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

Lars Peter Hansen

Carnegie-Mellon University

Thomas J. Sargent

University of Minnesota and
Federal Reserve Bank of Minneapolis

Research Department Staff Report 74
Federal Reserve Bank of Minneapolis

September 1981

ABSTRACT

This paper describes the continuous time stochastic process for money and inflation under which Cagan's adaptive expectations model is optimal. It then analyzes how data formed by sampling money and prices at discrete points in time would behave.



Lawrence Christiano performed all of the numerical calculations, and also helped to evaluate the integrals in formula (24). This research was supported in part by NSF Grant SES-8007016.

## 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, <u>univariate</u> 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 <u>bivariate</u> 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 non-causality 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.

This paper provides an exact answer to the questions raised by Benjamin Friedman [30] about the implications of aggregation over time for the appropriate interpretation of existing estimates of Cagan's model. Furthermore, the machinery of this paper readily supplies an exact, asymptotically efficient procedure for estimating the parameters of the continuous time version of Cagan's model from discrete time data. This exact procedure, which is briefly described in section 5, is a superior alternative to the approximate estimator used by Mohsin Khan [31] in his empirical research which was aimed at answering Benjamin Friedman's [30] questions about time aggregation. The approximate estimator used by Khan, which proceeds by replacing derivatives and levels in continuous time with certain linear combinations of levels and differences in discrete time, has no particular optimal approximating properties for the kind of continuous time inflation, money creation process that makes Cagan's model consistent with rational expectations.

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

(1) $$m(t) - p(t) = \alpha D^{+}\hat{E}_t p(t) + a(t), \ \alpha < 0$$

where p(t) is the logrithm of the price level, m(t) is the logarithm of the money supply, a(t) is a random disturbance to the portfolio balance schedule, $D^{+}$ is the mean square right time derivate operator, and $\hat{E}_t$ is the linear least squares projection operator onto an information set that includes at least current and past observations on p, m, and a. We do not require that the p process be mean square differentiable but only that $\{\hat{E}_t p(t+v): v \geq 0\}$ be mean square differentiable for $v > 0$ and that this process have mean square right derivative at $v = 0$. This right derivative is denoted $D^{+}\hat{E}_t p_t$. 2/

To obtain the solution to stochastic differential equation (1) we write equation (1) shifted ahead v time units as

(2) $$m(t+v) - p(t+v) = \alpha D^{+}\hat{E}_{t+v} p(t+v) + a(t+v).$$

Projecting both sides of (2) onto the t period information set gives

(3) $$\hat{E}_t m(t+v) - \hat{E}_t p(t+v) = \alpha D\hat{E}_t p(t+v) + \hat{E}_t a(t+v)$$

where D is the time derivative operator. 3/ The realizable, time invariant solution to (3) is just

$$\hat{E}_t p(t+v) = -\rho\hat{E}_t \int_0^{\infty} e^{\rho u}[a(t+v+u) - m(t+v-u)]du$$

where $\rho = 1/\alpha$. Taking limits as v declines to zero and noting that

$$\lim_{v \downarrow 0} \hat{E}_t p(t+v) = \hat{E}_t p(t) = p(t),$$

we obtain

$$p(t) = -\rho \hat{E}_t \int_0^{\infty} e^{\rho u}[a(t+u) - m(t+u)]du. \tag{4}$$

as the solution to (1).

We now specialize our assumptions to require that Cagan's adaptive expectations mechanism be optimal. That is, we wish to find specifications for a and m which together with (4) imply that

$$D\hat{E}_t p(t+v) = \beta \int_{-\infty}^{t} e^{-\beta(t-u)} Dp(u)du, \ \beta > 0, \ v > 0. \tag{5}$$

In expression (5) Dp is not necessarily required to be an ordinary stochastic process but rather can be a generalized stochastic process so long as the integral on the right-hand-side of (5) is well defined. 4/ On the other hand, $\{D\hat{E}_t p(t+v) : v > 0\}$ is assumed to be an ordinary stochastic process. Thus even though inflation may not be physically realizable, we assume that anticipated inflation is physically realizable. Equation (5) also implies that at each point in time inflation is expected to be constant over the entire future. This is true since the right-hand-side of (5) does not depend on v.

