation model used bears a very close resemblance to the two main equations of the condensed St. Louis model described by Anderson and this, with its linearity and simplicity make it an ideal structure in which to house the analysis. Section 6 gives some numerical comparisons of multiplier error using the method under feasible values for the parameters. Section 7 discusses further problems that arise using the algorithm for policy analysis even when a set of consistent estimates is used. Focus here is on the seriousness of ignoring mistakes during simulation by substituting actual outcomes for expectations and on problems raised by Lucas' critique of policy evaluation (Lucas 1976 ). Section 8 provides a summary and conclusion.

## 2. Outline of the Fair-Anderson method

In their two influential papers Anderson (1979) and Fair (1979) undertake dynamic deterministic ex post policy simulations on the Fair (1976) and St. Louis models with a general aim of deriving dynamic policy multipliers. The method is easily exposited in the context of the following simple two equation model explaining a price level and nominal income

$$(2.1) \qquad p_t = \alpha m_t + u_{1t}$$

$$(2.2) \qquad y_t = \beta m_t + \gamma p_t^e + u_{2t}$$

where superscript 'e' denotes on expectation formed at time t-1 and $p_t$, $y_t$, $m_t$ and $u_{1t}$ and $u_{2t}$ are the price level, nominal income, money supply and structural disturbances respectively (all variables in natural logs). The model is incomplete without an assumption about expectation formation and until the R.E. revolution macro models such as this were typically estimated incorporating extrapolative schemes such as

$$(2.3) \qquad p_t^e = \sum_{i=1}^{n} \delta_i \, p_{t-i}$$

The R.E. of prices $(p_t^*)$ however is defined as

$$p_t^* = E(p_t | \Omega_{t-1})$$

where $\Omega_{t-1}$ is an information set containing the model ((1.1) and (1.2) in this case) and all variables up to time t-1 and so

$$p_t^* = \alpha\hat{m}_t \tag{2.4}$$

where a hat denotes an optimal (minimum forecast error variance) extrapolative prediction.

The R.E. compares starkly with any extrapolative scheme such as (2.3) since it involves optimal prediction of the exogenous variables (in this case, the money supply) and the form of this predictor will obviously depend on the exogenous variable process.

Having obtained estimates of the structural parameters of (2.1) and (2.2) by incorporating (2.3) and imposing an arbitrary restriction on the $\delta$'s (such as that they sum to one) to identify $\gamma$ one might proceed to a simulation exercise. This would normally involve setting all structural errors at their means of zero (for deterministic simulation) and numerically solving (2.1) to (2.3) under different money supply settings with a general aim of deriving policy multipliers such as

$$\frac{\partial p_i}{\partial m_j} \text{ and } \frac{\partial y_i}{\partial m_j} \quad (i \geqslant j) \text{ .}$$ [1/]

Our simulated values would thus satisfy

---

1/ Obviously in our linear example the multipliers are independent of initial conditions. This is not so in nonlinear models where multipliers such as $\left.\frac{\partial p_i}{\partial m_j}\right|_{y_0, \ldots}$ are calculated.

$$(2.5) \qquad p_t^{\,s} = am_t^{\,s}$$

$$(2.6) \qquad y_t^{\,s} = bm_t^{\,s} + c \sum_{i=1}^{n} \delta_i p_{t-i}^{s} = bm_t^{s} + ca \sum_{i=1}^{n} \delta_i\, m_{t-i}^{s}$$

where superscript 's' denotes a simulated value. Numerical values for multipliers are easily calculated by comparing the time paths of endogenous variables (given the same initial conditions) under different money supply settings ( often , focus is on the effect of a 'blip' or a shock in the exogenous variables, in our case money supply, at time $t$ ).

To repeat the exercise under a 'maintained hypothesis of R.E.'s' Anderson and Fair suggest simulating under the expectations scheme

$$(2.7) \qquad p_t^{\,e} = p_t$$

in place of (2.3) so as to make 'expectations consistent with the predictions of the model'. Justification of this comes from the fact that R.E.'s differ from actual values only by a stochastic error consisting of current structural disturbances via the reduced form and current innovations in the exogenous variables. In our example the error in the R.E. is

$$p_t - p_t^{*} = \alpha m_t + u_{1t} - \alpha \hat{m}_t = u_{1t} + \alpha\varepsilon_t$$

where $\varepsilon_t$ is one step ahead prediction error of the money supply. In deterministic simulations however the model is solved with structural disturbances set at their expected values of zero so that imposing

(2.7) it is claimed provides approximate rationality. For example in our model such 'consistent' expectations are given as

$$(2.8) \qquad p_t^e = p_t^s = am_t^s$$

as an approximation to the R.E. in (2.4). Using (2.2) and (2.8) the substitution for income is then

$$(2.9) \qquad y_t^s = (b + ca)m_t^s$$

If the hypothesis of R.E.'s is maintained from the beginning, comparable forms for price and income are obtained by solving (2.1), (2.2) and (2.4) to give

$$(2.10) \qquad p_t = \alpha m_t + u_{1t}$$

$$(2.11) \qquad y_t = (\beta + \gamma\alpha)\hat{m}_t - \gamma\alpha\varepsilon_t + u_{2t}$$

By comparing (2.8) and (2.9) with (2.10) and (2.11) ignoring differences due both to structural disturbances and money supply innovations ((2.8) and (2.9) do, in fairness, describe a deterministic simulation) we see that the essential distinction rests on the estimates a , b and c . 1/ Making expectations consistent with the predictions of the model as in (2.8) and (2.9) is not imposing R.E.'s because the structural parameter estimates used therein were obtained under a different expectations hypothesis, in our example the ad hoc scheme in (2.3). The main force of this paper is to show that under a maintained hypothesis of R.E.'s, estimation incorporating an ad hoc expectations scheme will lead to a

---

1/ In his paper Fair replaces a term structure equation incorporating a simple extrapolative expectations scheme with an efficient markets R.E. condition. Whilst 2SLS estimates in other equations remain consistent the choice of instruments used in the first stage for these is no longer optimal.

biased and inconsistent set of structural parameters. These biases cannot be overcome by then simulating under a maintained hypothesis of R.E.'s in the manner described above.

## 3. Extrapolative proxies in common use

Of the extrapolative proxies typically used in macro modelling we may distinguish those of finite order from those of infinite order. Infinite schemes obviously require a restriction on the lag structure for implementation and the exponentially declining lag of adaptive expectations provides a classic example. Finite schemes may or may not be further restricted; an Almon lag would have the lag weights following a polynomial in the lag operator whereas the scheme in (2.3) allows the data to determine the weights up to the imposed truncation point (n) . We consider estimation incorporating each scheme in turn in the context of the income equation (2.2) above. Using the finite but otherwise unrestricted scheme of (2.3) we would substitute (2.3) into (2.2) and estimate freely by O.L.S.

$$(3.1) \qquad y_t = bm_t + c \sum_{i=1}^{n} \delta_i \, p_{t-i} + u_t$$

The restriction that the $\delta$'s sum to one would be imposed afterwards to identify $c$ . The maximum lag $n$ may be chosen 'sufficiently' large to capture the bulk of the distributed lag, leaving the error largely free of autocorrelation although it has often been set equal to one giving a simple lagged variable proxy. The advantage of this procedure lies in allowing the data to determine the form of the distributed lag where cues from the underlying theory are weak. Its weakness lies in the arbitrary restriction required to identify the structural parameter. In dynamic models further such restrictions are required.

