Federal Reserve Bank of Minneapolis
Research Department Working Paper

RATIONAL EXPECTATIONS, HYPERINFLATION, AND THE DEMAND FOR MONEY

Lawrence J. Christiano

Working Paper 163
PACS File 2700

Revised November 1981

This paper shows how to derive the family of models in which Cagan's model of hyperinflation is a rational expectations model. The slope parameter in Cagan's portfolio balance equation is identified in some of these models and in others it is not -- a fact which clarifies results obtained in several recent papers.

1. Introduction

In 1956 Phillip Cagan published a now-classic paper on the demand for money during hyperinflation. In that paper the demand for real cash balances is a function of the public's expectations of the future course of inflation. Cagan hypothesized that expectations of future inflation rates are formed by applying exponentially declining weights to past inflation rates (the "adaptive expectation" hypothesis). Later, Sargent and Wallace (1973) and Sargent (1977) produced a bivariate statistical model of money growth and inflation in which Cagan's portfolio balance equation is satisfied and adaptive expectations are rational in Muth's (1961) sense. Assuming that a given expectation scheme is rational amounts to providing that scheme with an economic rationale. In this sense the work of Sargent and Wallace can be viewed as exploring, more fully than did Cagan himself, the implications of Cagan's model for the data. (By "Cagan's model" we mean Cagan's portfolio balance equation with adaptive expectations.)

In this paper we reexamine the problem posed by Sargent and Wallace and find a family of observationally distinct bivariate models of money growth and inflation that satisfy both Cagan's portfolio balance equation and the requirement that adaptive expectations be rational. Following Hansen and Sargent (1980b), we shall refer to the problem of deriving the restrictions on a time series model implied by a given expectations scheme as the "inverse optimal predictor problem."[1] In this terminology, Sargent and Wallace find a particular solution to a given inverse optimal predictor problem, while we find the family of solutions. We believe that the results we obtain have important implications for estimating Cagan's model along the lines set out by Sargent and Wallace.

Sargent and Wallace's model ("Model 1" below) appears to be reasonably good one for several reasons. First, it is consistent with the exogeneity pattern between money creation and inflation found in the data. Also, the model implies that the disturbance term in Cagan's portfolio blance equation is highly serially correlated, a finding that is corroborated by other researchers. (See, e.g., Khan (1975).) Third, as Sargent (1976) has demonstrated, the model explains the apparantly anomalous empirical results obtained by Jacobs (1975) in his work on hyperinflation. Finally, results are reported in Sargent (1977) which suggest that the model holds up reasonably well when confronted with the data.

A difficulty in Sargent and Wallace's model is that the slope parameter in the portfolio balance equation is not statistically identifiable. On the other hand, it was his interest in this parameter that motivated the empirical part of Cagan's work in the first place. He hoped to use a statistical estimate of the slope parameter to infer whether a government was issuing money at the long-run revenue-maximizing rate. This, in turn, was expected to shed light on the causes of hyperinflation. Cagan also was interested in using an empirical estimate of the slope parameter to evaluate Friedman's (1956) conjecture that it is stable.

Cognizant of the importance of the slope parameter, Salemi and Sargent (1979) took the route of dropping Cagan's adaptive expectations assumption in their empirical work. They appear to have done so partly on the hunch that the lack of identifiability of the slope parameter is a consequence of the assumption that adaptive expectations are rational. According to Salemi and Sargent (1979, p. 751): "Sargent (1977) demonstrated that if Cagan's adaptively formed expectations are rational, α is _not_ econometrically identifiable." (Their emphasis.)

In this paper we show that the lack of identifiability of the slope parameter in Sargent and Wallace's model is not a consequence of their assumption of rational expectations. Rather, it is a property of the particular solution to the inverse optimal predictor problem that they happened to choose. For definiteness, we produce two other parameterizations ("Models 2 and 3") with the following properties: Cagan's model is a rational expectations model, the error term in the portfolio balance equation is highly serially correlated, the exogeneity pattern between money and prices just mentioned is duplicated, and the slope parameter of interest α is statistically identifiable. Furthermore, Models 2 and 3 make the same predictions regarding Jacob's estimator as does Sargent and Wallace's. (This fact can easily be verified by applying the techniques to Sargent (1976) to Model 2. We do not reproduce these calculations in this paper.) Model 3 has the additional implication that Cagan's estimator for the slope parameter in the portfolio balance equation is consistent.

The approach we take in this paper is the following. We begin by specifying a fairly general multivariate time series representation for the variables of Cagan's model. We then treat Cagan's model under rational expectations as a set of within- and cross-equations restrictions on the time series representation. We consider the inverse optimal predictor problem to be solved when we have found the minimum set of restrictions that the time series model must satisfy to be consistent with Cagan's model under rational expectations. This way of representing rational expectations model was recently spelled out in detail in a paper by Hansen and Sargent (1980a). There they domonstrate its applicability to many problems besides the one studied here. The application we consider is an example of what Hansen and Sargent call a rational expectations model with inexact cross-equations restrictions. The word "inexact" appears because one of the variables of the model--the disturbance term in the portfolio balance queation--is hidden from the econometrician.

