ON TWO METHODS FOR SOLVING AND ESTIMATING LINEAR SIMULTANEOUS EQUATIONS UNDER RATIONAL EXPECTATIONS

Gregory C. Chow

Philip J. Reny

Econometric Research Program
Research Memorandum No. 314

July 1984

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

ON TWO METHODS FOR SOLVING AND ESTIMATING LINEAR SIMULTANEOUS EQUATIONS UNDER RATIONAL EXPECTATIONS

Gregory C. Chow and Philip J. Reny

1. INTRODUCTION

This paper is concerned with solving and estimating linear simultaneous-equations models of the form

$$(1.1) \qquad By_t + A_1 y_{t-1} + \dots + A_p y_{t-p} + B_0 y_{t|t-1} + B_1 y_{t+1|t-1} + \dots + B_q y_{t+q|t-1} = \Gamma z_t + u_t$$

where y_t is a column vector of G endogenous variables, z_t is a vector of K exogenous variables, u_t is a vector of residuals and $y_{t+i|t-1}$ denotes the conditional expectation of y_{t+i} given all information up to the end of period t-1. The conditional expectations $y_{t+i|t-1}$ are assumed to be mathematical expectations computed from the same econometric model which generates (1.1) under the assumption of rational expectations. One consistent method for estimating the parameters of such models is given by McCallum [8]. Two other methods of estimation will be discussed in this paper.

A closely related problem to that of estimation is the solution of model (1.1). By solution is meant finding a stochastic mechanism to generate observations of y_t which is consistent with model (1.1). It is well-known that when expectations $y_{t+i|t-1}$ of future endogenous variables appear in a linear simultaneous-equations model, many solutions exist if the expectations are formed by the assumption of rational expectations. For example, see Gourieroux, Laffont and Montfort [6] for a treatment of the solutions of a univariate model with future expectations. Therefore, even if all parameters of (1.1) are known, an econometrician still faces the problem of finding a stochastic model which will generate future observations of y_t for the purpose of forecasting. Several approaches to finding a unique solution of (1.1) have been proposed and are discussed in Chow [2 , Chapter 11]. Two of the apparently promising approaches will be discussed in this paper, one by Chow [2] and the other by Dagli and Taylor [3]. Corresponding to each approach is a set of methods for estimating the parameters of model (1.1). Thus the problems of solving and estimating model (1.1) are closely related.

In section 2, Chow's method for solving the model (1.1) will be summarized. In section 3, Dagli and Taylor's method of solving and estimating (1.1) will be presented. In section 4 it is shown that any solution which is linear in y_t, z_t and u_t with constant coefficients is a special case of Chow's solution. Since the Dagli-Taylor solution is of this form, it is a special case of Chow's solution. In other words, by imposing certain restrictions on the parameters of the stochastic model proposed by Chow as solution to model (1.1), one obtains the solution of Dagli and Taylor. The last part of section 4 is devoted to deriving these restrictions explicitly. An important implication of this result is that if Chow's solution is accepted, the Dagli-Taylor method for estimating the parameters of (1.1) will be inconsistent unless the special restrictions imposed by their method happen to be valid. This and related issues will be discussed in section 5 which compares the two methods.

An illustrative model consisting of two simultaneous equations and involving the vector $y_{t+1|t-1}$ of two future endogenous variables is presented in section 6 for the purposes of demonstrating how the two methods work and of comparing them. When Chow's solution model [2] was presented, the methods of maximum likelihood and minimum distance were proposed for estimating its parameters, but no actual computations were reported. It is encouraging to find that these methods have worked well for the illustrative model. We can recommend the application of these methods to the estimation of simultaneous equations involving future expectations in econometric practice. The inconsistency of the Dagli-Taylor method is also illustrated in section 6.

2. CHOW'S METHOD OF SOLUTION

The method of Chow [2 , Chapter 11] for finding a solution model which is consistent with (1.1) but free of future expectations can be briefly summarized. To facilitate comparison with the Dagli-Taylor method to be presented in section 3, we follow Dagli and Taylor [3] by assuming that the residual u_t in (1.1) satisfies

$$u_t = \Delta(L)\varepsilon_{1t} \tag{2.1}$$

where ε_{1t} is an independent and identically distributed random vector with mean 0 and $\Delta(L)$ is a matrix polynomial in the lag operator L, with coefficients Δ_i, such notation being used throughout this paper. We also assume the exogenous variables z_t to satisfy

$$z_t = \Lambda(L)\varepsilon_{2t} \tag{2.2}$$

where ε_{2t} is an independent and identically distributed random vector with mean 0 and $\Lambda(L)$ is a matrix polynomial in the lag operator L. Economic agents described by the model (1.1) are assumed to know $\Lambda(L)$ and past realizations of the exogenous variables up to z_t, implying $z_{t|t-1} = z_t$, but not to know the future realizations $z_{t+1}, z_{t+2}, \ldots$. The random vector $\varepsilon_t' = (\varepsilon_{1t}', \varepsilon_{2t}')$ is assumed to be normally distributed with mean zero and covariance matrix Ω. The normalizations $\Lambda_0 = 0$, $\Lambda_1 = I$, and $\Delta_0 = I$ are applied .