Among all extrapolative schemes available those of the Almon lag and adaptive expectations stand out through frequency of use.

Using the former would mean estimating

$$y_t = bm_t + c \sum_{i=1}^{n} \delta(i)\, p_{t-i} + u_t$$

where $\delta(i) = K_0 + K_1 i + K_2 i^2 + \ldots K_m i^m$

with additional 'degrees of freedom' requirement that $n$ be greater than $m + 1$ ($m$ is commonly set equal to two giving a quadratic form) and this is achieved using restricted least squares methods.

Invoking adaptive expectations gives

$$(3.2) \qquad y_t = bm_t + c \left[\frac{(1 - \delta)}{(1 - \delta L)}\right] p_{t-1} + u_t$$

where $L$ is the lag operator.

A Koyck transform is often applied to give

$$(3.3) \qquad y_t = bm_t - b\delta m_{t-1} + c(1-\delta)p_{t-1} + \delta y_{t-1} + (1-\delta L)u_t$$

and this is typically estimated by nonlinear least squares methods incorporating the relevant nonlinear parameter restrictions but approximating the induced first order M.A. error term with an autoregressive error scheme.[1/]

The next three sections describe the implications for estimates derived using these proxies when expectations are really rational. In any linear dynamic simultaneous model which includes exogenous variables so generated there always exists a univariate representation for each endogenous variable (Prothero and Wallis , 1975 ) from which may be obtained a predictor that is unbiased and which has minimum one-step-ahead forecast error variance among the class of purely extrapolative predictors. This optimal extrapolative predictor (henceforth O.E.P.) is to our knowledge, never purposefully used in nonrational models although it may by chance coincide with an ad hoc scheme (we consider an example in the next section). It nevertheless provides a useful benchmark in the analysis because, like an R.E., it is a conditional expectation, but it differs from the R.E. by being conditioned on a smaller information set namely the past values of the variable itself and as a result it is less efficient than the R.E.

---

1/ A programme by Osborne recently written would incorporate the M.A. and so provides an exact ML estimation procedure.

## 4. Estimation using these proxies under a maintained hypothesis of rationality

Consider the following static two equation model explaining first differences in nominal income and the price level;

(4.1) $$p_t = \beta_{12} y_t + \alpha p_t^e + u_{1t}$$

(4.2) $$y_t = \beta_{21} p_t + \gamma g_t + u_{2t}$$

where $p_t$ , $y_t$ and $g_t$ are first differences of a price level, nominal income and real government expenditure respectively. (All variables are in natural logarithms). This has reduced form

(4.3) $$p_t = \Pi_{11} p_t^e + \Pi_{12} g_t + v_{1t}$$

(4.4) $$y_t = \Pi_{21} p_t^e + \Pi_{22} g_t + v_{2t}$$

where $\Pi_{11} = \alpha/\Delta$, $\Pi_{12} = \gamma\beta_{12}/\Delta$ , $\Pi_{21} = \alpha\beta_{21}/\Delta$ , $\Pi_{22} = \gamma/\Delta$ , $\Delta = 1 - \beta_{12}\beta_{21}$ and the reduced form errors $(v_{jt})$ are linear combinations of the structural errors $(u_{jt})$ . Equations (4.1) and (4.2) very closely resemble a static form of equations 11 and 12 in Anderson's paper (op. cit.), the two main equations of his 'condensed version' of the St. Louis model (Anderson has money supply in place of the price level in (4.2)). Expected price changes enter (4.1) to represent the effect on prices of a wage bargain made for the current period but based on previous information and the income term represents aggregate demand pressures. Because labour costs only form a fraction of total costs, then given a mark up rule for pricing (augmented by demand pressures) the elasticity $\alpha$ is less than unity. Making $\beta_{21}$ unity in (4.2) gives real income changes as a function of changes in government expenditure and so the model is clearly Keynesian in its structure.

To write (4.3) and (4.4) in terms of observables we require an assumption for expectations. Under a **rational** expectations regime the reduced form of the model is

$$(4.5) \qquad p_t = \frac{\Pi_{11}\,\Pi_{12}}{1 - \Pi_{11}} \hat{g}_t + \Pi_{12}\, g_t + v_{1t}$$

$$(4.6) \qquad y_t = \frac{\Pi_{21}\,\Pi_{12}}{1 - \Pi_{11}} \hat{g}_t + \Pi_{22}\, g_t + v_{2t}$$

Finally we require a process for government expenditure to obtain the predictor $\hat{g}_t$ of the current value. Under the general ARMA (p, q) representation

$$\phi(L) g_t = \theta(L) \varepsilon_t$$

where $\phi(L)$ and $\theta(L)$ have leading coefficients of one, the predictor for $g_t$ is given by

$$\hat{g}_t = L^{-1}\left[1 - \frac{\phi(L)}{\theta(L)}\right] g_{t-1}$$

and on substitution of this in (4.5) and (4.6) and multiplication throughout by $\theta(L)$ we have the final equations

$$(4.7) \qquad \theta(L)p_t = \frac{\Pi_{11}\Pi_{12}}{1-\Pi_{11}} L^{-1}\left[\theta(L)-\phi(L)\right]g_{t-1} + \theta(L)\Pi_{12}g_t + \theta(L)v_{1t}$$

$$(4.8) \qquad \theta(L)y_t = \frac{\Pi_{21}\Pi_{12}}{1-\Pi_{11}} L^{-1}\left[\theta(L)-\phi(L)\right]g_{t-1} + \theta(L)\Pi_{22}g_t + \theta(L)v_{2t}$$

(4.7) and (4.8) express $p_t$ and $y_t$ in terms of $q$ own lagged values and the current and (max $p, q$) lagged values of $g_t$. Note also that these final form equations have an M.A. error of order $q$.

Under the maintained hypothesis of R.E.'s (4.7) and (4.8) represent the data generation process and we now turn to consider estimation of the $\Pi$'s by incorporating an ad hoc proxy of some sort.

Consider first employing the finite scheme in (2.3)

$$(4.9) \qquad p_t = p_{11} \sum_{i=1}^{n} \delta_i p_{t-i} + p_{12}g_t + z_{1t}$$

$$(4.10) \qquad y_t = p_{21} \sum_{i=1}^{n} \delta_i p_{t-i} + p_{22}g_t + z_{2t}$$

(with $\sum_{i=1}^{n} \delta_i = 1$ imposed to identify the $p_{i1}$) .

