Historical

# Aggregation Over Time and the Inverse Optimal Predictor Problem for Adaptive Expectations in Continuous Time

Lars Peter Hansen
and
Thomas J. Sargent

July 1980

Working Paper #157
PACS File #2700

NOT FOR DISTRIBUTION

Carnegie-Mellon University
and
Federal Reserve Bank of Minneapolis

University of Minnesota
and
Federal Reserve Bank of Minneapolis



Larry Christiano performed all of the numerical calculations in this paper, and also helped to evaluate the integrals in formula (18).

# 1. Introduction

In 1956 Milton Friedman [5] and Phillip Cagan [2] formulated and applied the adaptive expectations hypothesis. Shortly thereafter, John F. Muth [13] solved the following "inverse optimal predictor"[1/] problem: for what discrete-time, univariate stochastic process is the discrete-time version of the adaptive expectations mechanism optimal in the sense of delivering linear least squares forecasts? Much later Sargent [20] solved the following extended inverse optimal predictor problem: in the context of a discrete-time version of Cagan's model of portfolio balance, for what bivariate money creation, inflation stochastic process does a discrete-time version of adaptive expectations deliver linear least squares forecasts for inflation?

This paper solves the continuous-time version of both of these inverse optimal predictor problems. In the context of a continuous-time version of Cagan's portfolio schedule, we find the continuous-time, generalized stochastic process for the money supply and price level that makes the adaptive expectations mechanism yield linear least squares forecasts for inflation. This problem is of interest, if only because Cagan actually formulated his model in continuous time, as have many others, even though he eventually ended up estimating an approximating discrete-time model. In order to determine the "optimal" discrete-time approximating model, this paper goes on to deduce the discrete-time process for point-in-time observations on the money supply and the price level that is implied by that continuous-time model which makes adaptive expectations rational in continuous time. This permits us to determine a sense in which the discrete-time adaptive expectations scheme can be viewed as approximating a model in which agents are optimally forming adaptive expectations in continuous time. We are also able to derive an exact formula linking the discrete-time adaptive expectations decay parameter $\lambda$ to the continuous-time decay parameter $\beta$. We compare this formula to the approximation $\lambda = \exp(-\beta)$ used by Cagan.

The continuous-time stochastic process for inflation and money creation which makes adaptive expectations optimal for predicting inflation *ipso facto* has the property that money creation fails to Granger cause [6] inflation in continuous time. However, for discrete-time samples drawn from this continuous-time process, money creation does Granger cause inflation. This is an example of the effects of aggregation over time in interrupting Granger noncausality patterns that hold for continuous time, a phenomenon that Sims [22, 23] has studied. The present model is simple enough that we are able to analyze this effect quite completely.

It is our hope that the calculations contained in this paper are interesting for their own sake, and also because they illustrate a way of analyzing the effects of aggregation over time that could be applied to a variety of linear rational expectations models.

## 2. The Continuous-Time Inverse Optimal Predictor Problem

We begin with Cagan's portfolio balance schedule in continuous time[2/]

(1) $$m(t) - p(t) = \alpha Dp(t) + a(t), \alpha < 0$$

where $p(t)$ is the logarithm of the price level, $m(t)$ is the logarithm of the money supply, $a(t)$ is a random disturbance to the portfolio balance schedule, and $D = d/dt$ is the time derivative operator. We find it convenient to define the derivatives

$$\mu(t) = Dm(t)$$

(2) $$x(t) = Dp(t)$$

$$\eta(t) = Da(t).$$

Differentiating (1) and using the definitions (2) gives

$$\mu(t) - x(t) = \alpha Dx(t) + \eta(t)$$

or

(3) $$(-\rho-D)x(t) = -\rho\mu(t) + \rho\eta(t)$$

where $\rho \equiv 1/2$.

We shall posit that $\mu(t)$ and $\eta(t)$ are generalized stochastic processes, specified in such a way that Cagan's adaptive expectations formula turns out to be implied by the hypothesis of rational expectations. Now the realizable, time invariant solution of the differential equation (2) for $x(t)$ is

$$x(t) = -\rho\hat{E}_t(\frac{1}{-\rho-D})\mu(t) + \rho\hat{E}_t(\frac{1}{-\rho-D})\eta(t)$$

or

(4) $$x(t) = -\rho\hat{E}_t\int_0^\infty e^{\rho u}\mu(t+u)du + \rho\hat{E}_t\int_0^\infty e^{\rho u}\eta(t+u)du$$

where $\hat{E}_t(\cdot)$ is the linear least squares projection of $(\cdot)$, conditioned on the information set $\{\mu(s),x(s): \ s \leq t\}$. Our objective is to determine stochastic processes for $\mu(t)$ and $\eta(t)$, which in conjunction with (4) imply the optimality of Cagan's adaptive expectation mechanism.[3/] That is, we wish to find specifications for $\mu(t)$ and $\eta(t)$, which together with (4) imply that

$$\hat{E}_t x(t+\tau) = \beta\int_{-\infty}^{t} e^{-\beta(t-s)}x(s)ds, \ \beta > 0, \ \tau > 0.$$

To produce the desired $(\mu(t),\eta(t))$ process,[4/] let $(w_1(t),w_2(t)) = w(t)^T$ be a continuous-time white noise vector with $Ew(t) = 0$, and

$$Ew(t)w(t-s)^T = V\delta(t-s)$$

where $V$ is a positive definite matrix and $\delta(\cdot)$ is the Dirac delta generalized function. Assume that $\mu(t)$, $\eta(t)$ are described by the continuous-time, generalized stochastic processes[5/]

$$\mu(t) = \frac{\beta}{D}w_1(t) + w_2(t), \ \beta > 0$$

(5)

$$\eta(t) = -\alpha Dw_1(t) + k_1w_1(t) + k_2w_2(t)$$

where $k_1$ and $k_2$ are arbitrary constants.[6/] To investigate the implications of (5) in conjunction with (4), we need a formula for evaluating the predictions of the terms of the form $\hat{E}_\rho\int_0^\infty e^{\rho u}z(t+a)du$ that appear in (4).[7/] Let a vector $z(t)$ have the representation

$$(6) \qquad z(t) = \sum_{j=1}^{r}\phi_j(D)w_j(t)$$

where $[w_1(t),\ldots,w_r(t)]$ is a fundamental white noise vector for $z(t)$, and $\phi_j(0)$ are rational polynomials in D.[8/] Then in the present context it can be established that the following formula of Hansen and Sargent [10] applies:[9/]

(7) $$\hat{E}_t\int_0^\infty e^{\rho u}z(t+a)du = \sum_{j=1}^{r}\left(\frac{\phi_j(D)-\phi_j(-\rho)}{D+\rho}\right)w_j(t).$$

By using formula (7), it follows from (4) that with the $\mu(t)$, $\eta(t)$ processes given by (5), the $x(t)$ process is