The reduced-form of the structural equations (1.1) is

$$B^{-1}(By_t + A_1y_{t-1} + \ldots + A_py_{t-p} + B_0y_{t|t-1} + B_1y_{t+1|t-1} + \ldots + B_qy_{t+q|t-1}) \tag{2.3}$$

$$= B^{-1}\Gamma z_t + B^{-1}u_t \equiv B^{-1}\Gamma z_t + v_t$$

where we have defined $B^{-1}u_t = v_t$. We will assume that any solution y_t of (1.1) satisfies for m = 0,1,...,q and for some $R_0,\ldots,R_q$, $K_0,\ldots,K_{q-1}$

$$(2.4) \quad y_{t+m} - y_{t+m|t-1} = R_0 d_{t+m,t-1} + R_1 d_{t+m-1,t-1} + \cdots + R_m d_{t,t-1} + K_0 e_{t+m,t-1} + K_1 e_{t+m-1,t-1} + \cdots + K_{m-1} e_{t+1,t-1}$$

where

$$(2.5) \quad \begin{aligned} d_{t+m,t-1} &= u_{t+m} - u_{t+m|t-1} \\ e_{t+m,t-1} &= z_{t+m} - z_{t+m|t-1} \end{aligned}$$

and

$$K_{-1} \equiv 0$$

This assumption is motivated by the solution method of Chow [2, Chapter 11] which begins by assuming that any solution to (1.1) takes the form

$$(2.6) \quad y_t = R_0 u_t + R_1 u_{t-1} + \cdots + K_0 z_t + K_1 z_{t-1} + \cdots$$

for some $R_0, R_1, \ldots, K_0, K_1, \ldots$. Chow shows [2, pp. 356-357] that (2.6) implies (2.4). As pointed out by Evans and Honkapohja [4], however, the infinite MA representation (2.6) runs into problems of well-definedness in nonstationary situations. As a result, an assumption of the form (2.6) restricts the class of solutions to stationary ones. Since, in the present paper, we do not wish to restrict the class of solutions a priori (i.e. before observing the data), an assumption of the form (2.6) is inappropriate. We therefore make the weaker assumption (2.4) which allows both stationary and nonstationary solutions to (1.1). (One should note that the solution given in [Chow, 2, Chapter 11] is identical to that given here (equation 2.7). Thus, keeping in mind the slight modification of the Chow [2] solution method made above, the Chow [2] solution is valid in both stationary and nonstationary situations.)

Taking expectations of (2.3) given the information set I_{t-1} available at the end of period t-1 and subtracting the result from (2.3), we have

$$y_t - y_{t|t-1} = B^{-1}u_t - B^{-1}u_{t|t-1} .$$

Putting m = 0 in (2.4) yields $y_t - y_{t|t-1} = R_0 u_t - R_0 u_{t|t-1}$ implying $R_0 = B^{-1}$.

The solution for (1.1) is found by using (2.4) to substitute for all $y_{t+m|t-1}$ (m = 0,1,...,q) in (1.1). This amounts to replacing $y_{t+m|t-1}$ by y_{t+m} minus a weighted sum of $d_{t+j,t-1}$ (j = 0,...,m) and $e_{t+j,t-1}$ (j = 1,...,m). The result is

(2.7) $$By_t + A_1 y_{t-1} + \dots + A_p y_{t-p} + B_0 y_t + B_1 y_{t+1} + \dots + B_q y_{t+q}$$

$$= \Gamma z_t + u_t + \tilde{R}_0 d_{t,t-1} + \tilde{R}_1 d_{t+1,t-1} + \dots + \tilde{R}_q d_{t+q,t-1}$$

$$+ \tilde{K}_0 e_{t+1,t-1} + \tilde{K}_1 e_{t+2,t-1} + \dots + \tilde{K}_{q-1} e_{t+q,t-1}$$

where

$$\tilde{R}_m = \sum_{i=0}^{q-m} B_{m+i} R_i \qquad (m = 0,\dots,q)$$

$$\tilde{K}_m = \sum_{i=0}^{q-m-1} B_{m+1+i} K_i \qquad (m = 0,\dots,q-1)$$

with $R_0 = B^{-1}$ as shown previously.

To show that the solution model (2.7) implies the original model (1.1), we simply take expectations of (2.7) given the information I_{t-1} to yield

$$By_{t|t-1} + A_1 y_{t-1} + \dots + A_p y_{t-p} + B_0 y_{t|t-1} + B_1 y_{t+1|t-1} + \dots + B_q y_{t+q|t-1}$$

$$= \Gamma\, z_t + u_{t|t-1}$$

We also take expectations of (2.7 given I_{t+q-1} and subtract the result from (2.7) to obtain

$$B_{q\ t+q} - B_q y_{t+q|t+q-1} = B_q B^{-1} u_{t+q} - B_q B^{-1} u_{t+q|t+q-1}$$