2. Inverse Optimal Predictor Problem

Consider the following portfolio balance equation.

$$(1) \qquad m_t - p_t = \alpha\pi_t + u_t, \ \alpha < 0.$$

Here, m_t and p_t are the money supply, measured in log deviations from the mean, u_t is a stochastic shock term, and π_t is the public's expectation of $p_{t+1} - p_t$, the rate of inflation over the next period. Cagan (1956) assumed that the latter expectation was formed adaptively, according to the following formula:

$$(2) \qquad \pi_t = (\frac{1-\lambda}{1-\lambda L})(p_t - p_{t-1}), \quad \lambda < 1$$

where L is the lag operator.

The assumption that adaptive expectations are rational amounts to

$$(3) \qquad E_t(p_{t+1} - p_t) = \pi_t$$

where $E_t(\cdot) \equiv \hat{E}(\cdot\ p_{t-s}, m_{t-s}, u_{t-s}; s>0)$, and $\hat{E}$ is the least squares projection operator.

Since we only have data on $\{m_t\}$ and $\{p_t\}$, the inverse optimal predictor problem we wish to solve is the following: "Assuming all variables are Gaussian, what is the class of observationally distinct <u>bivariate</u> representations for $\{p_t\}$ and $\{m_t\}$ which is consistent with (1) - (3)?" It is convenient to first derive the family of trivariate representations for $\{m_t\}$, $\{p_t\}$, and $\{u_t\}$ which is consistent with (1) - (3).

Let $y(t) = (p(t),m(t),u(t))^T$. Suppose that $\{y(t)\}$ can be modelled as

$y(t) = c(L)a(t)$, where $c(L) = \sum_{i=0}^{\infty} c_i L^i$. Writing these out in detail,

$$\text{(4a)} \qquad \begin{pmatrix} p(t) \\ m(t) \\ u(t) \end{pmatrix} = \begin{bmatrix} c_{11}(L) & c_{12}(L) & c_{13}(L) \\ c_{21}(L) & c_{22}(L) & c_{23}(L) \\ c_{31}(L) & c_{32}(L) & c_{33}(L) \end{bmatrix} \begin{matrix} a_1(t) \\ a_2(t) \\ a_3(t) \end{matrix} ,$$

with $a(t) = (a_1(t),a_2(t),a_3(t))^T$. Here,

$$\text{(4b)} \qquad Ea(t)a(t-\tau)^T = \begin{cases} V & \tau=0 \\ \underset{\sim}{0} & \tau\neq 0 \end{cases} ,$$

where V is positive semidefinite and symmetric. Write

$$V = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} .$$

Condition (3) specifies that $\{p_t\}$ is exogenous in the $\{y_t\}$ process. By a trivariate version of Sim's (1972b) theorem, this implies that two polynomials in the L operator in the first row of c(L) may be set to zero. Condition (3) also restricts the parameters of the $\{p(t)\}$ process. Thus, (3) implies that we can set[2/]

(5) $$c_{11}(L) = \frac{(1-\lambda L)}{(1-L)^2},\ c_{12}(L) = 0,\ c_{13}(L) = 0.$$

Substituting (2) into (1) yields

$$m_t - p_t = \alpha(\frac{1-\lambda}{1-\lambda L})(1-L)p_t + u_t,$$

or,

(6) $$m_t - u_t = (\frac{1-\lambda L+a(1-\lambda)(1-L)}{1-\lambda L})p_t.$$

Because (6) and (5) hold for every sample realization of {a(t){, these restrictions hold in (4) if, and only if,

(7) $$c_{21}(L) - c_{31}(L) = \frac{1-\lambda L+\alpha(1-\lambda)(1-L)}{(1-L)^2}$$

and

(8) $$\begin{aligned} c_{22}(L) &= c_{32}(L) \\ c_{23}(L) &= c_{33}(L). \end{aligned}$$

Equation (7) implies that $c_{21}(L)$ and $c_{31}(L)$ can be written[3/]

(9) $$(1-L)^2 c_{21}(L) = [(1-k_1)-\lambda L(1-k_2)+\alpha(1-k_3)-\alpha\lambda(1-k_4)$$

$$- \alpha(1-k_5)L+\alpha\lambda(1-k_6)L] + \Psi_1(L)(1-L)^2$$

and

$$(10)\qquad (1-L)^2 c_{31}(L) = -[k_1-\lambda L k_2+\alpha k_3-\alpha\lambda k_4-\alpha k_5 L+\alpha\lambda k_6 L]+\Psi_1(L)(1-L)^2$$

where the scalars k_i, $i=1, \ldots, 6$ and the one-sided polynomial in positive powers of L, $\Psi_1(L)$, are arbitrary.