(8) $$x(t) = \left(\frac{D+\beta}{D}\right)w_1(t).$$

Thus, with the $(\mu(t),\eta(t))$ processes (5), we have that the joint process for the variables $(x(t),\mu(t))$ can be represented as[10/]

(9) $$\begin{pmatrix} x(t) \\ \mu(t) \end{pmatrix} = \begin{pmatrix} \frac{D+\beta}{D} & 0 \\ \frac{\beta}{D} & 1 \end{pmatrix}\begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix}.$$

We proceed to verify that for the joint $(x,\mu)$ process (9), Cagan's adaptive expectations formulation is rational. The first equation of (9) can be written

$$x(t+\tau) = w_1(t+\tau) + \beta\int_t^{t+\tau}w_1(s)ds + \beta\int_{-\infty}^{t}w_1(s)ds,\ \tau > 0.$$

Using $\hat{E}_t w_1(s) = 0$ for $s > 0$, and $\hat{E}_t w_1(s) = w_1(s)$ for $s \leq 0$, gives

(10) $$\hat{E}_t x(t+\tau) = \beta\int_{-\infty}^{t}w_1(s)ds = \frac{\beta}{D}w_1(t),\ \tau > 0.$$

Next, notice that $x(t) = \frac{D+\beta}{D}w_1(t)$ implies that $w_1(t) = \frac{D}{D+\beta}x(t)$. Substituting this last equality into (10) gives

(11) $$\hat{E}_t x(t+\tau) = \frac{\beta}{D+\beta}x(t),\ \tau > 0.$$

Since $\frac{\beta}{D+\beta}x(t) = \beta\int_0^\infty e^{-\beta u}x(t-u)du$, equation (11) establishes that the adaptive expectations scheme for $x(t)$ is rational under the joint $(x,\mu)$ process given by (9).

The second line of (9) can be written as

$$\mu(t+\tau) = w_2(t+\tau) + \beta\int_{-\infty}^{t} w_1(s)ds + \beta\int_{t}^{t+\tau} w_1(s)ds, \ \tau > 0.$$

Hence we have

$$\hat{E}_t \mu(t+\tau) = \beta\int_{-\infty}^{t} w_1(s)ds = \frac{\beta}{D} w_1(t)$$

or

$$(12) \qquad \hat{E}_t \mu(t+\tau) = \frac{\beta}{D+\beta} x(t), \ \tau > 0.$$

Equations (11) and (12) characterize the Granger-causality structure of the system (9). In the continuous-time system (9), $\mu$ fails to Granger cause x, as (10) establishes. However, x does Granger cause $\mu$, as (12) establishes. Even stranger, current and lagged $\mu$'s fail to help predict $\mu$, once current and lagged x's are taken into account. That these features characterize our system (9) is not surprising, since we constructed (9) in order to guarantee that Cagan's adaptive expectation mechanism (11) is consistent with optimal forecasting. In light of equation (4), if Cagan's mechanism is to be rational, there must be extensive feedback from x(t) to $\mu(t)$.

## 3. Effects of Aggregation Over Time

We use the definitions $Dp(t) = x(t)$, $Dm(t) = \mu(t)$ and (9) to deduce the generalized stochastic process for p(t), m(t):

$$(13) \qquad \begin{pmatrix} p(t) \\ m(t) \end{pmatrix} = \begin{bmatrix} \frac{D+\beta}{D^2} & 0 \\ \frac{\beta}{D^2} & \frac{1}{D} \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}.$$

The presence of D and $D^2$ in the denominator of the "moving average" polynomials on the right side of (13) indicates that (p(t),m(t)) is a nonstationary process.

We are interested in deducing the implications of (13) for point-in-time sampled, discrete-time observations on (p(t),m(t)). We shall assume that point-in-time observations on (p(t),m(t)) are available at the integers t=0, 1, 2, .... It turns out that the second differences of (p(t),m(t)) form a stationary discrete-time process with a very simple representation.

We consider now the discrete-time process that is formed by taking second differences of point-in-time observations on (p(t),m(t)) at the integers. We first note that the lag operator L can be represented as $L = e^{-D}$. Then the first difference operator is $(1-L) = (1-e^{-D})$, while the second difference operator is $(1-L)^2 = (1-e^{-D})^2$. Applying this operator to (13) gives

$$(14) \qquad \begin{pmatrix} (1-L)^2 p(t) \\ (1-L)^2 m(t) \end{pmatrix} = \begin{pmatrix} \frac{(1-e^{-D})^2}{D^2}(\beta+D) & 0 \\ \frac{(1-e^{-D})^2}{D^2}\beta & \frac{(1-e^{-D})^2}{D} \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix}.$$

Now recall the following Laplace transform pairs:[11/]

$$\frac{1-e^{-s}}{s^2} \leftrightarrow \begin{cases} t & t \varepsilon [0,1] \\ 1 & t \geq 1 \end{cases}$$

(15)
$$\frac{(1-e^{-s})^2}{s^2} \leftrightarrow \begin{cases} t & t\varepsilon[0,1] \\ 2-t & t\varepsilon[1,2] \\ 0 & t > 2. \end{cases}$$

Using the Laplace transforms (15) and (14) gives the desired representation:

(16)
$$(1-L)^2p(t) = \int_0^1(\beta\tau+1)w_1(t-\tau)d\tau + \int_1^2(\beta(2-\tau)-1)w_1(t-\tau)d\tau$$

$$(1-L)^2m(t) = \int_0^1\beta\tau w_1(t-\tau)d\tau + \int_1^2\beta(2-\tau)w_1(t-\tau)d\tau$$

$$+ \int_0^1 w_2(t-\tau)d\tau - \int_1^2 w_2(t-\tau)d\tau.$$

To represent things compactly, we define

$$y(t) = \begin{bmatrix} (1-L)^2p(t) \\ (1-L)^2m(t) \end{bmatrix}.$$

Then we can write (16) as

(17)
$$y(t) = \int_0^1 \begin{bmatrix} (\beta\tau+1) & 0 \\ \beta\tau & 1 \end{bmatrix} \begin{bmatrix} w_1(t-\tau) \\ w_2(t-\tau) \end{bmatrix} d\tau$$

$$+ \int_1^2 \begin{bmatrix} \beta(2-\tau)-1 & 0 \\ \beta(2-\tau) & -1 \end{bmatrix} \begin{bmatrix} w_1(t-\tau) \\ w_2(t-\tau) \end{bmatrix} d\tau.$$