We assume that the joint process x given by

$$x(t) = \begin{bmatrix} p(t) \\ m(t) \\ a(t) \end{bmatrix}$$

has a time invariant Wold representation

$$x(t) = c(D)w(t). \tag{6}$$

In equation (6) c(D) is a one-sided matrix convolution operator and w is a continuous-time white noise vector with $Ew(t) = 0$ and

$$Ew(t)w(t-v)' = I\delta(t-v)$$

where $\delta$ is the Dirac delta generalized function. [5/] We assume that (6) holds for some t greater than or equal to a start up time T and that $w(t) = 0$ for $t < T$. [6/] The requirement that (6) is a Wold representation implies that instantaneous forecast errors in forecasting an element of x(t) using past x's are a linear combination of elements in w(t).

We write the first row of (6) as

$$p(t) = c_1(D)w(t) \quad . \tag{7}$$

The operator that shifts a time subscript v units ahead can be represented as $e^{vD}$. Therefore, shifting (7) forward v time units and taking expectations we find

$$\hat{E}_t p(t+v) = [c_1(D)e^{vD}]_+ w(t)$$

where $[\ ]_+$ is the annihilation operator that instructs to ignore portions of the convolution operator that are concentrated on the negative numbers. Equation (5) tells us that

$$[Dc_1(D)e^{vD}]_+ = \frac{\beta Dc_1(D)}{\beta+D}$$

for all $v > 0$. [7/] It is verified in the appendix that the solution to this operator equation is

$$c_1(D) = \frac{D+\beta}{D^2} k_0$$

where $k_0$ is an arbitrary row vector constant. We are free to normalize c and w such that

(8) $$p(t) = \frac{(D+\beta)}{D^2}k_1 w_1(t)$$

where $w_1$ is the first element of $w$ and $k_1$ is an arbitrary scalar constant.

Specification (8) implies that

(9) $$D^{+}\hat{E}_t p(t+v) = \frac{\beta}{D}k_1 w_1(t).$$

Substituting (9) into equation (1) we see that

(10) $$m(t) - p(t) = \frac{\alpha\beta D}{(\beta+D)}p(t) + a(t).$$

Equation (10) is a version of Cagan's model in continuous time, since

$$\frac{\beta D}{\beta+D}p(t) = \beta \int_{-\infty}^{t} e^{-\beta(t-u)} Dp(u)du .$$

Equation (10) informs us that there is an exact relationship among m, p, and a. This singularity means that the white noise vector w can have at most two elements. We assume that the x process has maximal rank so that w has two elements.

Partioning w and c we write

(11) $$\begin{bmatrix} p(t) \\ m(t) \\ a(t) \end{bmatrix} = \begin{bmatrix} c_{11}(D) & c_{12}(D) \\ c_{21}(D) & c_{22}(D) \\ c_{31}(D) & c_{32}(D) \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

where

$$c_{11}(D) = \frac{(D+\beta)k_1}{D^2}$$

$$c_{12}(D) = 0.$$

Substituting (11) into equation (1) and equating coefficients on $w_1(t)$ and $w_2(t)$ we see that

$$c_{22}(D) = c_{32}(D)$$

(12)

$$c_{21}(D) - \frac{(D+\beta)k_1}{D^2} = \frac{\alpha\beta k_1}{D} + c_{31}(D) .$$

The stochastic process a is assumed not to be observed by the econometrician. To give equation (1) empirical content we need to say something about the dynamic correlation between a and m. We adopt the requirement that for $v > 0$

(13) $$\hat{E}[a(t+v)|a(u):u \leq t] = \hat{E}_t a(t+v)$$

Assumption (13) says that no other variables observed by private agents Granger cause (help predict) a. It implies that

(14) $$c_{31}(D) = k_2 c_{32}(D) .$$

for some scalar constant $k_2$. 8/ Combining restrictions (14) and (12) we determine that

(15) $$c_{21}(D) = \frac{\alpha\beta k_1}{D} + \frac{(D+\beta)k_1}{D^2} + k_2 c_{22}(D).$$