Note that if $n$ is greater than one (4.2) is overidentified and so (4.9) and (4.10) are jointly restricted and in such a situation 2SLS is normally used. Since 2SLS is dealt with in the next section we shall consider using O.L.S. here to obtain estimates of the reduced form coefficients. (Multipliers are after all direct functions of these and not the structural coefficients). By comparison with (4.7) and (4.8) we see that estimates based on (4.9) and (4.10) are seriously inconsistent. Even if $n \geq q$ so that the correct number of lagged values of $p_t$ enter (4.9) omitted variables (lagged values of $g_t$ which enter as a result of optimal prediction associated with R.E.'s) causes inconsistency and this is aggrevated by ignoring the M.A. error.[1/] (4.10) excludes lagged values of $g_t$ and q lagged values of the dependent variable and also has an unrecognised M.A. error. Asymptotic values for the estimates in this general case are hard to derive and probably wholly uninformative. For $n = 1$, however (lagged variable proxy) the calculation is tractable and asymptotic values for the $p_{ij}$'s in this case are derived in Appendix A.

In section 6 some numerical results on these parameter (and multiplier) biases are discussed.

Using an adaptive expectations hypothesis in the reduced form (4.3) and (4.4) means that the following system will be estimated:

(4.11) $$p_t = [p_{11}(1-\delta)+\delta]p_{t-1} + p_{12}g_t - p_{12}\delta g_{t-1} + (1-\delta L)z_{1t}$$

(4.12) $$y_t = p_{21}(1-\delta)p_{t-1} + \delta y_{t-1} + p_{22}g_t - p_{22}\delta g_{t-1} + (1-\delta L)z_{2t}$$

If $p = q = 1$ in (4.7) and (4.8) and if there are no estimation restrictions imposed, then the estimates of the coefficients of all the variables in (4.11) and (4.12) will be consistent estimates of their counterparts in (4.7) and (4.8). However when these coefficient estimates are 'unscrambled' using the restrictions in (4.7) and (4.8) the structural parameters will be wrongly identified. If either p or q is greater than one or if the restrictions in (4.11) and (4.12) are

---

1/ A careful data miner may diagnose and attempt to correct autocorrelation so induced by imposing an autoregressive error (normally to first order). Even so, an AR error has an infinite number of autocovariances and can only approximate the finite number of the M.A.(q) errors.

imposed for estimation (this is rarely done) then all estimates are inconsistent.

It might be tempting to assert at this stage that the nearer the extrapolative proxy used is to the O.E.P. implied by the data generation process then the less inconsistent the estimates so obtained. This is not necessarily the case because the O.E.P. itself still provides inconsistent estimates as we now show.

Consider the O.E.P. for $p_t$ derived from (4.7) (full derivation is given in Appendix B).

(4.13) $$\hat{p}_t = L^{-1}\left[1 - \frac{\phi(L)}{Q(L)}\right]p_{t-1}$$

where $Q(L)$ is of order $K = \max(p, q)$, has leading coefficient of one and has coefficients which are functions of all the parameters involved in the model and the g-process.

Now using the identity

$$\hat{p}_t + \mu_t \equiv p_t \equiv p_t^* + \Pi_{12}\varepsilon_t + v_{1t}$$

(where $\mu_t$ is the one step ahead forecast error of the O.E.P.)

or more explicitly

(4.14) $$p_t^* \equiv \hat{p}_t + \mu_t - \Pi_{12}\varepsilon_t - v_{1t}$$

and substituting $\hat{p}_t$ into (4.3) and (4.4) we obtain

$$(4.15) \qquad p_t = \Pi_{11}\hat{p}_t + \Pi_{12}g_t + \zeta_{1t}$$

$$(4.16) \qquad y_t = \Pi_{21}\hat{p}_t + \Pi_{22}g_t + \zeta_{2t}$$

where the error terms are given by

$$\zeta_{jt} = \Pi_{j1}\mu_t - \Pi_{j1}\Pi_{12}\varepsilon_t - \Pi_{j1}v_{1t} + v_{jt} \quad (j = 1, 2)$$

then applying least squares to (4.15) and (4.16) gives inconsistent estimates since the $\zeta_{jt}$ are correlated with the exogenous regressand $g_t$ through both the presence of $\varepsilon_t$ and of $\mu_t$ ((B7) in Appendix B shows $\mu_t$ to be correlated with all past $\varepsilon$'s and $v$'s). In general the one step ahead error of a purely extrapolative predictor is correlated with all the exogenous variables in the model and use of any such predictor in place of the R.E. will always lead to inconsistent estimates.[1/]

To take a specific example consider the ARIMA(0, 1, 1) process for $g_t$ (obtained by setting $p = q = 1$ and $\phi = 1$)

$$(1 - L)g_t = (1 - \theta L)\varepsilon_t$$

so that the forecast $\hat{g}_t$ is given by the adaptive expectations scheme

---

1/ Such a warning was first signalled by Nelson (1975).

$$\hat{g}_t = \frac{(1-\theta)}{(1-\theta L)} g_{t-1}$$

The final equations (4.7) and (4.8) now become

$$(4.21) \qquad p_t = \theta p_{t-1} + \left[\frac{\Pi_{11}\Pi_{12}}{1-\Pi_{11}}(1-\theta) - \theta\Pi_{12}\right] g_{t-1} + \Pi_{12} g_t + (1-\theta L) v_{1t}$$

$$(4.22) \qquad y_t = \theta y_{t-1} + \left[\frac{\Pi_{21}\Pi_{12}}{1-\Pi_{11}}(1-\theta) - \theta\Pi_{22}\right] g_{t-1} + \Pi_{22} g_t + (1-\theta L) v_{2t}$$

and the O.E.P. (4.13) now becomes

$$(4.23) \qquad \hat{p}_t = \frac{1-Q}{1-QL} p_{t-1}$$

Note that (4.23) also coincides with an adaptive expectations scheme; of course it is always possible that an ad hoc proxy will coincide with the O.E.P. Using a series from the O.E.P. as a proxy for the R.E. in this case is tantamount to estimating (4.11) and (4.12) (with $Q$ in place of $\delta$ there) subject to the within and cross equation restrictions that the ratio of the coefficient on $g_t$ to that on $g_{t-1}$ in both equations be equal to $-Q$, the coefficient of the M.A. error. A glance at (4.21) and (4.22) shows this restriction to be false and so inconsistent estimates result. Again, analytical expressions for asymptotic values of the estimates are uninformative but some numerical results on these are discussed in section 6.

To reiterate then even if the extrapolative proxy used coincides with the O.E.P. inconsistent estimates still result.

It must be noted at this stage that whilst parameter estimates from the reduced form of (4.3) and (4.4) will be inconsistent when the O.E.P. proxy is used the multipliers $\delta p_t/\delta \hat{g}_t$ and $\delta y_t/\delta \hat{g}_t$ which are the responses of $p_t$ and $y_t$ to that part of a change in $g_t$ that is predictable at t-1 will be correct asymptotically. This however is a peculiarity arising from the fact that in our simple model there is only one source of assymptotic bias namely the correlation between g and the augmented errors in (4.15) and (4.16) (the $\zeta_j$). When there is more than one source of bias, that is when the model involves 2 or more predetermined or exogenous variables, then nothing in general may be said about the asymptotic bias in these multipliers. The intuition for this result is weak but a proof is made in Appendix A.