Evidently, by virtue of the white noise property of $(w_1(t),w_2(t))$, $y(t)$ sampled at the integers is a first-order, bivariate moving average process with unconditional mean $Ey(t) = 0$. The autocovariogram of the $y(t)$ process is readily computed from[12/]

$$\Gamma_0 = Ey(t)y(t)^T$$

$$= \int_0^1 \begin{bmatrix} \beta\tau+1 & 0 \\ \beta\tau & 1 \end{bmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{bmatrix} \beta\tau+1 & \beta\tau \\ 0 & 1 \end{bmatrix} d\tau$$

$$+ \int_1^2 \begin{bmatrix} \beta(2-\tau)-1 & 0 \\ \beta(2-\tau) & -1 \end{bmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{bmatrix} \beta(2-\tau)-1 & \beta(2-\tau) \\ 0 & -1 \end{bmatrix} d\tau$$

and

$$\Gamma_1 = Ey(t)y(t-1)^T$$

$$= \int_1^2 \begin{bmatrix} \beta(2-\tau)-1 & 0 \\ \beta(2-\tau) & -1 \end{bmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{bmatrix} \beta(\tau-1)+1 & \beta(\tau-1) \\ 0 & 1 \end{bmatrix} d\tau$$

and

$$\Gamma_{-1} = Ey(t)y(t+1)^T = \Gamma_1^T$$

$$\Gamma_j = \Gamma_{-j} = 0 \text{ for } j > 1.$$

Evaluating the above integrals, we obtain

$$\Gamma_0 = \begin{bmatrix} 2v_{11}[\frac{1}{3}\beta^2+1] & 2v_{11}[\frac{1}{3}\beta^2+(\frac{v_{12}}{v_{11}})] \\ 2v_{11}[\frac{1}{3}\beta^2 + \frac{v_{12}}{v_{11}}] & 2v_{11}[\frac{1}{3}\beta^2 + (\frac{v_{22}}{v_{11}})] \end{bmatrix}$$

(18)

$$\Gamma_1 = \begin{bmatrix} v_{11}[\frac{1}{6}\beta^2-1] & v_{11}\ \frac{\beta}{2}[\frac{1}{3}\beta-1]+v_{12}[\frac{1}{2}\beta-1] \\ v_{11}\ \frac{\beta}{2}[\frac{1}{3}\beta+1]-v_{12}[\frac{1}{2}\beta+1] & v_{11}[\frac{1}{6}\beta^2 - \frac{v_{22}}{v_{11}}] \end{bmatrix}$$

## 4. Predicting Inflation Using Information on Lagged Inflation Only

We first consider the univariate Wold representation for the $(1-L)^2p(t)$ process. From (18), $(1-L)^2p(t)$ is a first-order moving average with covariance generating function

$$(19) \qquad g(z) = c(1)z^{-1} + c(0) + c(1)z$$

where from (18) $c(0) = 2v_{11}[\frac{1}{3}\beta^2+1]$, $c(1) = v_{11}[\frac{1}{6}\beta^2-1]$. We seek the Wold moving average representation for $(1-L)^2p(t)$, which is of the form

$$(20) \qquad (1-L)^2p(t) = (1-\lambda_pL)\varepsilon_{pt}, \ |\lambda_p| < 1$$

with $\varepsilon_{pt}$ a discrete-time white noise that is fundamental for $(1-L)^2p(t)$; the variance of the one-step-ahead prediction error $\varepsilon_{pt}$ is $\sigma^2_{\varepsilon p}$. From a routine application of the spectral factorization theorem,[13/] we have the following formulas for $\lambda_p$ and $\sigma^2_{\varepsilon p}$;

$$(21) \qquad \lambda_p = -\frac{1}{2}\frac{c(0)}{c(1)} \pm \sqrt{\frac{c(0)^2}{4c(1)^2} - 1}$$

subject to $|\lambda_p| < 1$

$$\sigma^2_{\varepsilon p} = \frac{c(0)}{1+\lambda^2}.$$

Using the preceding formulas for $c(0)$ and $c(1)$ we have

$$(22) \qquad \lambda_p = -\frac{(\frac{1}{3}\beta^2+1)}{\frac{1}{6}\beta^2-1} \pm \sqrt{\frac{(\frac{1}{3}\beta^2+1)^2}{(\frac{1}{6}\beta^2-1)^2} - 1}$$

subject to $|\lambda_p| < 1$.

Now consider the discrete-time inflation rate $X(t)$ which we define as $X(t) = p(t) - p(t-1)$ for $t$ at the integers. Representation (20) can then be written

$$(23) \qquad (1-L)X(t) = (1-\lambda_p L)\varepsilon_{pt}.$$

As shown by John F. Muth [13], the optimal j-step-ahead forecast of X governed by process (23), given current and lagged values of X alone, is the discrete-time version of Cagan's adaptive expectation schemes

$$(24) \qquad \hat{E}X(t+j)\,|\,X(t),X(t-1),\ \ldots = (1-\lambda_p)\sum_{i=0}^{\infty}\lambda_p^i X(t-i),\ j \geq 1.$$

Now equation (24) is precisely the discrete-time representation which Cagan used for approximating the continuous-time adaptive expectations scheme

$$\hat{E}_t x(t+\tau) = \beta\int_0^{\infty} e^{-\beta s}x(t-s)ds,\ \tau > 0.$$

Cagan took $\lambda_p$ to be related to β via the equation

$$(25) \qquad \lambda_p = e^{-\beta}.$$

For various values of β, Table 1 reports the values of $\lambda_p$ given by formula (22) and Cagan's formula (25). For β close to zero, equation (25) provides a close approximation to (22). However, for large values of β, exp(-β) is approximately zero, while equation (22) implies a $\lambda_p$ of approximately -.25.

This comparison is of interest in the following context. Suppose that our continuous-time model is correct, and that an analyst possesses discrete-time observations on p(t), t belonging to the integers. A procedure recommended by Nerlove [15] and Nerlove, Grether, and Carvalho [14] would be to determine the optimal predictors for the univariate process for p(t), and then attribute them to the private agents in the model. This procedure is motivated by an appeal to the rational expectations hypothesis, and is termed the method of "quasi rational expectations" by Nerlove, Grether, and Carvalho [14]. In an infinitely large