Restriction (15) is a restriction on the bivariate moving average representation for the observable process

$$\begin{bmatrix} p(t) \\ m(t) \end{bmatrix} = \begin{bmatrix} c_{11}(D) & c_{12}(D) \\ c_{21}(D) & c_{22}(D) \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

where we have previously imposed the restrictions that

(16) $$c_{11}(D) = \frac{(D+\beta)k_1}{D^2}$$

$$c_{12}(D) = 0.$$

An identification question for this model is whether parameters $\alpha$, $\beta$, $k_1$, and $k_2$ can be identified from the continuous-time "reduced form" convolutions $c_{11}$, $c_{12}$, $c_{21}$, $c_{22}$. [9/] It is clear that $\beta$ and $k_1$ can be identified from (16). In general, $\alpha$ and $k_2$ can be identified from equation (15). However for a special and convenient parameterization of $c_{22}(D)$, they are not identified. Suppose that the derivative of a is a white noise. [10/] Thus

(17) $$c_{32}(D) = \frac{k_3}{D} = c_{22}(D).$$

When a is a Gaussian process, (17) implies that a is in fact a Brownian motion. Substituting (17) into (15) yields

(18) $$c_{21}(D) = \frac{\alpha\beta k_1 + \beta k_1 + k_2 k_3}{D} + \frac{\beta k_1}{D^2} .$$

The parameters $\alpha$ and $k_2$ are not identifiable in (18). It remains true, however, even in this case that the model imposes testable cross equation restrictions in that $\frac{\beta k_1}{D^2}$ enters both $c_{11}$ and $c_{21}$.

For the remaining part of this paper, as a simplifying assumption we adopt (17) and require that

$$\alpha\beta k_1 + \beta k_1 + k_2 k_3 = 0.$$

In this case

$$c_{21}(D) = \frac{\beta k_1}{D^2} ,$$

and the moving average representation for p and m is

$$(19)\qquad \begin{bmatrix} p(t) \\ m(t) \end{bmatrix} = \begin{bmatrix} \frac{(D+\beta)k_1}{D^2} & 0 \\ \frac{k_1}{D^2} & \frac{k_3}{D} \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

In the continuous-time system (19), m fails to Granger cause p since $c_{12}(D) = 0$. However, p does Granger cause m. That these features characterize our system is not suprising, since we constructed (19) in order to guarantee that Cagan's adaptive expectations mechanism (5) is optimal. In light of equation (4), if Cagan's mechanism is rational, there must be extensive feedback from p to m.

## 3. Effects of Aggregation Over Time

We are interested in deducing the implications of our continuous-time version of Cagan's model for point-in-time sampled, discrete-time observations on (p, m). We shall assume that point-in-time observations on (p, m) are available at the integers t=0, 1, 2, ... . For convenience we rewrite equation (19)

$$\begin{bmatrix} p(t) \\ m(t) \end{bmatrix} = \begin{bmatrix} \frac{(D+\beta)k_1}{D^2} & 0 \\ \frac{\beta}{D^2}k_1 & \frac{k_3}{D} \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \tag{19}$$

The presence of D and $D^2$ in the denominator of the "moving average" polynomials on the right side of (19) indicates that (p, m) is a nonstationary process. It turns out that the second differences of (p, m) form a stationary 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, m) 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 (19) gives

$$\begin{bmatrix} (1-L)^2 p(t) \\ (1-L)^2 m(t) \end{bmatrix} = \begin{bmatrix} \frac{(1-e^{-D})^2}{D^2}(\beta+D)k_1 & 0 \\ \frac{(1-e^{-D})^2}{D^2}\beta k_1 & \frac{(1-e^{-D})^2}{D}k_3 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

Now recall the following Laplace transform pairs: 11/

$$\frac{(1-e^{-s})^2}{s} \leftrightarrow \begin{cases} 1 & t\varepsilon[0,1] \\ -1 & t\varepsilon[1,2] \\ 0 & t > 2 \end{cases} \tag{20}$$

$$(21) \qquad \frac{(1-e^{-s})^2}{s^2} \leftrightarrow \begin{cases} t & t\epsilon[0,1] \\ 2-t & t\epsilon[1,2] \\ 0 & t > 2. \end{cases}$$