## 5. A dynamic extension

To make our simple model more general consider the dynamic extension

$$(5.1)\qquad \begin{bmatrix} -1 & \beta_{12} \\ \beta_{21} & -1 \end{bmatrix} \begin{bmatrix} p_t \\ y_t \end{bmatrix} + \begin{bmatrix} \alpha & 0 & \psi(L) \\ 0 & \gamma & 0 \end{bmatrix} \begin{bmatrix} p_t^e \\ g_t \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where $\psi(L)$ is of order $r$ and has leading coefficient of $\psi_1$. The dynamics in the first equation of (5.1) mimic those in Anderson's condensed model equation referred to above.

The second equation in (5.1) is now overidentified and structural estimates may be obtained through the reduced form if we incorporate the $r+1$ cross equation restrictions

$$\frac{\rho_{2i}}{\rho_{1i}} = \frac{\Pi_{21}}{\Pi_{11}}$$

where the $\Pi$'s are the reduced form coefficients as above and the $\rho_{2i}$, $\rho_{1i}$ are the reduced form coefficients on $y_{t-1-i}$ in the second and first equation, respectively. This requires computationally expensive FIML estimation and when confronted with this problem the modeller almost invariably resorts to 2SLS structural estimation.[1/] We now consider then 2SLS estimation of (5.1) incorporating an ad hoc scheme for expectations under a maintained hypothesis of R.E.'s.

The maintained structure under the general ARMA(p, q) process for $g$ may now be written as

---

1/ Although over identification is a major reason for the use of 2SLS, nonlinearity also leads to its use in a 'nonlinear 2SLS procedure' (e.g. see Fair, 1976 ).

(5.2) $$p_t = \beta_{12}y_t + \frac{\alpha\Pi_{12}}{1-\Pi_{11}} L^{-1}\left[1 - \frac{\phi(L)}{\theta(L)}\right]g_{t-1} + \frac{\psi(L)}{1-\Pi_{11}} y_{t-1} + u_{1t}$$

(5.3) $$y_t = \beta_{21}p_t + \gamma g_t + u_{2t}$$

A simple transformation of (5.2) yields a representation in terms of a finite number of right hand side variables.

(5.4) $$\theta(L)p_t = \beta_{12}y_t + \left\{L^{-1}[\theta(L)-1] + \frac{\theta(L)\psi(L)}{1 - \Pi_{11}}\right\} y_{t-1}$$
$$+ \frac{\alpha\Pi_{12}}{1-\Pi_{11}} L^{-1}[\theta(L) - \phi(L)]g_{t-1} + \theta(L)u_{1t}$$

$p_t$ is now in terms of $q$ own lagged values and $\max(p,q)$ and $qr$ lagged values of $g_t$ and $y_t$ respectively.

Employing the scheme in (2.3) we would estimate by O.L.S.

(5.5) $$p_t = b_{12}\tilde{y}_t + a \sum_{i=1}^{n} \delta_i p_{t-i} + S(L)y_{t-1} + z_{1t}$$

(5.6) $$y_t = b_{21}\tilde{p}_t + c\, g_t + z_{2t}$$

where ~ denotes a prediction from an unrestricted reduced form regression (the first stage of 2SLS). The first thing to note is that the set of instruments used in the first stage would be incomplete as simple comparison of (5.3) and (5.4) with (5.5) and (5.6) shows. If the degrees of the lag polynomials in (5.4) are large (greater than one, say) then all estimates are likely to be very inefficient. Secondly the estimates in (5.5) will be inconsistent due to omitted variables and failure to take account of the M.A. error process of (5.4). Those in (5.6) will however be consistent. If we invoked adaptive expectations we would estimate (5.6) and

$$(5.7) \qquad p_t = b_{12}\tilde{y}_t + [a(1-\delta)+\delta]p_{t-1} + [S(L)(1-\delta L)-b_{12}\delta]y_{t-1} + (1-\delta L)z_{1t}$$

Even if the restrictions in (5.7) were not imposed, it will always exclude variables (at least one lagged value of $g_t$) and ignore the M.A. error in (5.4) and so will provide inconsistent estimates. In addition when the estimates are 'unscrambled' using the restrictions in (5.7) the structural parameters will be wrongly identified. Again estimates of (5.6) will be consistent but inefficient, the degree of efficiency depending on the 'completeness' of the set of instruments used (the included predetermined and exogenous variables). Finally we may note that if the O.E.P. for this model is used (the O.E.P. in this case is complicated and so its derivation is neglected to Appendix B) then we would estimate (5.6) and

$$(5.8) \qquad p_t = b_{12}\tilde{y}_t + a\hat{p}_t + S(L)y_{t-1} + z_{1t}$$

Our instruments for the first stage are now $\hat{p}_t$ , $g_t$ and r past values of $y_t$ and these provide our $\tilde{y}_t$ and $\tilde{p}_t$ . Note however that substitution of an O.E.P. for the R.E. using (4.14) augments the error term in (5.2) by

$\zeta'_{1t} - v_{1t}$ (where $\zeta'_{1t}$ has the same form as $\zeta_{1t}$ in (4.16) but with $\beta_{21}$ in place of $\Pi_{11}$ ).

This error as noted earlier includes $\varepsilon$ and $\mu$ making it correlated with all current and past values of g so now not only is the regressor $g_t$ correlated with the error in (5.8) but so is $\tilde{y}_t$ .

In fact because $g$ is also correlated with the reduced form error (seen clearly if the reduced form is written in terms of $\hat{p}_t$, $y_{t-1-i}$ (i=1 to r) and $g_t$) then O.L.S. gives inconsistent and biased reduced form predictions for $\tilde{y}_t$ and $\tilde{p}_t$ for the second stage of 2SLS. Again then all estimates are likely to be highly inefficient and those in (5.8) will be inconsistent. Obviously the above results carry over to LIML and 3SLS estimation although both are rarely used. In general then using 2SLS results in inconsistency of the estimates in the equation(s) containing expectations. Other estimates although consistent, are likely to be highly inefficient.

6. Some numerical illustrations

The previous analysis has shown that following the Fair-Anderson method will yield asymptotically biased estimates and in general these will give rise to biased multipliers. However, analytical expressions for these biases in terms of the model's coefficients are generally uninformative about their direction and magnitude even in the simplest of models. This section then, provides some numerical illustrations, deriving multipliers using estimates based on ad hoc expectations mechanisms under alternative settings for the model's parameters. In particular we consider estimation of the static model in section 4 and its dynamic counterpart in section 5 (this latter has the degree of $\psi(L)$ set equal to zero) incorporating, in turn, lagged variable and O.E.P. proxies for the R.E. Two types of multiplier for each variable are reported for the static model; the Ma's, representing the current response to anticipated changes in $g_t$ which are given as

$$Ma^p = \frac{\delta p_t}{\delta \hat{g}_t} = \frac{\Pi_{12}}{1-\Pi_{11}}$$

and

$$Ma^y = \frac{\delta y_t}{\delta \hat{g}_t} = \frac{\Pi_{12}\Pi_{21}}{1-\Pi_{11}} + \Pi_{22}$$

and the Mu's, the current responses to unanticipated changes in $g_t$, given as

$$Mu^p = \frac{\delta p_t}{\delta \varepsilon_t} = \Pi_{12}$$

and

$$Mu^y = \frac{\delta y_t}{\delta \varepsilon_t} = \Pi_{22}$$

These are also reported for the dynamic model (they have the same analytical form as in the static model) and in addition to these the current responses to unanticipated and anticipated changes in $g_{t-1}$ are tabulated. (These latter are obviously zero in the static model).