or

$$y_t - y_{t|t-1} = B^{-1}(u_t - u_{t|t-1})$$

if B_q^{-1} exists. The case where B_q^{-1} does not exist is treated in Chow [2, pp. 357-358]. When this equation is used to substitute for $By_{t|t-1}$ in the previous equation, the original model (1.1) results. Note than in constructing the solution model (2.7), the matrices R_i $(i = 1,\ldots,q)$ and K_i $(i = 0,\ldots,q-1)$ can be arbitrary and (2.7) still implies (1.1). (Also, no assumption about stationarity was needed to obtain this solution.) The essence of Chow's solution is that it includes these additional parameters which characterize the multiple solutions of the original model (1.1). Given model (1.1) alone, one can arbitrarily choose the values of these parameters in the solution model (2.7) to generate different solutions of (1.1). Once the values of these parameters are fixed, the problem of multiple solutions for (1.1) is resolved. It is therefore proposed to estimate the values of these parameters empirically together with the parameters in (1.1). Methods of estimation have been discussed in Chow [2]. Two of the methods will be applied in section 6 to estimate an illustrative model.

3. DAGLI AND TAYLOR'S METHOD OF SOLUTION AND ESTIMATION

The method of Dagli and Taylor [3] for solving system (1.1) consists of the following five steps. First, take conditional expectation of (1.1) given information I_{t-1} to obtain a model of the expectations variables.

$$(3.1) \qquad B_q y_{t+q|t-1} + \dots + B_1 y_{t+1|t-1} + (B + B_0) y_{t|t-1} + A_1 y_{t-1|t-1} + \dots$$

$$+ A_p y_{t-p|t-1} = \Gamma z_{t|t-1} + u_{t|t-1} .$$

Denote by L_1 the special operator which decreases only the first time subscript of a variable $y_{t+m|t-1}$ by one without changing the conditioning time subscript, i.e., $L_1 y_{t+m|t-1} = y_{t+m-1|t-1}$. The ordinary lag operator L has the property $Ly_{t+m|t-1} = y_{t+m-1|t-2}$. Using the operator L_1, we rewrite (3.1) as

$$(3.2) \qquad (B_q L_1^{-q} + \dots + B_1 L_1^{-1} + (B + B_0) + A_1 L_1 + \dots + A_p L_1^p) y_{t|t-1}$$

$$= H(L_1) y_{t|t-1} = \Gamma z_{t|t-1} + u_{t|t-1}$$

where $H(L_1)$ is a matrix polynomial in the operator L_1 defined by the first line of (3.2).

Second, factor the polynomial H(L) in the form

$$(3.3) \qquad H(L\ ) = \Phi(L^{-1})\, \theta(L\ )$$

where

$$\Phi(L^{-1}) = I + \Phi_1 L^{-1} + \dots + \Phi_q L^{-q}$$

$$\theta(L\ ) = \theta_0 + \theta_1 L + \dots + \theta_p L^p .$$

and it is assumed that the determinantal polynomials of $\Phi(z)$ and $\theta(z)$ have all roots outside the unit circle.

Third, premultiply (3.2) by the inverse of $\Phi(L_1^{-1})$ to obtain

$$(3.4) \qquad \theta(L_1) y_{t|t-1} = [\Phi(L_1^{-1})]^{-1} \Gamma z_{t|t-1} + [\Phi(L_1^{-1})]^{-1} u_{t|t-1}$$

$$= \psi_2(L)\varepsilon_{2t} + \psi_1(L)\varepsilon_{1t}$$

where $\psi_1(L)$ and $\psi_2(L)$ are polynomials with only positive and zero powers of L and are obtained by using $[\Phi(L_1^{-1})]^{-1}$ and the assumptions (2.1) and (2.2) for u_t and z_t. Dagli and Taylor give more details. Equation (3.4) can be used to express $y_{t|t-1}$ as a function of lagged values of y_t and lagged disturbances

$$(3.5) \qquad y_{t|t-1} = -\theta_0^{-1}[\theta_1 y_{t-1} + \dots + \theta_p y_{t-p} - \psi_2(L)\varepsilon_{2t} - \psi_1(L)\varepsilon_{1t}]$$

Fourth, subtract equation (3.1) from equation (1.1) to obtain

$$(3.6) \qquad By_t = By_{t|t-1} + \varepsilon_{1t}$$

and substitute (3.5) for $y_{t|t-1}$ in (3.6) to yield

$$(3.7) \qquad By_t = -B\theta_0^{-1}[\theta_1 y_{t-1} + \dots + \theta_p y_{t-p}] + B\theta_0^{-1}\psi_2(L)\varepsilon_{2t} + [B\theta_0^{-1}\psi_1(L) + I]\varepsilon_{1t}$$

Fifth, substitute for ε_{2t} from (2.2) to obtain the solution model

$$(3.8) \qquad By_t = C(L)y_t + D(L)z_t + R(L)\varepsilon_{1t}$$

where

$$C(L) = B\theta_0^{-1}[\theta_0 - \theta(L)]$$

$$D(L) = B\theta_0^{-1}\psi_2(L)\Lambda^{-1}(L)$$

$$R(L) = B\theta_0^{-1}\psi_1(L) + I .$$

Equation (3.8) is a dynamic model free of expectations variables which is consistent with the original model (1.1). The parameters of (3.8) are functions of the parameters of (1.1). Dagli and Taylor have recommended applying maximum likelihood to (3.8) for estimating the parameters of (1.1).