It will be convenient to have equations (9) and (10) in the following alternate form:

$$(11)\qquad (1-L)c_{21}(L) = \frac{1-\lambda+D_1+D_2(1-L)+\Psi_1(L)(1-L)^2}{(1-L)}$$

and

$$(12)\qquad (1-L)c_{31}(L) = \frac{D_1+D_3(1-L)+\Psi_1(L)(1-L)^2}{(1-L)}$$

where

$$(13)\qquad \begin{aligned} D_1 &= \alpha k_5 - \alpha\lambda k_6 - k_1 - \alpha k_3 + \alpha\lambda k_4 + \lambda k_2 \\ D_2 &= \lambda(1-k_2) + \alpha(1-k_5) - \alpha\lambda(1-k_6) \\ D_3 &= \alpha\lambda k_6 - \alpha k_5 - \lambda k_2. \end{aligned}$$

Restrictions (8) are satisfied simply by setting

$$(14)\qquad \begin{aligned} c_{22}(L) &= c_{32}(L) = \Psi_2(L) \\ c_{23}(L) &= c_{33}(L) = \Psi_3(L) \end{aligned}$$

where $\Psi_2(L)$ and $\Psi_3(L)$ are arbitrary one-sided polynomials in positive powers of L.

Equations (5) and (11) - (14) solve the following inverse optimal predictor problem: find the class of trivariate representations for $y(t)$ that is consistent with (1) - (3).

3. Identification

In the previous section we used the condition that $\{y(t)\}$ satisfy Cagan's model to restrict the parameters of (4). In this section we use the requirement that the parameters be identifiable from the second moment properties of $\{m_t\}$ and $\{p_t\}$ only to restrict the parameters of (4) even further.

First, since we do not have observations on $\{u(t)\}$, we can set $\Psi_3(L)$ identically to zero without restricting the autocovariances of $\{p_t\}$ and $\{m_t\}$. Accordingly, assume

$$(15) \qquad \Psi_3(L) \equiv 0.$$

The implication for $\{u(t)\}$ of this assumption is that (using (4), (12), and (14)),

$$(16) \qquad (1-L)u(t) = \frac{D_1+D_3(1-L)+\Psi_1(L)(1-L)^2}{(1-L)} a_1(t) + (1-L)\Psi_2(L)a_2(t).$$

From (5) and (7) we see that $\{m_t\}$ and $\{p_t\}$ must be second-differenced in order to induce stationarity. Write

$$(17) \qquad \begin{bmatrix} \tilde{x}(t) \\ \tilde{\mu}(t) \end{bmatrix} \equiv \begin{bmatrix} (1-L)x(t) \\ (1-L)\mu(t) \end{bmatrix}, \quad \begin{bmatrix} x(t) \\ \mu(t) \end{bmatrix} \equiv \begin{bmatrix} (1-L)p_t \\ (1-L)m_t \end{bmatrix}.$$

Then, the autocovariance generating function of the $(\tilde{x}(t),\tilde{\mu}(t))$ process is

$$(18a) \qquad \tilde{S}(z) = B(z)WB(z^{-1})^T$$

where

$$(18b) \qquad B(z) = \begin{bmatrix} (1-z)^2 c_{11}(z) & 0 \\ (1-z)^2 c_{21}(z) & (1-z)^2 c_{22}(z) \end{bmatrix}, \quad W = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}.$$

Applying (5), (11), and (14), we get

$$(19a) \qquad B(z) = \begin{bmatrix} 1 - \lambda z & 0 \\ B_{21}^0 + B_{21}^1 z + B_{21}^2(z) z^2 & B_{22}^0 + B_{22}^1 z + B_{22}^2(z) z^2 \end{bmatrix} = B_0 + B_1 z + B_2(z) z^2$$

where

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21}^i & B_{22}^i \end{bmatrix} \qquad i = 0,1,2$$

$$B_{21}^0 = 1 - \lambda + D_1 + D_2 + \Psi_0^1$$

$$B_{21}^1 = -(D_2 + 2\Psi_0^1 - \Psi_1^1)$$

$$B_{21}^2(z) = (\Psi_0^1 - 2\Psi_1^1 + \Psi_2^1(z)) + (\Psi_1^1 - 2\Psi_2^1(z))z + \Psi_2^1(z) z^2$$

$$(19b) \qquad B_{22}^0 = \Psi_0^2$$

$$B_{22}^1 = \Psi_1^2 - 2\Psi_0^2$$

$$B_{22}^{2}(z) = (\Psi_0^2 - 2\Psi_1^2 + \Psi_2^2(z)) + (\Psi_1^2 - 2\Psi_2^2(z))z + \Psi_2^2(z)z^2.$$

$$\Psi_i(z) = \Psi_0^i + \Psi_1^i z + \Psi_2^i(z)z^2, \; i=1,\ 2.$$

It is well known that the parameters of $(B(z),W)$ fail to be identifiable in two senses. We consider these now in turn. Write $B(z) = \sum_{i=0}^{\infty} B_i z^i$. Then,

$$\tilde{S}(z) = B(z)WB(z^{-1})^T = B(z)P^{-1}PWP^T(P^{-1})^T B(z^{-1})^T = \overline{B}(z)\overline{W}\,\overline{B}(z^{-1})^T \tag{20}$$

where P is an arbitrary invertible matrix and $\overline{B}(z) \equiv B(z)P^{-1}$, $\overline{W} \equiv PWP^T$. Equation (20) shows that if we have a likelihood function which is a function only of $\tilde{S}(z)$, then models $(\overline{B}(z),\overline{W})$ are indistinguishable from $(B(z),W)$.