A glance at Table 6.1 (which shows the results for the static model) shows that although reduced form parameters obtained using an O.E.P. proxy are seriously inconsistent (biases of up to 200% are reported and some estimates are wrongly signed), the Ma's are consistent and this as noted earlier and as is proved in the Appendix is a peculiarity specific to our simple model. The Mu's however show serious bias and occasionally the wrong sign. For example in set (b), a value for $Mu^p$ is given as -0.436 for the lagged variable proxy and -0.264 for the O.E.P. when its actual value is 0.7. Note that the O.E.P. proxy provides slightly better estimates in most cases although the advantage over its lagged variable counterpart is often marginal.

The impact multipliers for the dynamic model reported in Table 6.2 show similar order of magnitude biases to their counterparts in the static model with the Ma's no longer being consistent. The 'one lag' dynamic multipliers , $\delta y_t/\delta \hat{g}_{t-1}$, $\delta p_t/\delta \hat{g}_{t-1}$ , $\delta y_t/\delta \varepsilon_{t-1}$ and $\delta p_t/\delta \varepsilon_{t-1}$ are the most seriously biased out of all the multipliers reported,most of them having the wrong sign. In fact this result emerged fairly consistently from the sets of parameter values that we used (only a small sample of our numerical results are presented in this paper). However, there seems to be no firm qualitative conclusions on signs or magnitudes of biases arising from the use of the two proxies except perhaps that use of an O.E.P. is marginally better than the simple lagged variable proxy. This supports our contention that results from numerical experiments using the Fair-Anderson method are likely to be arbitrary and uninformative about policy responses from a rational model.

Finally we must note that we have discussed large sample (asymptotic) properties only. In small samples the results may be even worse since in 2SLS the linear combinations of instruments used in the first stage although valid for equations not containing expectations terms are poorly correlated with the endogenous variables and so will give rise to highly inefficient estimates.

Table 6.1

O.L.S. estimates from (4.3) and (4.4) obtained using ad hoc proxies for the R.E.

| | Parameter/Multiplier | Actual Value | Lagged Variable Proxy | O.E.P. Proxy |
|---|---|---|---|---|
| (a) | $\alpha$ | 1.000 | 0.470 | 1.076 |
| | $\beta_{12}$ | 0.666 | -0.685 | 0.178 |
| | $\beta_{21}$ | 0.500 | 0.500 | 0.498 |
| | $\gamma$ | 0.400 | 0.400 | 0.391 |
| | $\Pi_{11}$ | 1.500 | 0.350 | 1.094 |
| | $\Pi_{12}$ | 0.400 | -0.204 | 0.076 |
| | $\Pi_{21}$ | 0.750 | 0.175 | 0.545 |
| | $\Pi_{22}$ | 0.600 | 0.298 | 0.429 |
| | $Ma^y$ | 0.000 | 0.243 | 0.003 |
| | $Ma^p$ | -0.800 | -0.313 | -0.813 |
| | $Mu^y$ | 0.600 | 0.298 | 0.429 |
| | $Mu^p$ | 0.400 | -0.204 | 0.076 |
| | g process :- $(1-L)\ g_t = (1-0.8L)\varepsilon_t$ | | | |
| (b) | $\alpha$ | 1.000 | 0.853 | 1.023 |
| | $\beta_{21}$ | 1.000 | 1.000 | 1.000 |
| | $\beta_{12}$ | 0.500 | -1.658 | -0.640 |
| | $\gamma$ | 0.700 | 0.699 | 0.680 |
| | $\Pi_{11}$ | 2.000 | 0.321 | 0.624 |
| | $\Pi_{12}$ | 0.700 | -0.436 | -0.264 |
| | $\Pi_{21}$ | 2.000 | 0.321 | 0.624 |
| | $\Pi_{22}$ | 1.400 | 0.263 | 0.4123 |
| | $Ma^y$ | 0.000 | 0.060 | -0.026 |
| | $Ma^p$ | -0.700 | -0.640 | -0.700 |
| | $Mu^y$ | 1.400 | 0.263 | 0.412 |
| | $Mu^p$ | 0.700 | -0.436 | -0.264 |
| | g process :- $(1-L)\ g_t = (1-0.3L)\varepsilon_t$ | | | |

Disturbance specification for all experiments:-

$\mathrm{cov}(u_1,u_2) = 0$ and $\mathrm{var}(u_1) = \mathrm{var}\ (u_2) = \sigma^2_\varepsilon$

Table 6.2

2SLS estimates from (5.1)(with the degree of $\psi(L)$ set equal to zero) obtained using ad hoc proxies for the R.E.

| | Parameter/Multiplier | Actual Value | Lagged Variable Proxy | O.E.P. Proxy |
|---|---|---|---|---|
| (a) | $\alpha$ | 1.000 | 0.258 | 0.043 |
| | $\beta_{12}$ | 0.400 | 0.112 | 0.031 |
| | $\beta_{21}$ | 1.000 | 1.005 | 1.000 |
| | $\gamma$ | 0.800 | 0.801 | 0.800 |
| | $\psi$ | 0.300 | -0.877 | -0.734 |
| | $\Pi_{11}$ | 1.666 | 0.290 | 0.044 |
| | $\Pi_{12}$ | 0.533 | 0.101 | 0.026 |
| | $\rho_1$ | 0.500 | -0.988 | -0.757 |
| | $\Pi_{21}$ | 1.666 | 0.448 | 0.080 |
| | $\Pi_{22}$ | 1.333 | 0.902 | 0.828 |
| | $\rho_2$ | 0.500 | -1.008 | -0.764 |
| | $Ma^y$ | 0.000 | 0.966 | 0.830 |
| | $Ma^p$ | -0.800 | 0.142 | 0.027 |
| | $Mu^y$ | 1.666 | 0.902 | 0.828 |
| | $Mu^p$ | 0.533 | 0.101 | 0.026 |
| | $MaD^y = \rho_2 \ Ma^y$ | 0.000 | -0.974 | -0.634 |
| | $MaD^p = \rho_1 \ Ma^p$ | -0.400 | -0.140 | -0.020 |
| | $MuD^y = \rho_2 \ Mu^y$ | 0.833 | -0.909 | -0.633 |
| | $MuD^p = \rho_1 \ Mu^p$ | 0.266 | -0.099 | -0.020 |

g process :- $(1-0.8L)\ g_t = \varepsilon_t$

Table 6.2 (contd)