4. DAGLI AND TAYLOR'S SOLUTION IS A SPECIAL CASE OF CHOW'S SOLUTION

In this section we prove that the Chow solution is the most general among the class of constant-coefficient solutions which are linear in y_t, z_t and u_t. As a result, we obtain as a corollary that the Dagli-Taylor (DT) solution (3.8) is a special case of the Chow solution (2.7).

Theorem: If for polynomials $D(L)$, $G(L)$, $Q(L)$ (containing only nonnegative powers of L) with $D(L) = D_0 + D_1L + \ldots + D_dL^d$, $D_0 \equiv I$ and $d < \infty$

(4.1) $$D(L)y_t = G(L)z_t + Q(L)u_t$$

is a linear constant-coefficient model satisfying (1.1), then (4.1) is a special case of the Chow solution in the following sense. Given any $\bar{y}_t$ satisfying (4.1), $\bar{y}_t$ also satisfies (2.7) when R_{i+1}, K_i, $i = 0,1,2,\ldots,q-1$ are appropriately restricted.

Proof: Define $D_q^*(L)$ by

$$D_q^*(L) \equiv D_0^* + D_1^*L + D_2^*L^2 + \ldots + D_q^*L^q$$

where $$\sum_{m=0}^{n} D_m^* D_{n-m} = 0 \quad \text{for } n = 1,2,\ldots,q$$

and $$D_0^* = I, \quad D_{d+i} \equiv 0 \quad i = 1,2,3,\ldots .$$

(Note: The D_i^* can be computed recursively; i.e. given $D_1^*,\ldots,D_K^*$ we have $D_{K+1}^* = -\sum_{m=0}^{K} D_m^* D_{K-m}$.)

Premultiplying (4.1) by $D_q^*(L)$ gives, for any $\bar{y}_t$ satisfying (4.1),

(4.2) $$\bar{y}_t + P(L)\bar{y}_{t-q-1} = D_q^*(L)G(L)z_t + D_q^*(L)Q(L)u_t$$

where P(L) is a polynomial consisting of only nonnegative powers of L and is of degree at most d-1. ($P(L) \equiv 0$ is possible.)

Advancing the time subscripts in (4.2) by m for $0 \leq m \leq q$ gives

$$(4.3) \qquad \bar{y}_{t+m} + P(L)\bar{y}_{t+(m-q)-1} = V(L)z_{t+m} + W(L)u_{t+m}$$

where $V(L) \equiv D_q^*(L)G(L)$ and $W(L) = D_q^*(L)Q(L)$.

Taking the expectation of (4.3) conditioned on information at date t-1 and subtracting the result from (4.3) gives for m = 1,2,...,q (since P(L) consists of only nonnegative powers of L)

$$(4.4) \qquad \bar{y}_{t+m} - \bar{y}_{t+m|t-1} = W_0(u_{t+m} - u_{t+m|t-1}) + W_1(u_{t+m-1} - u_{t+m-1|t-1}) + \dots + W_m(u_t - u_{t|t-1}) + V_0(z_{t+m} - z_{t+m|t-1}) + V_1(z_{t+m-1} - z_{t+m-1|t-1}) + \dots + V_{m-1}(z_{t+1} - z_{t+1|t-1})$$

and $\bar{y}_t - \bar{y}_{t|t-1} = W_0(u_t - u_{t|t-1})$.

Now, as shown in section 2, (1.1) implies

$$\bar{y}_t - \bar{y}_{t|t-1} = B^{-1}(u_t - u_{t|t-1})$$

so that $W_0 = B^{-1}$.

Using (4.4) to substitute for all $\bar{y}_{t+m|t-1}$ (m = 0,1,...,q) in (1.1) (as in section 2) implies that $\bar{y}_t$ satisfies (2.7) with $R_{i+1} = W_{i+1}$ and $K_i = V_i$ for i = 0,1,...,q-1.

□

<u>Remark</u>: Given any solution to (1.1) of the form (4.1) the proof of the theorem provides a method for deriving the restrictions which that solution

places on the free parameters of the Chow solution. This fact will be exploited below.

Corollary: The DT solution (3.8) is a special case of the Chow solution (2.7).

Proof: (3.8) is a special case of (4.1) in the theorem. □

We now derive explicitly the restrictions on the R_{i+1} and K_i, $i = 0,1,2, \ldots, q-1$ imposed by the DT method.

(3.8) can be written as

$$\text{(4.5)} \qquad \theta(L)y_t = \psi_2(L)\varepsilon_{2t} + [\psi_1(L) + \theta_0 B^{-1}]\varepsilon_{1t}$$

or, using (2.1) and (2.2),

$$\text{(4.6)} \qquad \theta(L)y_t = \psi_2(L)\Lambda^{-1}(L)z_t + [\psi_1(L) + \theta_0 B^{-1}]\Delta^{-1}(L)u_t \ .$$

Finally, the DT root assumption allows us to premultiply by $\theta^{-1}(L)$ giving

$$\text{(4.7)} \qquad y_t = \theta^{-1}(L)\psi_2(L)\Lambda^{-1}(L)z_t + \theta^{-1}(L)[\psi_1(L) + \theta_0 B^{-1}]\Delta^{-1}(L)u_t \ .$$