If we take $\hat{B}(z) \equiv B(z)B_0^{-1}$ and $\hat{W} = B_0WB_0^T$, so that $\hat{B}_0 = I$, then the only invertible matrix P that satisfies

$$\hat{B}(z)\hat{W}\hat{B}(z^{-1})^T = \hat{B}(z)P^{-1}P\hat{W}P^T(P^{-1})^T\hat{B}(z^{-1})^T$$

is $P = I$. In this sense, setting $P = B_0$ in (20) resolves the identification difficulty suggested there. For this reason, from here on we work with the $(\hat{B}(z),\hat{W})$ system. Write $\hat{B}(z) \equiv I + \hat{B}_1 z + \hat{B}_2(z)z^2$. Then, using (19), we get

$$\hat{B}_1 = \begin{bmatrix} -\lambda & 0 \\ -(D_2+2\Psi_0^1-\Psi_1^1) + (1-\lambda+D_1+D_2+\Psi_0^1)\left(\dfrac{2\Psi_0^2-\Psi_1^2}{\Psi_0^2}\right) & -\left(\dfrac{2\Psi_0^2-\Psi_1^2}{\Psi_0^2}\right) \end{bmatrix}$$

$$\hat{B}_2(z) = \begin{bmatrix} 0 & 0 \\ B_{21}^2(z) - \left(\frac{1-\lambda+D_1+D_2+\Psi_0^1}{\Psi_0^2}\right)B_{22}^2(z) & \frac{B_{22}^2(z)}{\Psi_0^2} \end{bmatrix} \tag{21}$$

$$= \begin{bmatrix} 0 & 0 \\ \hat{B}_{21}^2(z) & \hat{B}_{22}^2(z) \end{bmatrix}, \text{ say.}$$

The second identification problem alluded to above is the following. In general, a polynomial matrix $P(z)$ can be found where $P(z)P(z^{-1})^T = I$ and

$$\begin{aligned} \tilde{S}(z) &= \hat{B}(z)\hat{W}\hat{B}(z^{-1})^T = \hat{B}(z)P(z)P(z^{-1})^T\hat{W}P(z)P(z^{-1})^T\hat{B}(z^{-1})^T \\ &= \hat{B}^*(z)\hat{W}^*\hat{B}^*(z^{-1})^T. \end{aligned} \tag{22}$$

Here, $\hat{B}^*(z) \equiv \hat{B}(z)P(z)$ and $\hat{W}^* \equiv P(z^{-1})^T\hat{W}P(z)$. (See Rozanov (1967) for a rigorous analysis of this source of identification difficulty.) In general, the number of distinct parameterizations $(\hat{B}^*(z),\hat{W}^*)$ that satisfy (22) is 2^n, where n is the number of distinct roots of det $\hat{B}^*(z)$. In the restricted case studied in this paper, the problem is less severe, however. From (21), we have

$$\det \hat{B}(z) = \hat{B}_{11}(z)\hat{B}_{22}(z) = (1-\lambda z)\left[1-\left(\frac{2\Psi_0^2-\Psi_1^2}{\Psi_0^2}\right)z + \left(\frac{B_{22}^2(z)}{\Psi_0^2}\right)z^2\right]. \tag{23}$$

From (23) we see that the assumption that Cagan's model of hyperinflation is a rational expectations model restricts one of the roots of det $\hat{B}(z)$ to the $\frac{1}{\lambda}$. This, by assumption, lies outside the unit circle (see equation (1)). The other root(s) (those of $\hat{B}_{22}(z)$) are left unrestricted by the assumption. Following Rozanov (1967), we restrict the set of admissible $P(z)$ transforms in (22) to $P(z) = I$ by assuming

$$(24) \qquad \hat{B}_{22}(z) \neq 0 \text{ for } z < 1.$$

We have established that, under (24), the parameters of $\hat{B}(z)$ can be identified from data. $\hat{B}(z)$ and $\hat{W}$ are the reduced form parameters of the model. The structural parameters are λ, α, k_i, $i=1, \ldots, 6$, and $\Psi_j(L)$, $j=1, 2$. Clearly--λ excepted--these parameters are not all identifiable given estimates of the reduced form parameters. To identify these, further restrictions need to be imposed. In this paper we do not investigate what minimum set of additional restrictions is necessary to achieve identification of the structural parameters. Presumably, such a minimum set is not unique, with each set having a different implication for the identifiability of the important parameter α.

4. Three Examples

In this section we consider three alternative sets of restrictions on the free structural parameters of the model. As we shall see, the three sets have different implications for the identifiability of α. In the first case (the one studied by Sargent and Wallace) α fails to be identified. In the case of Models 2 and 3, α is identified.