| | Parameter/Multiplier | Actual Value | Lagged Variable Proxy | O.E.P. Proxy |
|---|---|---|---|---|
| (b) | $\alpha$ | 1.000 | 0.128 | 0.281 |
| | $\beta_{12}$ | 0.666 | -0.166 | -0.074 |
| | $\beta_{21}$ | 0.500 | 0.500 | 0.500 |
| | $\gamma$ | 0.400 | 0.400 | 0.400 |
| | $\psi$ | 0.600 | -0.266 | -1.401 |
| | $\Pi_{11}$ | 1.500 | 0.118 | 0.271 |
| | $\Pi_{12}$ | 0.400 | -0.060 | -0.028 |
| | $\rho_1$ | 0.900 | -2.083 | -1.351 |
| | $\Pi_{21}$ | 0.750 | 0.091 | 0.142 |
| | $\Pi_{22}$ | 0.600 | 0.360 | 0.380 |
| | $\rho_2$ | 0.450 | -1.090 | -0.671 |
| | $Ma^y$ | 0.000 | 0.354 | 0.375 |
| | $Ma^p$ | -0.400 | -0.068 | 0.038 |
| | $Mu^y$ | 0.600 | 0.360 | 0.380 |
| | $Mu^p$ | 0.400 | -0.060 | -0.028 |
| | $MaD^y$ | -0.160 | -0.386 | -0.252 |
| | $MaD^p$ | -0.360 | 0.142 | -0.051 |
| | $MuD^y$ | 0.270 | -0.392 | -0.255 |
| | $MuD^p$ | 0.360 | 0.125 | 0.038 |

g process :- $(1-0.8L)g_t = \varepsilon_t$

## 7. Further properties of the method

The preceeding sections show that to impose R.E.'s during simulation requires a set of parameter estimates obtained under the same maintained hypothesis, and the set obtained using an ad hoc expectations hypothesis is inadequate. This section exposes further properties of the method assuming that a set of parameter estimates has been obtained under the correct hypothesis of R.E.'s.

### 7.1 Properties associated with Lucas' critique

Because an R.E. differs from the actual outcome only by an innovation uncorrelated with the 'past' as contained in the information set then substitution of the actual value for the R.E. will ensure immunity from the structural variation noted by Lucas (1976), since no varying parameters associated with optimal prediction enter the simulation.

To make this clear recall the model as in (3.5) and (3.6)

(3.5) $$p_t = \Pi_{11} p_t^e + \Pi_{12} g_t + v_{1t}$$

(3.6) $$y_t = \Pi_{21} p_t^e + \Pi_{22} g_t + v_{2t}$$

and recall that under R.E.'s

(7.1) $$p_t^e = p_t^* = \frac{\Pi_{12}}{1-\Pi_{11}} \hat{g}_t = \frac{\Pi_{12}}{1-\Pi_{11}} L^{-1}\left[1 - \frac{\phi(L)}{\theta(L)}\right] g_{t-1}$$

Substitution of (7.1) into (3.5) and (3.6) gives a reduced form in terms of observables which includes the parameters of the g-process (the $\phi$'s and the $\theta$'s ). Changes in the latter bring about changes in this reduced form and these must be accounted for when such changes are simulated. Making expectations consistent with the predictions of the model we would simulate ( in deterministic simulation)

$$(7.2) \qquad p_t^s = \frac{\Pi_{12}}{1-\Pi_{11}} g_t^s$$

$$(7.3) \qquad y_t^s = \frac{\Pi_{21}\,\Pi_{12}}{1 - \Pi_{11}} g_t^s$$

and changes in the g-process do not raise a problem during simulation. In effect then the method keeps separate the parameters of the g-process from those of the economic structure (a distinction drawn by Wallis, 1980 ) allowing policy simulation to proceed in the traditional fashion. This feature of the method combined with the removal of the need to solve the model to obtain an expression for the R.E. form its most attractive properties.

7.2 Properties associated with deterministic simulation

The substitution of an actual outcome for expectations in deterministic simulations, in which structural errors are set to zero, to obtain a 'mean path solution' for the endogenous variables is an approximation when the actual exogenous outcomes and not their expected values are used. (In dynamic models both the mean solution and the expected values are conditional on the initial values). In effect, this makes current shocks in the exogenous processes enter expectations and become perfectly anticipated. It is claimed that this oversight is unimportant. However, much attention in current economics focusses on the effects of shocks in exogenous processes on the short run behaviour of real variables,(such shocks together with structural disturbances forming the mistakes in R.E.'s.) For example Lucas (1976) discusses the effects of increasing the one step ahead error variance (decreasing the 'predictability') of stochastic policy rules, money supply in his case, and an open loop low 'noise' money supply rule is recommended by Sargent and Wallace (1975) to stabilise output because it is easily predictable. However, the neglect of unpredictable elements in exogenous processes can be overcome by building predictors for them using multivariate time series methods and this can provide mean paths for the exogenous variables (conditioned on information at the first period of simulation) for deterministic simulations.

### 7.3 Stochastic simulation and problems associated with perfect foresight in neutrality models

Even were this undertaken there are certain classes of models for which mean paths obtained from deterministic simulations, are very uninteresting and uninformative from a policy analysis point of view. For such models structural disturbances must be added to all the equations and actual exogenous variable values (as opposed to predicted or mean values) used i.e. stochastic simulation must be undertaken. In this situation setting outcomes equal to expectations is imposing not R.E.'s but perfect foresight since outcomes include all current stochastic disturbances.

One class of models that requires stochastic simulation and for which the imposition of perfect foresight instead of R.E.'s has very serious consequences is the so called neutrality class of Lucas (1972) and Sargent and Wallace (1975). Broadly speaking these models are built to give a steady state in which the level of real output is independent of the money supply. Typically they have as a central feature a Lucas supply function determining output, and a quantity theory equation or a demand for money equation determining the price level. The former has real output deviating from a natural level only through the effect of current and past errors in predicting the price level. If expectations are rational these errors are unpredictable and consist of innovations in the exogenous variables and in the structure, both being uncorrelated with past events contained in the information set. In effect then mistakes in expectations consisting of these innovations drive a trade cycle, and if they are ignored by invoking perfect foresight then output will not deviate from its natural level for the period of simulation. As an illustration consider the simple neutrality model

$$(7.4) \qquad y_t = y^n + \sum_{i=0}^{m} \gamma_i (p_{t-i} - p^*_{t-i}) + u_{1t}$$

$$(7.5) \qquad p_t = V + m_t - y_t + u_{2t}$$

where $y^n$ is a natural level of output. The reduced form is

$$(7.6) \qquad y_t = y^n + (1+\gamma_0)^{-1} \sum_{i=0}^{m} \gamma_i (\varepsilon_{t-i} + u_{2t-i} - u_{1t-i}) + u_{1t}$$

$$(7.7) \qquad p_t = V + m_t - y^n - (1+\gamma_0)^{-1} \sum_{i=0}^{m} \gamma_i (\varepsilon_{t-i} + u_{2t-i} - u_{1t-i}) + u_{2t} - u_{1t}$$

$$p_t^* = p_t^s - \eta_t^s = p_t^s - \varepsilon_t^s + \gamma_0(1+\gamma_0)^{-1}(\varepsilon_t^s + u_{2t}^s - u_{1t}^s)$$

Because this substitution is exact, traditional policy evaluation may proceed and in this particular example the question of the optimal money supply rule to adopt to stabilise output may be resolved. This would normally be undertaken using a quadratic loss minimisation algorithm which would minimise an objective (say)

$$Q = \sum_{i=1}^{K} (y_{t+i} - y_{t+i}^*)^2$$

where $y_{t+i}^*$ is the desired or target level for income at $t + i$, subject to the economic model above. The algorithm would iterate on money supply paths until a minimum value for $Q$ was achieved.