Now (4.7) is of the form (4.1) with

$$D(L) \equiv I$$
$$G(L) \equiv \theta^{-1}(L)\psi_2(L)\Lambda^{-1}(L)$$
$$\text{and} \quad Q(L) \equiv \theta^{-1}(L)[\psi_1(L) + \theta_0 B^{-1}]\Delta^{-1}(L) \ .$$

Accordingly, we easily obtain $D_q^*(L) = I$. So, using the proof of the theorem gives that y_t satisfies (2.7) with

(4.8) $$R_{i+1} = G_{i+1} \text{ and } K_i = Q_i , \quad i = 0,1,\ldots,q-1$$

(where $G(L) = G_0 + G_1L + G_2L^2 + \ldots$, $Q(L) = Q_0 + Q_1L + Q_2L^2 + \ldots$). Hence (4.8) provides explicitly the restrictions imposed on the free parameters of the Chow solution by the solution proposed by Dagli and Taylor.

5. COMPARISON OF THE TWO METHODS

As remarked in the introduction, in general we cannot hope to obtain a unique solution to (1.1) under the rational expectations assumption. As a result, in addition to the rational expectations hypothesis, stationarity and/or covergent expectations have often been imposed on the solution in order to obtain uniqueness. Recently there have been attempts to relax some of these ex ante restrictions and use instead empirical information to determine their appropriateness. Examples are the works of Flood and Garber [5] and Burmeister and Wall [1]. The reasoning here is simple enough. If we are sufficiently unsure about the validity of ex ante restrictions on the solution to (1.1) we ought not impose them, but rather we should test for them. Indeed we shall see below that imposing false restrictions can lead to serious problems with the resulting estimates.

Let us consider the problem faced by an econometrician of obtaining and estimating the reduced-form solution to (1.1). (We say "the reduced-form solution" since although (1.1) has a multiplicity of solutions we suppose that only one describes the data.) The only information available (other than data) is (1.1) and the knowledge that the solution is linear with constant coefficients. One available procedure would be to start with a linear constant-coefficient model $D(L)y_t = G(L)z_t + Q(L)u_t$ and use (1.1) to obtain the necessary restrictions on the parameters. This method was employed in part by McCallum [8] to obtain consistent parameter estimates, but in McCallum's method the number of lagged endogenous as well as exogenous variables to include in the reduced form was left unspecified and the cross-equation restrictions were not explicitly taken into account.

As we know from the theorem of section 4, the method given in section 2 provides the most general solution to (1.1) among the class of linear constant-coefficient solutions. Thus, our econometrician need not make any possibly erroneous assumptions concerning the nature of the solution to apply this method. Also, the method of section 2 is free of both shortcomings of the McCallum procedure.

As described in section 3, in order to obtain their solution Dagli and Taylor employ a factorization technique. In particular they factorize $H(L)$ as

$$H(L) = \Phi(L^{-1})\theta(L)$$

and assume that $\theta(L)$ and $\Phi(L^{-1})$ be of degrees p and q respectively and that $\theta(z)$ and $\Phi(z)$ have all roots outside the unit circle. Under these assumptions (which have previously been made in the literature; for instance see pp. 333-337 in Sargent [9] for a univariate example) they are able to obtain a unique solution to (1.1). It is not clear however, into which class these assumptions restrict the solution to (1.1). For example, one might suspect that these assumptions are equivalent to stationarity, but they are in fact stronger. Stationarity can be achieved without these assumptions. One sees this by noting that as long as the determinantal polynomial of $H(L) = \Phi(L^{-1})\theta(L)$ has all roots outside the unit circle any solution to (2.7) will be stationary. But if this is the case then $|\Phi(z)|$ has all roots inside the unit circle contrary to the DT assumption. Hence there are many stationary time series which satisfy (1.1) but are ruled out by the Dagli-Taylor assumptions. With this in mind one may question the reason (apart from their ability to induce a unique solution) behind the restrictions imposed on the roots of the determinantal polynomials of $\theta(L)$ and $\Phi(L^{-1})$. Indeed, not only have we discovered that the Dagli-Taylor assumption implies arbitrary restrictions on the free parameters of the more general Chow solution, but that this assumption can lead to inconsistent estimates of the structural parameters as well as unresolvable identification problems. The last two points are illustrated by the following simple example.

Consider the univariate model

$$y_t + \alpha y_{t+1|t-1} = \beta z_t + u_t \tag{5.1}$$

where $z_{t|t-1} = z_t$, $z_{t|t-2} = 0$, $E(z_t z_{t-i}) = 0$, $\forall\, i \geq 1$, and u_t is a white noise disturbance term uncorrelated with z_t. Using the DT solution method (which requires $|\alpha| < 1$ since $\theta(L) = 1$ and $\Phi(L^{-1}) = 1+\alpha L^{-1}$) we obtain the reduced form

$$y_t = \beta z_t + u_t \,. \tag{5.2}$$

Advancing the time subscripts in (5.2) by one and taking expectations as of date t-1 gives

(5.3) $$y_{t+1|t-1} = 0 .$$

Looking at (5.2) we notice immediately that the parameter α is absent and hence cannot be estimated. The reason for this is clear by observing (5.3). Combining (5.3) with (5.1) not only gives (5.2) but helps us realize that α is unidentifiable by the DT solution method.