Model 1 (Sargent-Wallace)

The following case is studied in Sargent and Wallace (1973) and

Sargent (1977):

(25) $k_1 = 1,\ k_2 = 0,\ k_3 = 1,\ k_4 = 1,\ k_5 = \frac{\alpha+1}{\alpha},$

$k_6 = 1,\ \Psi_2(L) = \frac{1}{1-L},\ \Psi_1(L) = 0.$

This implies, by (13) and (19b),

(26) $D_1 = 0,\ D_2 = \lambda - 1,\ D_3 = -[1+\alpha(1-\lambda)],$

$B^2_{21}(z) = B^2_{22}(z) = 0,\ \Psi^1_0 = \Psi^1_1 = \Psi^1_2(z) = 0,$

$\Psi^2_0 = \Psi^2_1 = 1,\ \Psi^2_2(L) = \frac{1}{1-L}.$

Substituting (26) into (21), we get

(27a) $\hat{B}_1 = \begin{bmatrix} -\lambda & 0 \\ 1-\lambda & -1 \end{bmatrix},\ \hat{B}_2(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

so that [4/]

(27b) $\begin{bmatrix} (1-L)x(t) \\ (1-L)\mu(t) \end{bmatrix} = (I+\hat{B}_1 L) \begin{pmatrix} \hat{a}_1(t) \\ \hat{a}_2\ (t) \end{pmatrix} = \begin{bmatrix} 1-\lambda L & 0 \\ (1-\lambda)L & (1-L) \end{bmatrix} \begin{pmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \end{pmatrix},$

where $(\hat{a}_1(t), \hat{a}_2(t))^T = B_o(a_1(t)_1\ a_2(t))^T$ and B_o is defined in (19a).
Also, from (16),

$$(28) \qquad (1-L)u(t) = -(1+\alpha(1-\lambda))a_1(t) + a_2(t).$$

Two features of the model are the following. First, α cannot be identified from $\hat{B}(z)$ because it does not appear in (27a). Second, $\{u(t)\}$ in (28) follows a random walk. As it happens, the former is a consequence of the latter. This is easily demonstrated. From (16) we have that necessary and sufficient conditions for $\{u(t)\}$ to follow a random walk are that

$$(29) \qquad \Psi_1(L) = \frac{\kappa_1}{1-L},\ \Psi_2(L) = \frac{\kappa_2}{1-L},\ D_1 = 0$$

with κ_1 and κ_2 arbitrary constants. (Note, this means that $\Psi_0^i = \Psi_1^i = \kappa_i$ and $\Psi_2^i(z) = \frac{\kappa_i}{1-z}$ for i=1, 2.) Substituting (29) into (21), we obtain (27), as asserted. It follows that, if α is to be identified from the coefficients $\hat{B}(z)$, then $\{u(t)\}$ must not be specified to follow a random walk. In the case of Model 1, α cannot be identified at all since an estimate of $\hat{W}$ also provides no information on α. This is because $B_0 = I$, so that $\hat{W} = W$. (Sargent (1977) succeeds in identifying α from W by imposing further <u>ad hoc</u> restrictions on the parameters of Model 1.) In the rest of this section we consider one case ("Model 2") where α can be identified from $\hat{B}(z)$ and another ("Model 3") in which α is not identifiable from $\hat{B}(z)$, although it can be recovered from an estimate of $\hat{W}$.

<u>Model 2</u>

Let k_3 and k_4 be arbitrary and set

$$(30) \qquad k_5 = k_2 = 0,\ k_6 = \frac{\alpha-1}{\alpha},\ k_1 = -\lambda(\alpha-1) - \alpha k_3 + \alpha\lambda k_4,$$

$$\Psi_1(L) = 0, \ \Psi_2(L) = \frac{1+\delta_1 L}{1-L}.$$

By (13) and (19b), (30) implies

(31) $$D_1 = 0, \ D_2 = \alpha, \ D_3 = \lambda(\alpha-1), \ B^2_{21}(z) = 0, \ B^2_{22}(z) = -\delta_1, \ \Psi^1_i = 0$$

for all i,

$$\Psi^2_0 = 1, \ \Psi^2_1 = 1 + \delta_1, \ \Psi^2_2(L) = \frac{1+\delta_1}{1-L}.$$

Substituting (31) into (21) we get

$$\hat{B}_1 = \begin{bmatrix} -\lambda & 0 \\ -\alpha\delta_1 + (1-\lambda)(1-\delta_1) & \delta_1-1 \end{bmatrix}$$

(32a)

$$\hat{B}_2(z) = \begin{bmatrix} 0 & 0 \\ (1-\lambda+\alpha)\delta_1 & -\delta_1 \end{bmatrix},$$

so that

(32b) $$\begin{bmatrix} (1-L)x(t) \\ (1-L)\mu(t) \end{bmatrix} = (I+\hat{B}_1 L+\hat{B}_2 L^2) \begin{bmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \end{bmatrix}$$

Also, from (16)

$$(33) \qquad (1-L)u(t) = \lambda(\alpha-1)a_1(t) + (1+\delta_1 L)a_2(t).$$

Substituting (31) into (23), we see that the roots of det $\hat{B}(z)$ in (32) are $\frac{1}{\lambda}$, 1 and $-\frac{1}{\delta_1}$. By the argument given in the previous section, if we add the constraint that

$$(34) \qquad \frac{1}{\delta_1} > 1,$$

then $\hat{B}_1$ and $\hat{B}_2$ in (32b) are uniquely identified from the second moments of $((1-L)x(t), (1-L)\mu(t))$. But, by inspection of (32a), we see that if the elements of $\hat{B}(z)$ are identified, then λ, α, and δ_1 are too. This example shows that the assumption that adaptive expectations are rational does not imply that the slope parameter in Cagan's portfolio balance equation fails to be identified.