Table 1

| $\beta$ | $\lambda_p$ | $\exp(-\beta)$ |
|---|---|---|
| 0 | 1.000000 | 1.000000 |
| .25 | .778290 | .778801 |
| .50 | .603289 | .606531 |
| .75 | .463584 | .472367 |
| 1.00 | .351000 | .367879 |
| 1.25 | .259528 | .286505 |
| 1.50 | .184661 | .223130 |
| 1.75 | .122966 | .173774 |
| 2.00 | .071797 | .135335 |
| 2.25 | .029094 | .105399 |
| 2.50 | -.006757 | .082085 |
| 2.75 | -.037033 | .063928 |
| 3.00 | -.062746 | .049787 |
| 3.25 | -.084705 | .038774 |
| 3.50 | -.103558 | .030197 |
| 3.75 | -.119828 | .023518 |
| 4.00 | -.133939 | .018316 |
| 4.25 | -.146237 | .014264 |
| 4.50 | -.157003 | .011109 |
| 4.75 | -.166469 | .008652 |
| 5.00 | -.174828 | .006738 |
| 5.25 | -.182238 | .005248 |
| 5.50 | -.188832 | .004087 |
| 5.75 | -.194722 | .003183 |
| 6.00 | -.200000 | .002479 |
| 6.25 | -.204746 | .001930 |
| 6.50 | -.209027 | .001503 |
| 6.75 | -.212899 | .001171 |
| 7.00 | -.216413 | .000912 |
| 7.25 | -.219609 | .000710 |
| 7.50 | -.222524 | .000553 |
| 7.75 | -.225189 | .000431 |
| 8.00 | -.227632 | .000335 |
| 8.25 | -.229875 | .000261 |
| 8.50 | -.231940 | .000203 |
| 8.75 | -.233845 | .000158 |
| 9.00 | -.235605 | .000123 |
| 9.25 | -.237234 | .000096 |
| 9.50 | -.238746 | .000075 |
| 9.75 | -.240150 | .000058 |
| 10.00 | -.241457 | .000045 |
| $+\infty$ | -.267949 | 0.000000 |

sample, the analyst could recover the parameter $\lambda_p$ given by formula (21), if he followed Nerlove, Grether, and Carvalho's method. Using formula (22) or Table 1, the analyst could then infer the value of $\beta$. Table 1 provides a fairly complete characterization of Cagan's approximation (25) as a vehicle for inferring $\beta$ from $\lambda$.

## 5. Predicting Inflation Using Information on Lagged Inflation and Lagged Money Creation

We now turn to the bivariate moving average of the discrete-time process for inflation and money creation. A Wold moving average representation for $((1-L)^2 p(t),(1-L)^2 m(t))^T = y(t)$ is

(26) $$y(t) = u_t + Fu_{t-1}$$

where $u_t$ is a (2x1) vector discrete-time white noise with $Eu_t u_t^T = \overline{V}$, where $\overline{V}$ is a positive semidefinite matrix; $u_t = y(t) - \hat{E}y(t)|y(t-1),y(t-2), \ldots$; and the eigenvalues of F are less than or equal to unity in absolute value. Given $\Gamma_0$ and $\Gamma_1$ from (18), F and $\overline{V}$ are determined by solving the following equations

$$(I+Fz)\overline{V}(I+Fz^{-1})^T = \Gamma_1^T z^{-1} + \Gamma_0 + \Gamma_1 z$$

or

(27) $$\Gamma_0 = V + F\overline{V}F^T$$
$$\Gamma_1 = F\overline{V}.$$

The spectral factorization theorem discussed by Rozanov [19] implies that these equations have a unique solution with the properties indicated above. In practice, we have solved the above equations for $\overline{V}$ and F by using an algorithm described by Rozanov [19]. By following Rozanov's suggestions, Hansen and Sargent [9, Appendix B] describe explicit closed-form formulas for $\overline{V}$ and F as functions of the elements of $\Gamma_0$ and $\Gamma_1$.

Letting $X_t = p(t) - p(t-1)$, $M_t = m(t) - m(t-1)$, we can write (26) as

(28) $$\begin{bmatrix}(1-L)X_t \\ (1-L)M_t\end{bmatrix} = u_t + Fu_{t-1}.$$

By carrying out a series of calculations paralleling these of Muth [13], it is straightforward to verify that (28) admits the alternative representation

$$\begin{bmatrix} X_{t+1} \\ M_{t+1} \end{bmatrix} = (I+F)(I+FL)^{-1} \begin{bmatrix} X_t \\ M_t \end{bmatrix} + u_{t+1}.$$

or

$$(29) \qquad \begin{bmatrix} X_{t+1} \\ M_{t+1} \end{bmatrix} = (I+F) \sum_{i=0}^{\infty} (-F)^i \begin{bmatrix} X_{t-i} \\ M_{t-i} \end{bmatrix} + u_{t+1}.$$

From the fact that u is fundamental for (X,M), it can be readily verified that there obtains the following bivariate generalization of Cagan's adaptive expectations scheme:

$$(30) \qquad \left[ \hat{E} \begin{matrix} X_{t+j} \\ M_{t+j} \end{matrix} | X_t, M_t, X_{t-1}, M_{t-1}, \ldots \right] = (I+F) \sum_{i=0}^{\infty} (-F)^i \begin{bmatrix} X_{t-i} \\ M_{t-i} \end{bmatrix}, \; 2 \geq 1.$$

The one-step-ahead prediction error vector is $u_{t+1}$, which has covariance matrix $\overline{V}$.

Representation (28) is usefully compared to the one constructed by Sargent [20]. He posited a discrete-time model of the inflation, money creation, process which makes the discrete-time version of adaptive expectations rational when taken in conjunction with a discrete-time version of Cagan's portfolio balance schedule. Sargent's model is the discrete-time, bivariate, first-order moving average

$$(31) \qquad \begin{bmatrix} (1-L)X_t \\ (1-L)M_t \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} + \begin{pmatrix} -\lambda & 0 \\ (1-\lambda) & -1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{pmatrix}$$

where $(\varepsilon_{1t}, \varepsilon_{2t})^T = \varepsilon_t$ is a discrete-time vector white noise with arbitrary contemporaneous covariance matrix $E\varepsilon_t\varepsilon_t^T = W$; $\varepsilon_t$ is fundamental for $((1-L)X_t, (1-L)M_t)$; and $|\lambda| < 1$. It is evident from the first equation of (31) that Cagan's discrete-time adaptive expectations formulation for inflation is rational, given (31).14/

In form, (31) matches (28). One of our tasks now is to study the relation between the (2x2) matrix F in (28) and the corresponding matrix

$$E = \begin{pmatrix} -\lambda & 0 \\ (1-\lambda) & -1 \end{pmatrix}$$

in (31). Notice that the eigenvalues of the matrix E are -1 and $-\lambda$. It can be proved[15/] that one of the eigenvalues of F in (28) is -1. A comparison between the value of $-\lambda_p$ given by equation (22) and the nonunit eigenvalue of F is one interesting measure of the effects of time aggregation.