The solution model for (5.1) obtained by the Chow method is

(5.4) $$y_t = -(1/\alpha)y_{t-1} + Kz_t + (\beta/\alpha)z_{t-1} + u_t + Ru_{t-1}$$

where R and K are arbitrary. As is readily seen this solution employs two parameters which characterize the multiple solutions associated with (5.1) in addition to the two structural parameters α and β. The restrictions of R and K implied by the DT solution are that $K = \beta$, $R = 1/\alpha$. Putting $K = R = 0$ and assuming $|\alpha| > 1$ (which is contrary to the DT assumption) we obtain from (5.4)

(5.5) $$y_t = \sum_{i=0}^{\infty} (-\frac{1}{\alpha})^i \left[u_{t-i} + (\beta/\alpha)z_{t-i-1} \right] .$$

To demonstrate the possibility of inconsistent estimates arising from the DT method we shall assume that the actual time series process satisfies (5.5) and hence (5.1). Estimating β from (5.2) via the DT solution gives

$$\hat{\beta} = \sum_{t=1}^{T} z_t y_t \Big/ \sum_{t=1}^{T} z_t^2 .$$

Using (5.5) and the assumptions on z_t and u_t above give

$$\text{plim}\ \hat{\beta} = 0 \neq \beta.$$

This example illustrates that the DT solution method can lead to inconsistent estimates if the DT root conditions are not met and to identification problems.

We have pointed out that the two methods for solving and estimating a system of linear simultaneous equations under rational expectations imply different assumptions about the parameters of the model. Specifically, the DT method implies a set of restrictions on the parameters of a more general model assumed by the Chow method. A choice between the two methods can be made by testing the restrictions implied by the DT method. To make such a choice, the specification test of Hausman (1978) can be applied. Let θ be a vector of parameters of the model (1.1), $\hat{\theta}^0$ be the maximum likelihood estimator of Dagli and Taylor, and $\hat{\theta}$ be the maximum likelihood estimator using the Chow solution. Under the assumption that the DT restrictions are correct, $\hat{\theta}^0$ is asymptotically efficient, but if the DT restrictions are not correct, $\hat{\theta}^0$ is inconsistent. When the DT restrictions are incorrect, the estimator $\hat{\theta}$ is still consistent. Let $\hat{q} = \hat{\theta} - \hat{\theta}^0$. The test statistic is

$$\hat{q}'(\mathrm{Cov}\ \hat{q})\hat{q}$$

where the covariance matrix Cov $\hat{q}$ can be estimated by

$$\mathrm{Cov}\ \hat{q} = \mathrm{Cov}\ \hat{\theta} - \mathrm{Cov}\ \hat{\theta}^0 .$$

Under the null hypothesis that the DT restrictions are correct, this statistic is asymptotically distributed as χ^2 with the number of degrees of freedom equal to the number of elements in the vector $\hat{q}$.

To illustrate the Hausman test we generated time series data (fifty observations) using (5.4) with $z_t \sim N(0,1)$ (treated as exogenous, $u_t \sim N(0,.25)$, $\beta = 1$, $R = K = 0$ and $\alpha = 2$ (so that $|\alpha| > 1$ and the DT root assumptions do not hold nor do their implied restrictions on R and K). Since (5.2) contains only one parameter (namely β), we have $\hat{q} = \hat{\beta}_c - \hat{\beta}_{DT}$ where $\hat{\beta}_c$ and $\hat{\beta}_{DT}$ are the estimates of β obtained from (5.4) and (5.2) respectively. Estimation of the parameters was performed using a full-information maximum likelihood technique (as in section 6 below). The results are as follows:

$$\hat{\beta}_c = 1.35 \qquad \text{var}(\hat{\beta}_c) = .169$$

$$\hat{\beta}_{DT} = -.117 \qquad \text{var}(\hat{\beta}_{DT}) = .0132 \; .$$

Hence the test statistic is

$$\chi^2 = (1.35 + .117)^2(.169 - .0132)^{-1} = 13.8 \; .$$

As a result of this test, we may reject, at the 99.5 percent confidence level, the null hypothesis that the DT assumptions hold. This is both comforting and not surprising since the example was designed to violate the DT assumptions. For this simple example, then, the Hausman test directs us toward the more parsimonious specification (5.4) whose complete set of estimates turn out to be

$$\hat{\alpha}_c = 2.31 \qquad \hat{\beta}_c = 1.35$$

$$\hat{K}_c = -.003 \qquad \hat{R}_c = .003$$

$$\hat{\sigma}^2_u = .18$$

6. ESTIMATING AN ILLUSTRATIVE MODEL

To illustrate how well the Chow method works and what results might be obtained by the DT method, we have employed data generated from a two-equation model to estimate its parameters using both techniques. The model takes the form

$$(6.1) \qquad By_t + y_{t|t-1} + B_1 y_{t+1|t-1} = \Gamma z_t + u_t$$

where the matrices B, B_1 and Γ are all 2x2, u_t and z_t are uncorrelated random disturbances with covariance matrices $\sigma_u^2 I$ and I respectively, z_t is treated as exogenous with $z_{t|t-1} = z_t$ and $z_{t|t-2} = u_{t|t-1} = 0$. This gives rise to the Chow solution