Now consider the following model:

Model 3

Let k_3 and k_4 be arbitrary and set

$$k_1 = -\alpha k_3 + \alpha\lambda k_4 + r,\ k_2 = \frac{r}{\lambda},\ k_5 = k_6 = 0,\ \psi_1(L) = 0,\ \psi_2(L) = \frac{1}{1-L}.$$

By (13) and (19b), this implies

$$D_1 = 0,\ D_2 = \lambda - r + \alpha(1-\lambda),\ B^2_{21}(z)=B^2_{22}(z)=0,\ \psi^1_i$$

$$= 0 \text{ all } i,\ \psi^2_0 = \psi^2_1 = 1.$$

Substituting the latter into (21), we get

$$\hat{B}_1 = \begin{bmatrix} -\lambda & 0 \\ 1-\lambda & -1 \end{bmatrix}, \qquad \hat{B}_2(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} .$$

Evidently, α fails to be identifiable from estimates of $\hat{B}(z)$.

Note that from (19),

$$B_o = \begin{bmatrix} 1 & 0 \\ 1 - r + \alpha(1-\lambda) & 1 \end{bmatrix} .$$

Consequently,

$$\begin{bmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \end{bmatrix} = \begin{bmatrix} a_1(t) \\ (1 - r + \alpha(1-\lambda))a_1(t) + a_2(t) \end{bmatrix} .$$

Let $\hat{W} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} \end{bmatrix}$. Then,

$$\begin{aligned} \hat{\sigma}_{12}/\hat{\sigma}_{11} &= (1 - r + \alpha(1-\lambda)) + \sigma_{12}/\sigma_{11} \\ &= 1 + \alpha(1-\lambda), \end{aligned}$$

because $\hat{\sigma}_{12} = (1 - r + \alpha(1-\lambda))\sigma_{11}$, $\hat{\sigma}_{11} = \sigma_{11}$ and $(\sigma_{12}/\sigma_{11}) = r$. Thus, since λ is identifiable from $\hat{B}(z)$, α can be identified by exploiting the fact that $\alpha = ((\hat{\sigma}_{12}/\hat{\sigma}_{11}) - 1)/(1 - \lambda)$ in Model 3.

Since Model 3 satisfies (29), we conclude that the lack of identifiability of the parameter α is not a consequence of Sargent and Wallace's assumtion that $\{u(t)\}$ follows a random walk.

5. (In) consistency of Cagan's Estimator for α

In this section we compute the population projection of $\{\mu(t)-x(t)$ on $\{x(t)\}$ under the assumption that Cagan's model of hyperinflation is a rational expectations model. We then use a formula due to Sims (1972a) to evaluate the consistency of Cagan's estimator under alternative specifications of the "true" rational expectations model. We confirm Sargent and Wallace's results that when Model 1 is the true model, then Cagan's estimator for λ is consistent, while his estimator for α is inconsistent. We then show that when Model 3 obtains, the regression Cagan computed is a projection equation, so that his estimator for α is consistant.

From (5), (11), and (15) we have

$$(35)\qquad \begin{pmatrix} x(t) \\ \mu(t) \end{pmatrix} = \begin{bmatrix} \frac{1-\lambda L}{1-L} & 0 \\ \frac{1-\lambda+D_1+D_2(1-L)+\Psi_1(L)(1-L)^2}{(1-L)} & (1-L)\Psi_2(L) \end{bmatrix} \begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix}$$

$$= F(L)a^1(t),$$

say, where $a^1(t) = (a_1(t),a_2(t))^T$. The autocovariance generating function of this process is

$$(36)\qquad S(z) = \begin{bmatrix} S_{xx}(z) & S_{x\mu}(z) \\ S_{\mu x}(z) & S_{\mu\mu}(z) \end{bmatrix} = F(z)VF(z^{-1})^T.$$

Substituting from (35) into (36), we get

$$S_{xx}(z) = \frac{(1-\lambda z)(1-\lambda z^{-1})\sigma_{11}}{(1-z)(1-z^{-1})}$$