For various values of V and $\beta$ we have calculated F and $\overline{V}$. In addition, we calculated $\lambda_p$ and $\sigma^2_{\varepsilon p}$ in the univariate Wold moving average representation for $(1-L)X_t$:

$$(1-L)X_t = (1-\lambda_p L)\varepsilon_{pt}, \ |\lambda_p| < 1$$

where $\varepsilon_p$ is a fundamental white noise for $(1-L)X_t$ and $\sigma^2_{\varepsilon p} = E\varepsilon^2_p$ is the one-step-ahead prediction error variance for $(1-L)X_t$. We also calculated the univariate Wold moving average representation for $(1-L)M_t$

$$(1-L)M_t = (1-\lambda_m L)\varepsilon_{mt}$$

where $\varepsilon_{mt}$ is a fundamental white noise for $(1-L)M_t$ and $\sigma^2_{\varepsilon m} = E\varepsilon^2_m$ is the one-step-ahead prediction error variance for $(1-L)M_t$.[16/] Recall that $\overline{V}_{11}$ is the discrete-time, one-step-ahead prediction error variance for predicting $X_t$ on the basis of lagged X's and lagged M's, while $\overline{V}_{22}$ is the discrete-time, one-step-ahead prediction error variance in predicting $M_t$ on the basis of lagged X's and lagged M's. Therefore, $(\sigma^2_{\varepsilon p}-V_{11})|\sigma^2_{\varepsilon p}$ is a measure of the marginal assistance of lagged M's in predicting $x_t$, while $(\sigma^2_{\varepsilon m}-V_{22})|\sigma^2_{\varepsilon m}$ is a measure of the marginal assistance of lagged X's in predicting $M_t$. These quantities, which we call "percentage gains" in Tables 2-7, are measures of the strength of the Granger causality that occur between the discrete-time X and M processes. We recall that in the continuous-time model (9), which we are maintaining, M(t) fails to Granger cause X(t). However, in the discrete-time model, $M_t$ will in general Granger cause $X_t$ due to the effects of aggregation over time.[17/] The percentage gain $(\sigma_{\varepsilon p}^2 - V_{11}) | \sigma_{\varepsilon p}^2$ is a measure of the failure of the discrete-time process to reveal the Granger causality structure of the underlying continuous-time model.

Tables 2-4 report complete characterizations of $\overline{V}$, F, $\sigma_{\varepsilon p}^2$, $\sigma_{\varepsilon m}^2$, $\lambda_p$, and $\lambda_m$ for three values of β, and for three settings for the "intensity" matrix V. Tables 5-7 give less complete characterizations of $\overline{V}$ and F for a large number of values of β.

One outstanding characteristic that emerges from these tables is that for small values of β, not only does exp(-β) approximate $\lambda_p$ well, but the matrix F approximates the matrix

$$\begin{pmatrix} -\lambda & 0 \\ (1-\lambda) & -1 \end{pmatrix}$$

well with λ taken to be $\lambda_p$ or exp(-β). Further, for small β, money creation only very weakly Granger-causes inflation in the discrete-time data.

On the other hand, for large values of β, exp(-β) fails to approximate $\lambda_p$ well, and F fails to resemble the matrix

$$\begin{pmatrix} -\lambda & 0 \\ (1-\lambda) & -1 \end{pmatrix}.$$

In addition, for large β, substantial Granger causality can extend from money creation to inflation in discrete time.

For values of $\lambda_p$ in the range estimated by Cagan [2] and Sargent [20], these results are moderately comforting, since they suggest that aggregation over time imparts at most a very small asymptotic bias to Cagan's estimator of β.[18/] They also are compatible with the weak evidence in discrete time for Granger causality extending from money creation to inflation.

On the other hand, the tables also indicate that for high values of β the effects of aggregation over time can be considerable. In particular, while

Table 2

<u>$V = I,\ \beta = .05$</u>

$$F = \begin{pmatrix} -.9512240 & -.5079E\text{-}05 \\ .0487756 & -1.00000 \end{pmatrix},\quad \overline{V} = \begin{pmatrix} 1.05084 & .02840 \\ .02840 & 1.00084 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 1.05084$; % gain = 0; $\lambda_p = .951224$

$\sigma^2_{\varepsilon m} = 1.05084$; % gain = 4.758; $\lambda_m = .951224$

eigenvalues of F: -1.0, -.951229

<u>$V = I,\ \beta = 2.05$</u>

$$F = \begin{pmatrix} .0139974 & -.133791 \\ 1.0140000 & -1.133790 \end{pmatrix},\quad \overline{V} = \begin{pmatrix} 4.76243 & 2.73743 \\ 2.73743 & 2.71243 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 4.78290$; % gain = .428; $\lambda_p = .0626363$

$\sigma^2_{\varepsilon m} = 4.78290$; % gain = 43.289; $\lambda_m = .0626363$

eigenvalues of F: -1.0, -.119794

<u>$V = I,\ \beta = 10.05$</u>

$$F = \begin{pmatrix} 1.83797 & -1.75125 \\ 2.83797 & -2.75125 \end{pmatrix},\quad \overline{V} = \begin{pmatrix} 60.9131 & 54.8881 \\ 54.8881 & 50.8631 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 65.5079$; % gain = 7.014; $\lambda_p = -.241708$

$\sigma^2_{\varepsilon m} = 65.5079$; % gain = 22.356; $\lambda_m = -.241708$

eigenvalues of F: -1.0, .0867212

Table 3

$V = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}$, $\beta = .05$

$$F = \begin{pmatrix} -.9512230 & -.507230E-04 \\ .0487767 & 1.00005 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 10.508400 & .258385 \\ .258385 & 1.008390 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 10.5084$; % gain = 0; $\lambda_p = .951224$

$\sigma^2_{\varepsilon m} = 1.16661$; % gain = 13.563; $\lambda_m = .853612$

eigenvalues of F: -1.0, -.951274

$V = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}$, $\beta = 2.05$

$$F = \begin{pmatrix} .329567 & -.694458 \\ 1.329570 & -1.694460 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 46.7499 & 26.4999 \\ 26.4999 & 17.2499 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 47.8290$; % gain = 2.256; $\lambda_p = .0626363$

$\sigma^2_{\varepsilon m} = 28.7633$; % gain = 40.028; $\lambda_m = -.208744$

eigenvalues of F: -1.0, -.364891

$V = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}$, $\beta = 10.05$

$$F = \begin{pmatrix} 5.07515 & -5.39143 \\ 6.07515 & -6.39143 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 526.424 & 466.174 \\ 466.174 & 416.924 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 655.079$; % gain = 19.640; $\lambda_p = -.241708$

$\sigma^2_{\varepsilon m} = 630.971$; % gain = 33.924; $\lambda_m = -.265206$

eigenvalues of F: -1.0, -.316278

Table 4

$V = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}$, $\beta = .05$

$$F = \begin{pmatrix} -.9512240 & -.507996E-06 \\ .0487755 & -1.00000 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 1.0508400 & .0258385 \\ .0258385 & 10.0008000 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 1.05084$; % gain = 0; $\lambda_p = .951224$

$\sigma^2_{\varepsilon m} = 10.1589$; % gain = 1.556; $\lambda_m = .984313$

eigenvalues of F: -1.0, -.951225

$V = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}$, $\beta = 2.05$

$$F = \begin{pmatrix} -.0540333 & -.0149769 \\ .9459670 & -1.0149800 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 4.78062 & 2.75562 \\ 2.75562 & 11.73060 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 4.78290$; % gain = .048; $\lambda_p = .0626363$