$$(6.2) \qquad y_t = -B_1^{-1}(B+I)y_{t-1} + B^{-1}Kz_t + B_1^{-1}\Gamma z_{t-1} + B^{-1}u_t + B^{-1}Ru_{t-1}$$

where K and R are arbitrary real 2x2 matrices and B_1 is assumed to be nonsingular for simplicity. The corresponding DT solution is

$$(6.3) \qquad y_t = (B+I)^{-1}\Gamma z_t + B^{-1}u_t$$

implying the restrictions,

$$R = BB_1^{-1}(B+I)B^{-1}$$

and $\qquad K = B(B+I)^{-1}\Gamma$

As in our simple example in section 5, there is an identification problem here in that B_1 cannot be estimated. Nevertheless we can compare the estimates of those parameters which can be estimated.

Four samples of the time series y_t, with fifty observations each, were generated by using (6.2). The parameter values for the four samples are summarized in Table 1 below.

Table 1. Parameter Values Generating Four Samples

Sample	σ_u	K, R
1	.01	K=R=0
2	.1	"
3	.01	K, R subject to DT restrictions
4	.1	" "

For each sample, the structural parameters have the values

$$B = \begin{pmatrix} 1 & 5 \\ 5/6 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1/2 & 6 \\ 0 & 2 \end{pmatrix}, \quad \Gamma = \begin{pmatrix} -5 & 0 \\ -2 & 1 \end{pmatrix}$$

which imply that y_t is a stationary time series. Each of the time series generated by (6.2) satisfies the original model (6.1). With those parameter values the Dagli-Taylor implicit restrictions on R and K are,

$$R = \begin{pmatrix} 0.447 & 0.763 \\ -0.044 & -0.447 \end{pmatrix}$$

$$K = \begin{pmatrix} -5 & -30 \\ -1 & 13 \end{pmatrix}$$

Given the sample data, we used an approximate full-information maximum likelihood (FIML) method described below to estimate the parameters of both (6.2) and (6.3). In addition, the method of minimum distance as suggested by Malinvaud (1970) was also used to estimate (6.2). This method consists of minimizing the expression

$$\sum_{t=1}^{T} u_t' \hat{\Omega}^{-1} u_t$$

iteratively, where $\hat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} u_t u_t'$ is obtained from the previous iteration. The log-likelihood function to be maximized in conjunction with the approximate FIML method is

$$\log L = T \log |B| - \frac{T}{2} \log \sigma_u^4 - \frac{1}{2\sigma_u^2} \sum_{t=1}^{T} u_t' u_t$$

where (6.2) or (6.3) is used to substitute for u_t, but the initial value u_0 was assumed to be zero as an approximation. In each case for the Chow method the starting values used in the optimization routine were obtained by applying two-stage least squares to the model (6.1) with $y_{t+1|t-1}$ replaced by y_{t+1}. For the DT method the starting values were equal to the actual parameter values. For identification purposes the structural parameters were restricted as follows (the β's and γ's were to be estimated).

$$B = \begin{pmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 5/6 & 1 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} \beta_3 & \beta_4 \\ 0 & \beta_5 \end{pmatrix} = \begin{pmatrix} 1/2 & 6 \\ 0 & 2 \end{pmatrix}$$

and $$\Gamma = \begin{pmatrix} \gamma_1 & 0 \\ -2 & \gamma_2 \end{pmatrix} \quad \begin{pmatrix} -5 & 0 \\ -2 & 1 \end{pmatrix}$$

When estimating (6.2)

$$K = \begin{pmatrix} k_1 & k_2 \\ k_3 & k_4 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} r_1 & r_2 \\ r_3 & r_4 \end{pmatrix}$$

were left unrestricted. Tables 2 to 5 summarize the results. The following abbreviations are used in Tables 2 to 5.

DT: Dagli-Taylor Method - FIML

CM: Chow Method - Minimum Distance

CF: Chow Method - FIML

The gradient algorithm of Davidon, Fletcher and Powell included in the GQOPT program of Goldfeld and Quandt was used to perform the numerical maximization.

In each case the Chow method provides reasonably close estimates of all parameters and the method of minimum distance seems to perform about as well as the approximate FIML method. With $\sigma_u = .01$ we see from Table 2 that the DT method does not provide close estimates of the structural parameters, with some missing the parameter values by almost a factor of two and others being less than one half the true parameter values. The estimate of σ_u is by far the worst, being 8.99. These poor estimates are not surprising since the data used do not satisfy the implicit restrictions imposed by the DT method. When σ_u is increased to .1 and again the implicit restrictions are not satisfied, the DT method breaks down by giving parameter estimates in the range $10^{12} - 10^{14}$. (see Table 3). On the other hand, when the data satisfy the DT implicit restrictions, the DT method as well as the Chow method combined with FIML and minimum distance all perform very well (see Tables 4 and 5). The reader should be reminded, however, that even under this circumstance, the DT method was unable to provide estimates for β_3, β_4 or β_5 due to the identification problem discussed earlier.