(37) $$S_{\mu x}(z) = \{\frac{[1-\lambda+D_1+D_2(1-z)+\Psi_1(z)(1-z)^2]\sigma_{11}[1-\lambda z^{-1}]}{(1-z)(1-z^{-1})}$$

$$+ \frac{(1-z)\Psi_2(z)\sigma_{12}(1-\lambda z^{-1})}{(1-z^{-1})}\}.$$

The projection of $\{\mu(t)\}$ on $\{x(t)\}$ is defined as

(38) $$\mu(t) = h(L)x(t) + \nu(t)$$

where $Ex(t)\nu(t-s) = 0$ for all s. It may be shown that

(39) $$h(z) = \frac{S_{\mu x}(z)}{S_{xx}(z)}.$$ 5/

To get the projection of $\{\mu(t)-x(t)\}$ on $\{x(t)\}$, subtract $x(t)$ from both sides of (39) to get

(40) $$\mu(t) - x(t) = b^P(L)x(t) + \nu(t)$$

where $b^P(L) = h(L) - 1$. Substituting (37) into (39) and using the definition of $b^P(z)$, we get

(41) $$b^P(z) = \frac{1-\lambda+D_1+D_2(1-z)+(\Psi_1(z)+r\Psi_2(z))(1-z)^2}{1-\lambda z} -1$$

where $r = \sigma_{12}/\sigma_{11}$. 6/ (The fact that $b^P(z)$ is one sided in positive powers of

z is a consequence of the exogeneity of $\{x(t)\}$.) Least squares regression of $\{\mu(t)-x(t)\}$ on $\{x(t)\}$ gives consistent estimates of $\{b_i^p, i\geq 0\}$, where $b(z) = \sum_{i=0}^{\infty} b_i^p z^i$. Cagan estimated $\{b_i^c, i\geq 0\}$ in

$$\mu(t) - x(t) = b^c(L)x(t) + v^c(t) \tag{42}$$

subject to the following (nonlinear) constraints

$$b^c(z) = \alpha^c(1-\lambda^c)\left(\frac{1-z}{1-\lambda^c z}\right) = \frac{\alpha^c(1-\lambda^c)}{\lambda^c} + \frac{\alpha^c(1-\lambda^c)(1-(\lambda^c)^{-1})}{1-\lambda^c L}, \tag{43}$$

or,

$$b_0^c = \alpha^c(1-\lambda^c)$$

$$b_i^c = \alpha^c(1-\lambda^c)\left(1 - \frac{1}{\lambda^c}\right)(\lambda^c)^i, \quad i > 0.$$

Sims' formula asserts that in large samples, when $b^p(z)$ is the true distributed lag and $b^c(z)$ is the one imposed by the econometrician, then least squares picks the free parameters of $b^c(z)$ to minimize

$$\int_{-\pi}^{\pi} b^p(e^{-i\omega})-b^c(e^{-i\omega})^2 S_{xx}(e^{-i\omega})d\omega. \tag{44}$$ [7]

We use this formula to evaluate the consistency properties of Cagan's estimators for λ and α. Consider the following case.

<u>Model 1</u>

Under the Sargent-Wallace parameterization, we have, substituting (13) and (25) into (41),

(45) $$b^p(L) = (\frac{1-L}{1-\lambda L})(r-1) = \frac{r-1}{\lambda} + \frac{(r-1)(1-\lambda^{-1})}{1-\lambda L},$$

so that,

$$b_0^p = r - 1$$

$$b_i^p = (r-1)(1-\lambda^{-1})\lambda^i, \; i > 0.$$

Substituting (37), (43), and (45) into (44), we get

$$\int_{-\pi}^{\pi} (\frac{1-e^{-i\omega}}{1-\lambda e^{-i\omega}})(r-1)-(\frac{1-e^{-i\omega}}{1-\lambda^c e^{-i\omega}})\alpha^c(1-\lambda^c) \;^2\; \frac{1-\lambda e^{-i\omega}}{1-e^{-i\omega}} \;^2\sigma_{11} d\omega,$$

or,

$$\int_{-\pi}^{\pi} \frac{r-1}{1-\lambda e^{-i\omega}} - (\frac{\alpha^c(1-\lambda^c)}{1-\lambda^c e^{-i\omega}} \;^2\; 1-\lambda e^{-i\omega} \;^2\sigma_{11} d\omega.$$

The unique minimzer of the above integral is $\lambda^c = \lambda$ and $\alpha^c = (r-1)/(1-\lambda)$. It follows that, in the case of Model 1, Cagan's estimator for α (α^c) is inconsistent, while his estimator for λ (λ^c) is consistent.

Now, suppose that $\{u(t)\}$ follows a random walk. In Section 4 we showed that this is equivalent with (29). Substituting (29) into (41), we get

(46) $$b^p(z) = \frac{1-z}{1-\lambda z}(D_2-\lambda+\kappa_1+r\kappa_2).$$

Substituting (43) and (46) into (44), we get

$$\int_{-\pi}^{\pi} \left| \frac{D_2-\lambda+\kappa_1+r\kappa_2}{1-\lambda e^{-i\omega}} - \frac{\alpha^c(1-\lambda^c)}{1-\lambda^c e^{-i\omega}} \right|^2 \left| 1-\lambda e^{-i\omega} \right|^2 \sigma_{11} d\omega.$$