## 8. Summary and Conclusion

Reading the abstract to Anderson's paper (op. cit. p. 67) we may be led to believe that the paper provides a method for deriving policy multipliers from an existing non rational model when expectations are rational. The main purpose of this paper, however, has been to show that multipliers derived using the Fair-Anderson methodology are likely to be seriously biased and so rational policy responses may not be extracted from non rational models. In numerical experiments using simple macromodels asymptotic multiplier biases of up to 200% are reported, with no consistent qualitative results on sign or magnitude emerging. This gives quantitative support to our claim that numerical values of policy responses so derived are likely to be arbitrary.

The simulation concept itself however is very useful provided that a consistent set of estimates is used and provided that 'mistakes' in expectations are reintroduced into the simulation procedure, where stochastic simulation is required. This latter point is especially relevant for policy analysis in "neutrality" models where deterministic simulation is uninformative.

Finally the most attractive feature of the algorithm is that it allows policy evaluation to proceed in a traditional manner because making expectations consistent with the predictions of a macro model (one which is consistently estimated of course) has the effect of separating the economic structure from the "structure" of the exogenous processes and thus removes the source of structural variation referred to by Lucas, namely these processes themselves.

# Appendix A

Expressions for asymptotic estimates from (4.9) and (4.10) with (n = 1).

Using the fact that

$$p_t^* = p_t - \Pi_{12}\varepsilon_t - v_{1t}$$

we may write (4.3) and (4.4) as

$$\text{(A1)} \qquad p_t = (K_1 + \Pi_{12})g_t - K_1\varepsilon_t + v_{1t}$$

$$\text{(A2)} \qquad y_t = (K_2 + \Pi_{22})g_t - K_2\varepsilon_t + v_{2t}$$

where $K_j = \Pi_{j1}\Pi_{12}/(1 - \Pi_{11})$ . O.L.S. estimation of (4.9) and (4.10), (n = 1) gives

$$\text{(A3)} \qquad \begin{aligned} \text{plim}\begin{bmatrix} p_{11} \\ p_{12} \end{bmatrix} &= \frac{1}{\Delta}\begin{bmatrix} \sigma_{gg}\sigma_{pp}(1) - \sigma_{gp}(1)\sigma_{gp} \\ \sigma_{pp}\sigma_{gp} - \sigma_{gp}(1)\sigma_{pp}(1) \end{bmatrix}, \\ \text{plim}\begin{bmatrix} p_{21} \\ p_{22} \end{bmatrix} &= \frac{1}{\Delta}\begin{bmatrix} \sigma_{gg}\sigma_{yp}(1) - \sigma_{gp}(1)\sigma_{gy} \\ \sigma_{pp}\sigma_{gy} - \sigma_{yp}(1)\sigma_{gp}(1) \end{bmatrix}, \end{aligned}$$

where $\Delta = \sigma_{gg}\sigma_{pp} - (\sigma_{gp}(1))^2$ and $\sigma$ denotes a data moment. Thus, for example

$$\sigma_{pp}(1) = \lim_{T\to\infty}(T)^{-1}\sum_{1}^{T} p_t p_{t-1}\ , \quad \sigma_{gp} = \lim_{T\to\infty}(T)^{-1}\sum_{1}^{T} p_t g_t\ .$$

Appendix B

Derivation of the O.E.P for the static and dynamic models.

Substituting

$$g_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t$$

in (A1) and multiplying through by $\phi(L)$ gives

$$\text{(B1)} \quad \phi(L)p_t = \left[(K_1 + \Pi_{12})\ \theta(L) - K_1\phi(L)\right] \varepsilon_t + \phi(L)v_{1t}$$

The right-hand side of (B1) has a MA representation in terms of a single innovation $\mu_t$ (Granger and Morris, 1976),

$$\text{(B2)} \quad \phi(L)\ p_t = Q(L)\mu_t\ ,$$

where $Q(L)$, of order max $(p,q)$, has leading coefficient of one, and the remaining coefficients and the variance of $\mu$ are obtained by canonical factorisation of the autocovariance function of the right-hand side of (B1).

The one-step ahead predictor then comes directly from (B2) as

$$\text{(B3)} \quad \hat{p}_t = L^{-1} \left[ 1 - \frac{\phi(L)}{Q(L)} \right] p_{t-1}$$

Derivation for the dynamic model in (5.1) is similar though more complicated. A form for $p_t$ and $y_t$ in terms of $g_t$ and $y_{t-1}$ is

$$\text{(B4)} \quad p_t = (K_1 + \Pi_{12})\ g_t + \frac{\rho_1(L)}{1-\Pi_{11}}\ y_{t-1} - K_1\varepsilon_t + v_{1t}$$

$$\text{(B5)} \quad y_t = \frac{(K_2 + \Pi_{22})}{(1 - \rho_1(L)K_3L)} g_t + \frac{V_{2t} - K_2\varepsilon_t}{(1 - \rho_1(L)K_3L)}$$

where $K_3 = \frac{\Pi_{21}}{\Pi_{11}(1-\Pi_{11})}$ and $\rho_1$ (L) is the reduced form lag polynomial on $y_{t-1}$ in the equation for $p_t$ from (5.1).

Lagging (B5) one period, substituting this in (B4) substituting for $g_t$ throughout as above and multiplying through by lag polynomial terms gives an ARMA ( (ŗ + 1) p, m) representation.

$$( 1 - \rho_1(L)K_3L ) \; \phi(L)p_t = Q_D(L) \; \mu_t$$

where $Q_D$ has leading coefficient of one and is of order $m = \left[(r+1)\max(p,q)\right]$ .

To show that the error in the univariate representation is correlated with all predetermined and exogenous variables in our static model rewrite (B2) as

$$\mu_t = \frac{\phi(L)}{Q(L)} p_t$$

Now substituting from (B1) for $p_t$ gives

$$\text{(B6)} \quad \mu_t = Q(L)^{-1} \left[(K_1 + \Pi_{12}) \; \theta(L) - K_1 \; \phi(L) \right] \varepsilon_t + Q(L)^{-1}\phi(L)v_{1t}$$

Expanding $Q(L)^{-1}$ gives the representation in terms of past ε's and v's as

$$\text{(B7)} \quad \mu_t = (\Pi_{12} \; \varepsilon_t + V_{it}) + ( \sum_{i=1}^{\infty} \omega_i \; \varepsilon_{t-1} + \sum_{i=1}^{\infty} \eta_i \, V_{1t}{-i} )$$

The first term in brackets is simply the error in the R.E. The second term is the source of correlation of μ with all past values of p and y (through the v's and ε's), and g (through the ε's).