Using the Laplace transforms (21) and (20) gives the desired representation:

$$(22) \qquad \begin{aligned} (1-L)^2 p(t) &= k_1\int_0^1 (\beta\tau+1)w_1(t-\tau)d\tau + k_1\int_1^2 [\beta(2-\tau)-1]w_1(t-\tau)d\tau \\ (1-L)^2 m(t) &= k_1\int_0^1 \beta\tau w_1(t-\tau)d\tau + k_1\int_1^2 \beta(2-\tau)w_1(t-\tau)d\tau \\ &\quad + k_3\int_0^1 w_2(t-\tau)d\tau - k_3\int_1^2 w_2(t-\tau)d\tau. \end{aligned}$$

To represent things compactly, we define

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

Then we can write (16) as

$$(23) \qquad \begin{aligned} y(t) &= \int_0^1 \begin{bmatrix} (\beta\tau+1)k_1 & 0 \\ \beta\tau k_1 & k_3 \end{bmatrix} \begin{bmatrix} w_1(t-\tau) \\ w_2(t-\tau) \end{bmatrix} d\tau \\ &\quad + \int_0^2 \begin{bmatrix} [\beta(2-\tau)-1]k_1 & 0 \\ \beta(2-\tau)k_1 & -k_3 \end{bmatrix} \begin{bmatrix} w_1(t-\tau) \\ w_2(t-\tau) \end{bmatrix} d\tau. \end{aligned}$$

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

$$\Gamma_0 = Ey(t)y(t)'$$

$$= \int_1^2 \begin{bmatrix} \beta\tau+1 & 0 \\ \beta\tau & 1 \end{bmatrix} \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix} \begin{bmatrix} \beta\tau+1 & \beta\tau \\ 0 & 1 \end{bmatrix} d\tau$$

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

and

$$\Gamma_1 = Ey(t)y(t-1)'$$

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

and

$$\Gamma_{-1} = Ey(t)y(t+1)' = \Gamma_1'$$

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

where $v_{11} = (k_1)^2$ and $(k_3)^2$. Evaluating the above integrals, we obtain

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

(24)

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

The matrix covariogram (24) of the discrete time process $y_t$ contains all of the information required to compute the Wold moving average representations in discrete time. By studying the univariate and bivariate discrete time Wold representations for $y_t$, we are able to characterize the effects of aggregation over time. This is accomplished in the following two sections.

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

We first consider the univariate Wold representation for the $(1-L)^2p$ process. From (24), $(1-L)^2p$ is a first-order moving average with covariance generating function

$$g(z) = c(1)z^{-1} + c(0) + c(1)z \tag{25}$$

where from (19) $c(0) = 2v_{11}(\frac{1}{3}\beta^2+1)$, $c(1) = v_{11}(\frac{1}{6}\beta^2-1)$. We seek the Wold moving average representation for $(1-L)^2p$, which is of the form

$$(1-L)^2p(t) = (1-\lambda_p L)\varepsilon_{pt}, \quad |\lambda_p| < 1 \tag{26}$$

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

$$\lambda_p = -\frac{1}{2}\frac{c(0)}{c(1)} \pm \sqrt{\frac{c(0)^2}{4c(1)^2} - 1} \tag{27}$$

subject to $|\lambda_p| < 1$

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

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

$$\lambda_p = -\frac{(\frac{1}{3}\beta^2+1)}{\frac{1}{6}\beta^2-1} \pm \sqrt{\frac{(\frac{1}{3}\beta^2+1)^2}{(\frac{1}{6}\beta^2-1)^2} - 1} \tag{28}$$

subject to $|\lambda_p| < 1$.