In this case $\lambda^c = \lambda$ and $\alpha^c = (D_2-\lambda+\kappa_1+r\kappa_2)/(1-\lambda)$ are the unique minimizers. Evidently, the estimator for λ, λ^c, is consistent. The estimator for α is consistent if, and only if,

$$\alpha(1-\lambda) = D_2 - \lambda + \kappa_1 + r\kappa_2.$$

Substituting from (13), we see that this is equivalent with

(47) $$D_3 + \kappa_1 + r\kappa_2 = 0.$$

Now, when $\{u(t)\}$ follows a random walk, (29) applies and

$$(1-L)u(t) = (D_3+\kappa_1)a_1(t) + \kappa_2 a_2(t).$$

Using the above relation, it is easy to verify that the condition $E(1-L)u(t) a_1(t-s) = 0$ for all s is equivalent with (47). This in turn is equivalent with the condition $E(1-L)u(t)\tilde{x}(t-s) = 0$ for all s and (comparing (1) with (42)) with the condition $(1-L)u(t) = v^c(t)$. Since the parameters of Model 3 satisfy (47), we conclude the following: the assumption that Cagan's model is a rational expectations model fails to imply that his estimator for α is inconsistent.

6. Conclusion

What we have shown in this paper is that Sargent and Wallace's model is only one of a family of observationally distinct time series representations of money growth and inflation in which Cagan's model is a rational expectations model. The slope parameter in Cagan's portfolio balance equation is identified in some of these models, and in others it is not. We have brought no evidence to bear on the question of which of these models lies closest to the truth. This is an empirical question that remains to be answered.

Footnotes

1/The inverse optimal predictor problem is related to problems studied in Mosca and Zappa (1979).

2/This is a solution to the following inverse optimal predictor problem: "what univariate representation for $\{p(t)\}$ is consistent with (3)?" This problem was posed and solved in Muth (1960).

3/What follows would be even more general if we set $k_i = k_i(L)$, where $k_i(L)$ is a polynomial in non-negative powers of L and i=1, ..., 6. We do not do this in order to keep the notation simple and because the generality in which we consider the problem is sufficient to establish the claims made in the introduction.

4/The notation used in this paper matches that used in Sargent (1977) to facilitate comparisons.

5/See, for example, Sargent (1979, Chapter XI).

6/It can easily be shown that, in equation (40), $\nu(t) = (1-L)\Psi_2(L)(a_2(t)-ra_1(t))$.

7/See Sargent (1979, p. 293) for a heuristic deriation of this formula. For an application of the formula in a context similar to the one in this paper, see Sargent (1976).

References

Cagan, Phillip, 1956, "The monetary dynamics of hyperinflation," in M. Friedman, ed., Studies in the Quantity Theory of Money (Chicago: University of Chicago Press).

Friedman, Milton, "The Quantity Theory of Money - A Restatement," in M. Friedman, ed. Studies in the Quantity Theory of Money (Chicago: University of Chicago Press).

Hansen, Lars Peter, and Thomas J. Sargent, 1980a, "Frequency domain methods for specifying and estimating exact rational expectations model, July, mimeo.

Hansen, Lars Peter, and Thomas J. Sargent, 1980b, "Aggregation over time and the inverse optimal predictor problem for adaptive expectations in continuous time," Federal Reserve Bank of Minneapolis Working Paper #157, July.

Jacobs, Rodney, 1975, "A difficulty with monetarist models of hyperinflation, "Economic Inquiry 13, September.

Khan, M.S., 1975, "The monetary dynamics of hyperinflation: A note," Journal of Monetary Economics 1, 355-362.

Mosca, Edoardo, and Giovanni Zappa, 1979, "Consistency conditions for the asymptotic innovations representation and an equivalent inverse regulation problem," IEEE Transactions on Automatic Control, Vol. AC-24, No. 3, June, 501-503.

Muth, J.F., 1960, "Optimal properties of exponentially weighted forecasts, "Journal of the American Statistical Association, Vol. 55, No. 290, 299-306.

Muth, J.F., 1961, "Rational expectations and the theory of price movements, "Econometrica 29, July, 315-335.

Rozanov, Yu. A., 1967, "Stationary random processes," (Holden-Day, San Francisco, California).

Salemi, Michael K., and Thomas J. Sargent, 1979, "The demand for money during hyperinflation under rational expectations: II," International Economic Review 20, 741-758.

Sargent, Thomas J., 1976, "Econometric exogeneity and alternative estimators of portfolio balance schedules for hyperinflations," Journal of Monetary Economics 2, 511-521.

Sargent, Thomas J., 1977, "The demand for money during hyperinflation under rational expecations: I," International Economic Review 18, 59-82.

Sargent, Thomas J., 1979, "Macroeconomic theory," (New York: Academic Press).

Sargent, Thomas J., and Neil Wallace, 1973, "Rational ecpectations and the dynamics of hyperinflation," International Economic Review, June, 328-350.

Sims, Christopher, 1972a, "Approximate prior restrictions in distributed lag estimation, "Journal of the American Statistical Association 67, 169-175.

Sims, Christopher, 1972b, "Money, income and causality," American Economic Review 62, 540-552.