$\sigma^2_{\varepsilon m} = 17.9960$; % gain = 34.816; $\lambda_m = .516757$

eigenvalues of F: -1.0, -.0690102

$V = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}$, $\beta = 10.05$

$$F = \begin{pmatrix} .466067 & -.244613 \\ 1.466070 & -1.244610 \end{pmatrix}, \quad \overline{V} = \begin{pmatrix} 64.8440 & 58.8190 \\ 58.8190 & 63.7940 \end{pmatrix}$$

$\sigma^2_{\varepsilon p} = 65.5079$; % gain = 1.013; $\lambda_p = -.241708$

$\sigma^2_{\varepsilon m} = 86.7970$; % gain = 26.502; $\lambda_m = -.0787326$

eigenvalues of F: -1.0, .221454

Table 5

V=I

| β | exp(-β) | % gain p | % gain m | $\lambda_p$ | eigenvalue of F |
|---|---|---|---|---|---|
| .05 | .951229 | .000 | 4.758 | .951224 | -.951229 |
| .15 | .860708 | .000 | 12.957 | .860587 | -.860708 |
| .25 | .778801 | .000 | 19.662 | .778290 | -.778799 |
| .35 | .704688 | .000 | 25.133 | .703413 | -.704680 |
| .45 | .637628 | .001 | 29.581 | .635156 | -.637604 |
| .55 | .576950 | .003 | 33.181 | .572824 | -.576890 |
| .65 | .522046 | .005 | 36.073 | .515811 | -.521922 |
| .75 | .472367 | .010 | 38.375 | .463584 | -.472138 |
| .85 | .427415 | .017 | 40.186 | .415673 | -.427030 |
| .95 | .386741 | .026 | 41.585 | .371661 | -.386136 |
| 1.00 | .367879 | .032 | 42.153 | .351000 | -.367138 |
| 2.00 | .135335 | .396 | 43.475 | .071797 | -.127017 |
| 3.00 | .049787 | 1.216 | 38.864 | -.062746 | -.026334 |
| 4.00 | .018316 | 2.253 | 34.399 | -.133939 | .021749 |
| 5.00 | .006738 | 3.297 | 30.901 | -.174828 | .047553 |
| 6.00 | .002479 | 4.251 | 28.251 | -.200000 | .062746 |
| 7.00 | .000912 | 5.090 | 26.228 | -.216413 | .072361 |
| 8.00 | .000335 | 5.817 | 24.656 | -.227632 | .078800 |
| 9.00 | .000123 | 6.445 | 23.409 | -.235605 | .083308 |
| 10.00 | .000045 | 6.989 | 22.401 | -.241457 | .086582 |
| 11.00 | .000017 | 7.462 | 21.573 | -.245871 | .089031 |
| 12.00 | .000006 | 7.876 | 20.881 | -.249278 | .090909 |
| 13.00 | .000002 | 8.240 | 20.297 | -.251959 | .092380 |
| 14.00 | .000001 | 8.562 | 19.797 | -.254106 | .093554 |
| 15.00 | .000000 | 8.850 | 19.364 | -.255850 | .094504 |
| 16.00 | .000000 | 9.107 | 18.987 | -.257287 | .095285 |
| 17.00 | .000000 | 9.339 | 18.655 | -.258483 | .095933 |
| 18.00 | .000000 | 9.548 | 18.361 | -.259490 | .096478 |
| 19.00 | .000000 | 9.738 | 18.098 | -.260345 | .096940 |
| 20.00 | .000000 | 9.911 | 17.863 | -.261077 | .097335 |

Table 6

$$V = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}$$

| β | exp(-β) | % gain p | % gain m | $\lambda_p$ | eigenvalue of F |
|---|---|---|---|---|---|
| .05 | .951229 | .000 | 13.563 | .951224 | -.951274 |
| .15 | .860708 | .000 | 30.723 | .860587 | -.861784 |
| .25 | .778801 | .001 | 39.841 | .778290 | -.783221 |
| .35 | .704688 | .003 | 44.530 | .703413 | -.715367 |
| .45 | .637628 | .010 | 46.743 | .635156 | -.657499 |
| .55 | .576950 | .023 | 47.568 | .572824 | -.608589 |
| .65 | .522046 | .046 | 47.620 | .515811 | -.567484 |
| .75 | .472367 | .083 | 47.253 | .463584 | -.533032 |
| .85 | .427415 | .135 | 46.673 | .415673 | -.504170 |
| .95 | .386741 | .205 | 45.996 | .371661 | -.479957 |
| 1.00 | .367879 | .248 | 45.644 | .351000 | -.469338 |
| 2.00 | .135335 | 2.120 | 40.205 | .071797 | -.367138 |
| 3.00 | .049787 | 5.169 | 37.661 | -.062746 | -.339408 |
| 4.00 | .018316 | 8.313 | 36.325 | -.133939 | -.328647 |
| 5.00 | .006738 | 11.104 | 35.519 | -.174828 | -.323453 |
| 6.00 | .002479 | 13.460 | 34.985 | -.200000 | -.320572 |
| 7.00 | .000912 | 15.421 | 34.606 | -.216413 | -.318813 |
| 8.00 | .000335 | 17.054 | 34.323 | -.227632 | -.317663 |
| 9.00 | .000123 | 18.426 | 34.105 | -.235605 | -.316871 |
| 10.00 | .000045 | 19.586 | 33.931 | -.241457 | -.316303 |
| 11.00 | .000017 | 20.578 | 33.790 | -.245871 | -.315881 |
| 12.00 | .000006 | 21.434 | 33.672 | -.249278 | -.315559 |
| 13.00 | .000002 | 22.179 | 33.573 | -.251959 | -.315309 |
| 14.00 | .000001 | 22.831 | 33.489 | -.254106 | -.315110 |
| 15.00 | .000000 | 23.408 | 33.416 | -.255850 | -.314949 |
| 16.00 | .000000 | 23.920 | 33.352 | -.257287 | -.314817 |
| 17.00 | .000000 | 24.379 | 33.296 | -.258483 | -.314708 |
| 18.00 | .000000 | 24.791 | 33.246 | -.259490 | -.314617 |
| 19.00 | .000000 | 25.164 | 33.202 | -.260345 | -.314539 |
| 20.00 | .000000 | 25.502 | 33.162 | -.261077 | -.314473 |