Table 2. Parameter Estimates Using Sample Data Set 1

	DT	CM	CF		DT	CM	CF		CM	CF		CM	CF
σ	8.99	.016	.016	β_4	---	6.00	6.00	k_1	-.003	-.003	r_1	0.00	0.0
β_1	6.09	4.99	4.99	β_5	---	2.00	2.00	k_2	.002	.002	r_2	-.002	0.0
β_2	1.64	.834	.833	γ_1	-1.88	-5.00	-4.99	k_3	-.002	-.002	r_3	-.001	0.0
β_3	----	.501	.501	γ_2	1.31	1.00	1.00	k_4	.002	.002	r_4	.002	0.0

Table 3. Parameter Estimates Using Sample Data Set 2.

	DT	CM	CF		DT	CM	CF		CM	CF		CM	CF
σ	2.2×10^{14}	.080	.078	β_4	---	6.00	6.00	k_1	-.013	-.016	r_1	0.00	0.0
β_1	9.8×10^{13}	5.00	5.00	β_5	---	2.00	2.00	k_2	.008	.008	r_2	-.002	-.003
β_2	6.1×10^{14}	.835	.834	γ_1	1×10^{12}	-4.99	-4.99	k_3	-.009	-.010	r_3	0.00	-.002
β_3	---	.504	.503	γ_2	-2×10^{13}	1.01	1.00	k_4	.01	.012	r_4	.003	.006

Table 4. Parameter Estimates Using Sample Data Set 3

	DT	CM	CF		DT	CM	CF		CM	CF		CM	CF
σ	.020	.019	.020	β_4	---	6.00	6.00	k_1	-5.00	- 5.00	r_1	.447	.447
β_1	5.00	5.00	5.00	β_5	---	2.00	2.00	k_2	-30.00	-30.00	r_2	.763	.763
β_2	.833	.833	.833	γ_1	-5.00	-5.00	-5.00	k_3	-1.00	- 1.00	r_3	-.044	-0.44
β_3	----	.500	.500	γ_2	1.00	1.00	1.00	k_4	13.00	13.00	r_4	-.447	-.447

Table 5. Parameter Estimates Using Sample Data Set 4

	DT	CM	CF		DT	CM	CF		CM	CF		CM	CF
σ	.100	.094	.100	β_4	---	6.00	6.00	k_1	-5.00	-5.00	r_1	.447	.447
β_1	5.00	5.00	5.00	β_5	---	2.00	2.00	k_2	-30.00	-30.00	r_2	.763	.765
β_2	.833	.833	.833	γ_1	-5.00	-5.00	-5.00	k_3	- 1.00	- 1.00	r_3	-.044	-.044
β_3	---	.500	.500	γ_2	1.00	1.00	1.00	k_4	13.00	13.00	r_4	-.447	-.447

7. Conclusion

The method of Dagli-Taylor [3] for solving linear simultaneous-equation models under rational expectations has been shown to be a special case of the method of Chow [2]. Furthermore, it has been shown that the Dagli-Taylor solution implies certain arbitrary restrictions on free parameters of the Chow solution which serve to characterize the multiple solutions arising from the original model. We call these restrictions arbitrary because they do not result in a solution set characterized by any of the more common restrictions employed in the literature (stationarity, convergent expectations). We have shown that there are many nonpathological (i.e., stationary with convergent expectations) time series processes which satisfy a given rational expectations model but for which the DT method provides inconsistent parameter estimates. The Chow solution has been shown to be the most general solution of the rational expectations model (1.1) which takes the form of a linear constant-coefficient process involving y_t, the disturbance u_t and the exogenous variable z_t. The methods of FIML and Minimum distance when applied to Chow's solution have worked well for our illustrative model, and can be recommended for econometric work employing linear simultaneous equations with expectations formed under the assumption of rational expectations.

REFERENCES

[1] Burmeister, E. and K. Wall: "Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to German Hyperinflation," Journal of Econometrics, 10 (1982), 255-284.

[2] Chow, G. C.: Econometrics, New York: McGraw-Hill, 1983.

[3] Dagli, C. A. and J. B. Taylor: "Estimation of Linear Rational Expectations Models Using a Matrix Polynomial Factorization," Princeton University, mimeo, November 1982.

[4] Evans, George and Seppo Honkapohja: "A Complete Characterization of ARMA Solutions to Linear Rational Expectations Models," unpublished paper, Stanford University, 1983.

[5] Flood, R. P. and P. M. Garber: "Market Fundamentals versus Price-Land Bubbles: The First Tests," Journal of Political Economy, 88 (1980), 745-770.

[6] Gourieroux, C., J. J. Laffont and A. Montfort: "Rational Expectations in Dynamic Linear Models," Econometrica, 50 (1982), 409-425.

[7] Malinvaud, E.: Statistical Methods of Econometrics, Amsterdam: North-Holland, 1970.

[8] McCallum, B. T.: "Rational Expectations and the Natural Rate Hypothesis: Some Consistent Estimates," Econometrica, 44 (1976), 43-52.

[9] Sargent, Thomas J.: Macroeconomic Theory, New York: Academic Press, 1979.