# APPENDIX C

Proof of the consistency of $M_a^p$ and $M_a^y$ from (4.15) and (4.16)

Asymptotic biases in the $p_{ij}$ from (4.15) and (4.16) are given as

$$\text{(C1)} \quad \begin{bmatrix} B_{i1} \\ B_{i2} \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \sigma_{gg} & -\sigma_{g\hat{p}} \\ -\sigma_{g\hat{p}} & \sigma_{\hat{p}\hat{p}} \end{bmatrix} \begin{bmatrix} 0 \\ \sigma_g \zeta_j \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} -\sigma_{g\hat{p}} \sigma_g \zeta_j \\ \sigma_{\hat{p}\hat{p}} \sigma_g \zeta_j \end{bmatrix}$$

where $\Delta = \sigma_{gg}\sigma_{\hat{p}\hat{p}} - (\sigma_{g\hat{p}})^2$ and where $B_{ij}$ is the inconsistency in $p_{ij}$.

We need to prove

$$\text{(C2)} \quad Ma^p = \frac{\Pi_{12}}{1-\Pi_{11}} = \text{plim} \frac{p_{12}}{1-p_{11}} = \frac{\Pi_{12} + B_{12}}{1-\Pi_{11}-B_{11}}$$

and
$$\text{(C3)} \quad Ma^y = \Pi_{22} + \frac{\Pi_{21}\Pi_{11}}{1 - \Pi_{11}} = \text{plim}\left(p_{22} + \frac{p_{21}p_{11}}{1 - p_{11}}\right) = \Pi_{22}+B_{22}+ \frac{(\Pi_{21}+B_{21})(\Pi_{11}+B_{11})}{1 - \Pi_{11} - B_{11}}$$

Rearranging we see that (C2) is satisfied if

$$\frac{B_{12}}{B_{11}} = \frac{-\Pi_{12}}{1-\Pi_{11}}$$

We know from (C1) that

$$\text{(C4)} \quad \frac{B_{i2}}{B_{i1}} = \frac{-\sigma\hat{p}\hat{p}}{\sigma\hat{p}g}, \quad i = 1, 2 .$$

Taking moments from (A1) gives

$$\text{(C5)} \quad \sigma\hat{p}\hat{p} = \frac{\Pi_{12}}{1-\Pi_{11}} \quad \sigma\hat{p}g$$

Now noting that $\hat{p} = p - \mu$ and that $\hat{p}$ is uncorrelated with $\mu$ by construction so that

$$(C6)\quad \sigma\hat{p}\hat{p} = \sigma\hat{p}p$$

then (C5) and (C6) give the result

$$(C7)\quad \frac{B_{12}}{B_{11}} = \frac{-\sigma\hat{p}\hat{p}}{\sigma\hat{p}g} = \frac{-\Pi_{12}}{1-\Pi_{11}} \quad \text{Q.E.D.}$$

Proof of (C3) is more difficult requiring two extra conditions on the biases, which are given by the consistency of the ILS estimates of $\beta_{21}$ and $\gamma$. These conditions are

$$(C8)\quad \beta_{21} = \frac{\Pi_{21}}{\Pi_{11}} = \text{plim}\ \frac{P_{21}}{P_{11}} = \frac{\Pi_{21} + B_{21}}{\Pi_{11} + B_{11}}$$

$$\Rightarrow \frac{\Pi_{21}}{\Pi_{11}} = \frac{B_{21}}{B_{11}}$$

$$(C9)\quad \gamma = \Pi_{22} \frac{-\Pi_{12}\Pi_{21}}{\Pi_{11}} = \text{plim}(p_{22} - \frac{p_{12}p_{21}}{p_{11}}) = \Pi_{22}+B_{22}$$

$$- \frac{(\Pi_{12}+B_{12})(\Pi_{21}+B_{21})}{\Pi_{11} + B_{11}}$$

Using (C1) and (C7) gives

$$(C10)\quad \frac{B_{22}}{B_{21}} = \frac{B_{12}}{B_{11}} = -\frac{\Pi_{12}}{1-\Pi_{11}}$$

After rearranging (C3) and then substituting in this from (C8) and (C10) we can see that we need to prove

$$(C11)\quad \frac{B_{22}}{B_{11}}(1- \Pi_{11}) + \frac{\Pi_{21}\Pi_{12}}{\Pi_{11}} = 0$$

This is most easily verified by using (C8) and (C10) to get

$$\frac{B_{22}}{B_{11}} \Pi_{11} - \frac{\Pi_{12}\Pi_{21}}{1-\Pi_{11}} = 0$$

which on multiplication throughout by $1-\Pi_{11}/\Pi_{11}$ gives (C11) for the proof.

Note that if the zero element in the vector in the middle of C1 was replaced by a term representing a second source of correlation with $\zeta_j$ (this could be provided for example by use of a lagged variable proxy which, unlike $\hat{p}$, is correlated with the augmented errors) then the biases are weighted combinations of the two covariances with the $\zeta_j$ and nothing in general may be said about the form of the biases purely in relation to the reduced form parameters ( the $\Pi$'s ) and hence, the multipliers. Similarly in the n-regressor case, there will be only one zero element in the vector of covariances between the $\zeta_j$ and the regressors and again the biases will be weighted combinations of these covariances and will not in general satisfy the consistency conditions in (C2) and (C3). The numerical results for our dynamic model show biases in all multipliers and this bears witness to the latter point for the 3 regressor case.

Anderson, P.A. (1979), 'Rational expectations forecasts from non-rational models', Journal of Monetary Economics, 5, 67-80.

Fair, R.C. (1976), A Model of Macroeconomic Activity II. The Empirical Model, Cambridge, Mass: Ballinger.

——— (1979), 'An analysis of a macroeconometric model with rational expectations in the bond and stock markets', American Economic Review, 69, 539-52.

Granger, C.W.J. and Morris, M.J. (1976) 'Time series modelling and interpretation', Journal of the Royal Statistical Society A, 139, 246-257.

Lucas, R.E. Jr. (1972) 'Expectations and the neutrality of money' Journal of Economic Theory, 4, 103-24.

——— (1973) 'Some International evidence on output-inflation trade offs', American Economic Review, 63, 326-34.

——— (1976) 'Econometric policy evaluation: A critique', in The Phillips Curve and Labor Markets ed. by K, Brunner and A.H. Meltzer, Amsterdam: North-Holland, 19-46.

Nelson, C.R. (1975), 'Rational expectations and the estimation of econometric models', International Economic Review, 16, 555-61.

Prothero, D.L. and Wallis, K.F. (1976) 'Modelling macroeconomic time series (with discussion)' Journal of the Royal Statistical Society A, 139, 468-500.

Sargent, T.J. and Wallace, N.(1975) '"Rational" expectations, the optimal monetary instrument, and the optimal money supply rule', Journal of Political Economy, 83, 241-54.

Wallis, K.F. (1980) 'Econometric implications of the rational expectations hypothesis', Econometrica, 48, 49-73.