Table 7

$$V = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}$$

| $\beta$ | $\exp(-\beta)$ | % gain p | % gain m | $\lambda_p$ | eigenvalue of F |
|---|---|---|---|---|---|
| .05 | .951229 | -.000 | 1.556 | .951224 | -.951225 |
| .15 | .860708 | .000 | 4.524 | .860587 | -.860599 |
| .25 | .778801 | .000 | 7.308 | .778290 | -.778341 |
| .35 | .704688 | .000 | 9.917 | .703413 | -.703540 |
| .45 | .637628 | .000 | 12.361 | .635156 | -.635403 |
| .55 | .576950 | .000 | 14.647 | .572824 | -.573236 |
| .65 | .522046 | .001 | 16.785 | .515811 | -.516434 |
| .75 | .472367 | .001 | 18.782 | .463584 | -.464461 |
| .85 | .427415 | .002 | 20.645 | .415673 | -.416843 |
| .95 | .386741 | .003 | 22.382 | .371661 | -.373162 |
| 1.00 | .367879 | .003 | 23.205 | .351000 | -.352679 |
| 2.00 | .135335 | .044 | 34.465 | .071797 | -.077935 |
| 3.00 | .049787 | .145 | 38.882 | -.062746 | .052290 |
| 4.00 | .018316 | .284 | 39.554 | -.133939 | .120272 |
| 5.00 | .006738 | .433 | 38.302 | -.174828 | .158938 |
| 6.00 | .002479 | .575 | 36.158 | -.200000 | .182582 |
| 7.00 | .000912 | .704 | 33.687 | -.216413 | .197926 |
| 8.00 | .000335 | .819 | 31.186 | -.227632 | .208380 |
| 9.00 | .000123 | .920 | 28.804 | -.235605 | .215792 |
| 10.00 | .000045 | 1.009 | 26.607 | -.241457 | .221222 |
| 11.00 | .000017 | 1.087 | 24.614 | -.245871 | .225312 |
| 12.00 | .000006 | 1.157 | 22.825 | -.249278 | .228465 |
| 13.00 | .000002 | 1.218 | 21.226 | -.251959 | .230945 |
| 14.00 | .000001 | 1.272 | 19.800 | -.254106 | .232929 |
| 15.00 | .000000 | 1.321 | 18.529 | -.255850 | .234541 |
| 16.00 | .000000 | 1.365 | 17.394 | -.257287 | .235867 |
| 17.00 | .000000 | 1.405 | 16.378 | -.258483 | .236971 |
| 18.00 | .000000 | 1.441 | 15.467 | -.259490 | .237900 |
| 19.00 | .000000 | 1.474 | 14.648 | -.260345 | .238688 |
| 20.00 | .000000 | 1.504 | 13.910 | -.261077 | .239363 |

Cagan's approximation $\lambda = \exp(-\beta)$ prevents $\lambda$ from assuming negative values, negative $\lambda$'s can occur in the appropriate discrete-time model.

Fortunately, there is no need to count on the parameter $\beta$ staying in the range in which time aggregation effects are small. It is straightforward to implement procedures for estimating the parameters of the continuous-time model, $\beta$ and V, given records of discrete-time data. Equation (28) is a bivariate moving average representation for $\{(1-L)X_t, (1-L)M_t\}$, which can be estimated using either time domain or frequency domain versions of method of maximum likelihood.[19/] The likelihood function would be maximized over the free parameters, $\beta$ and V, of the continuous-time model.[20/]

## 6. Conclusions

We have produced a continuous-time model which solves the inverse optimal predictor problem for a continuous-time version of Cagan's model of hyperinflation with adaptive expectations. We have gone on to deduce the restrictions which this continuous-time model places on discrete-time data. This has permitted us to describe exact formulas linking the parameters of the discrete-time representation to the parameters of the continuous-time model. These formulas permit us to evaluate the quality of the approximations that Cagan and others have used in linking the discrete-time and continuous-time parameterizations.

The computational techniques used in this paper are useful for studying the effects of aggregation over time in a variety of dynamic models under rational expectations. In subsequent research we plan to use these tools to study the effects of aggregation over time in substantially richer dynamic contexts.

Footnotes

1/ Linear inverse optimal control and linear inverse optimal predictor problems are analyzed in discrete time by Mosca and Zappa [12].

2/ In this paper we use the following operational calculus. Let $F(t)$ be a function or generalized function defined on $t\varepsilon(-\infty,+\infty)$. Let $f(s)$ be the Laplace transform of $F(t)$, which we denote by $f(s) \leftrightarrow F(t)$. Let $D$ be the time derivative operator, and let $x(t)$ be a stochastic process or generalized stochastic process. Then we have $f(D)x(t) = \int_{-\infty}^{\infty} F(\tau)x(t-\tau)d\tau$. In conjunction with this equality, we use the following Laplace transform pairs in this paper: $1/s \leftrightarrow 1$; $1/s^2 \leftrightarrow t$;

$$\frac{1}{s-a} \leftrightarrow \begin{cases} -e^{at}, & t \leq 0 \\ 0, & t > 0 \end{cases} \quad , a > 0;$$

$$\frac{1}{s-a} \leftrightarrow \begin{cases} e^{at}, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad , a \leq 0;$$

$e^{-as} \leftrightarrow \delta(t-a)$ where $\delta(\cdot)$ is the Dirac delta generalized function; and $e^{-as}/s \leftrightarrow u(t-a)$ where $u(t)$ is the Heaviside unit step function, $u(t) = 1$, $t \geq 0$, $u(t) = 0$, $t < 0$. For descriptions of Laplace transforms, see Churchill [3] or Doetsch [4]. For a useful treatment of the operational properties of delta functions and other generalized functions, see Papoulis [17].

3/ This is the continuous-time version of the inverse optimal predictor problem studied by Sargent [20].

4/ Kwakernaak and Sivan [11] and Papoulis [16] contain useful introductions to the properties of continuous-time white noises and continuous-time stochastic processes constructed by integrating and differentiating them. Arnold [1] contains an introduction to the Ito stochastic integrals which are used to define and manipulate continuous-time white noises in a mathematically rigorous way.

5/ Notice that $\mu(t)$ and $\eta(t)$ contain white noise and derivatives of white noise components, respectively. Consequently, neither of these processes is physically realizable. Our defense for positing these ideal or generalized stochastic processes is that they lead to well-defined, physically realizable stochastic processes for the discrete-time observations on the money supply and price level.

6/ Using the operational calculus described in footnote 2, we have

$$\frac{\beta}{D}w_1(t) = \beta\int_0^{\infty} w_1(t-s)ds$$

which is a Wiener process that has infinite variance at t because it is viewed as starting up in the infinite past. Two alternative devices could be used to assure that all the objects manipulated have finite variances for all finite t.

The first would be to start up the system with fixed or random initial conditions $(x(t_0),\eta(t_0))$ for some finite time $t_0$. The second would be to start the system up infinitely far in the past, but to approximate terms like $\frac{\beta}{D}x(t)$ by

$$\frac{\beta}{D+\gamma}x(t) = \beta\int_0^\infty e^{-\gamma s}x(t-s)ds$$

for very small positive values of $\gamma$. The results in this paper can be produced by replacing $\frac{\beta}{D}$ by $\frac{\beta}{D+\gamma}$, $\gamma > 0$, and by taking the limit as $\gamma \to 0$. In this way, the results in this paper could all be obtained without resorting to manipulations involving terms with infinite variances, such as $\frac{\beta}{D}w_1(t)$.

7/ Whittle [24] describes the continuous-time versions of the Wiener-Kolmogorov linear least squares prediction formulas used here.

8/ Rozanov [19] uses the concept of a fundamental white noise. A white noise process $(w_1(t),\ldots,w_r(t))$ is said to be fundamental for $z(t)$ if for $\tau > 0$ $z(t+\tau) - \hat{E}[z(t+\tau)|z(s),s\leq t]$ can be expressed as an integral of $[w_1(z),\ldots,w_r(z)]$ over the interval $t \leq z \leq t + \tau$.

9/ Hansen and Sargent [10] derived the formula under more restrictive conditions than apply here. Nevertheless, it can be verified directly using the continuous-time prediction formulas given by Whittle [24] that the formula is correct.

10/ Note that neither the $x(t)$ process nor the $\mu(t)$ process is physically realizable due to the presence of white noise components in each.

11/ See Churchill [3] or Doetsch [4].

12/ We are using the rules for taking expected values of products of integrals of white noises that are described by Kwakernaak and Sivan [11, pp. 97-99].

13/ This theorem is discussed by Rozanov [19] in generality. Sargent [20, pp. 265-268] provides a nontechnical discussion of factoring the covariance generating function of a first-order moving average process.

14/ This is because the first difference of inflation is a first-order moving average, and because the Wold moving average representation for $(1-L)X_t$, $(1-L)M_t$ is triangular, implying that $(1-L)M_t$ fails to Granger cause $(1-L)X_t$. See Sims [23].

15/ First, note that

$$(I+Fz)\overline{V}(I+Fz^{-1})^T = \Gamma_{-1}z^{-1} + \Gamma_0 z + \Gamma_1 z$$

and that therefore the zeroes of $\det(\Gamma_{-1}z^{-1}+\Gamma_0 z+\Gamma_1 z)$ are comprised of the zeroes of $\det(I+Fz)$ and the reciprocals of the zeroes of $\det(I+Fz)$. Next, note that the zeroes of $\det(I+Fz)$ are minus the reciprocals of the eigenvalues of F. By using formulas (18) it can be proved that unity is a zero of $\det(\Gamma_1^T z^{-1}+\Gamma_0+\Gamma_1 z)$, which implies that -1 is an eigenvalue of F.

16/ Each of these univariate moving averages was calculated by using the covariances given in (18) together with formulas (21).

17/ In a more general context, Sims [22, 23] has emphasized that $\tilde{y}$'s failing to Granger cause $\tilde{x}$ in continuous time does not imply that $\tilde{y}$ fails to Granger cause $\tilde{x}$ in discrete time.

18/ See Sargent [20] for an argument that Cagan's procedure for estimating $\lambda$ is statistically consistent, provided that expectations are rational and that the money creation inflation process is given by (31).

19/ Such approximations are discussed by Hannan [7], Hansen and Sargent [8], and Phadke and Kadem [18].

20/ We note that without a priori restrictions on V, Cagan's parameter $\alpha$ is not identifiable. Identification could presumably be achieved by employing the strategy which Sargent [20] employed in discrete time, namely, the strategy of a priori restricting the covariance matrix of the shocks to the portfolio balance schedule and the money supply rule.

References


[1] Arnold, Ludwig, 1974, "Stochastic Differential Equations: Theory and Applications," (John Wiley, New York).

[2] Cagan, Phillip, 1956, "The Monetary Dynamics of Hyperinflation," in Studies in the Quantity Theory of Money (M. Friedman, ed.), (Chicago, University of Chicago), 25-117.

[3] Churchill, R. V., 1972, Operational Mathematics, Third Edition, (McGraw-Hill).

[4] Doetsch, Gustav, 1961, Guide to the Applications of Laplace Transforms, D. Van Nostrand, (London).

[5] Friedman, Milton, 1957, A Theory of the Consumption Function, (Princeton, New Jersey, Princeton University Press).

[6] Granger, C. W. J., 1969, "Investigating Causal Relations by Econometric Models and Cross Spectral Methods," Econometrica 37, 424-438.

[7] Hannan, E. J., 1970, Multiple Time Series, (Wiley, New York).

[8] Hansen, L. P., and T. J. Sargent, 1980, "Formulating and Estimating Dynamic Linear Rational Expectations Models," Journal of Economic Dynamics and Control, forthcoming.

[9] Hansen, L. P., and T. J. Sargent, 1980, "Linear Rational Expectations Models for Dynamically Interrelated Variables," in R. E. Lucas, Jr., and T. J. Sargent, eds., Rational Expectations and Econometric Practice, (University of Minnesota Press, Minneapolis).

[10] Hansen, L. P., and T. J. Sargent, 1980, "Methods For Estimating Continuous Time Rational Expectations Models From Discrete Time Data," manuscript.

[11] Kwakernaak, H., and R. Sivan, 1972, Linear Optimal Control Systems, (Wiley, New York).

[12] Mosca, Edoardo, and Giovanni Zappa, June 1979, "Consistency Conditions for the Asymptotic Innovations Representation and An Equivalent Inverse Regulation Problem," IEEE Transactions on Automatic Control, Vol. AC-24, No. 3, 501-503.

[13] Muth, J. F., 1960, "Optimal Properties of Exponentially Weighted Forecasts," Journal of the American Statistical Association, Vol. 55, No. 290, 299-306.

[14] Nerlove, Marc, David M. Grether, and Jose L. Carvalho, 1979, Analysis of Economic Time Series: A Synthesis, (Academic Press, New York).

[15] Nerlove, Marc, 1967, "Distributed Lags and Unobserved Components in Economic Time Series," in W. Fellner, et al., Ten Economic Studies in the Tradition of Irving Fisher, (John Wiley, New York).

[16] Papoulis, Athanasios, 1965, Probability, Random Variables, and Stochastic Processes, (McGraw-Hill, New York).

[17] Papoulis, Athanasios, 1962, The Fourier Integral and Its Applications, (McGraw-Hill, New York).

[18] Phadke, M. S., and G. Kadem, 1978, "Computation of the Exact Likelihood Function of Multivariate Moving Average Models," Biometrika 65, 511-519.

[19] Rozanov, Y. A., 1967, Stationary Random Processes, (Holden-Day, San Francisco, California).

[20] Sargent, T. J., 1977, "The Demand for Money During Hyperinflations Under Rational Expectations: I," International Economic Review 18, 59-82.

[21] Sargent, T. J., 1979, Macroeconomic Theory, (Academic Press, New York).

[22] Sims, C. A., 1971, "Discrete Approximations to Continuous Time Lag Distributions in Econometrics," Econometrica 39, 545-564.

[23] Sims, C. A., 1972, "Money Income and Causality," American Economic Review 62, 540-552.

[24] Whittle, P., 1963, Regulation and Prediction by Linear Least-Square Methods, D. Van Nostrand, (Princeton, New Jersey).