Now consider the discrete-time inflation rate $X$ which we define as $X(t) = p(t) - p(t-1)$ for $t$ at the integers. Representation (26) can then be written

(29) $$(1-L)X(t) = (1-\lambda_p L)\varepsilon_{pt}.$$

As shown by John F. Muth [13], the optimal j-step-ahead forecast of X governed by process (29), given current and lagged values of X alone, is the discrete-time version of Cagan's adaptive expectation schemes

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

Now equation (30) 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 $\beta$ via the equation

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

For various values of $\beta$, Table 1 reports the values of $\lambda_p$ given by formula (28) and Cagan's formula (31). For $\beta$ close to zero, equation (31) provides a close approximation to (28). However, for large values of $\beta$, $\exp(-\beta)$ is approximately zero, while equation (28) 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, at integer points in time. 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, 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 (27), if he followed Nerlove, Grether, and Carvalho's method. Using formula (28) or Table 1, the analyst could then infer the value of $\beta$. Table 1 provides a fairly complete characterization of Cagan's approximation (31) 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 $((1-L)^2 p(t),(1-L)^2 m(t))' = y(t)$ is

$$y(t) = u_t + Fu_{t-1} \tag{32}$$

where $u_t$ is a (2x1) vector discrete-time white noise with $Eu_t u_t = \overline{V}$, where $\overline{V}$ is a positive semidefinite matrix; $u_t = y(t) - \hat{E}y(t)|t(t-1),y(t-2),\ldots$; and the eigenvalues of F are less than or equal to unity in absolute value. Given $\Gamma_0$ and $\Gamma_1$ from (24), F and $\overline{V}$ are determined by solving the following equations

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

or

$$\Gamma_0 = \overline{V} + F\overline{V}F'$$

$$\Gamma_1 = F\overline{V}. \tag{33}$$

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 (32) as

$$\begin{bmatrix} (1-L)X_t \\ (1-L)M_t \end{bmatrix} = u_t + Fu_{t-1}\ . \tag{34}$$

By carrying out a series of calculations paralleling those of Muth [13], it is straightforward to verify that (34) 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

$$(35) \qquad \begin{bmatrix} X_{t+1} \\ M_{t+1} \end{bmatrix} = (I+F) \sum_{i=0}^{\infty} (-F)^i \begin{bmatrix} X_{t-i} \\ M_{t-i} \end{bmatrix} + u_{t+1}.$$

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:

$$(36) \qquad \hat{E} \begin{bmatrix} X_{t+j} \\ M_{t+j} \end{bmatrix} | X_t, M_t, X_{t-1}, M_{t-1}, \ldots \Bigg] = (I+F) \sum_{i=0}^{\infty} (-F)^i \begin{bmatrix} X_{t-i} \\ M_{t-i} \end{bmatrix}, \; 2 \geq 1.$$

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

Representation (34) 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

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

where $(\epsilon_1, \epsilon_2)' = \epsilon$ is a discrete-time vector white noise with arbitrary contemporaneous covariance matrix $E\epsilon_t\epsilon_t = W$; $\epsilon$ is fundamental for $((1-L)X, (1-L)M)$; and $|\lambda| < 1$. It is evident from the first equation of (37) that Cagan's discrete-time adaptive expectations formulation for inflation is rational, given (37). 14/

In form, (37) matches (34). One of our tasks now is to study the relation between the (2x2) matrix F in (34) and the corresponding matrix

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

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 $\beta$ and V where

$$V = \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix},$$

we have calculated F and $\overline{V}$. In addition, we calculated $\lambda_p$ and $\sigma^2_{\epsilon p}$ in the univariate Wold moving average representation for $(1-L)X$:

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

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

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

where $\epsilon_m$ is a fundamental white noise for (1-L)M and $\sigma^2_{\epsilon m} = E\epsilon^2_m$ is the one-step-ahead prediction error variance for (1-L)M. 16/ Recall that $\bar{v}_{11}$ is the discrete-time, one-step-ahead prediction error variance for predicting X on the basis of lagged X's and lagged M's, while $\bar{v}_{22}$ is the discrete-time, one-step-ahead prediction error variance in predicting M on the basis of lagged X's and lagged M's. Therefore, $(\sigma^2_{\epsilon p}-\bar{v}_{11})/\sigma^2_{\epsilon p}$ is a measure of the marginal assistance of lagged M's in predicting x, while $(\sigma^2_{\epsilon m}-\bar{v}_{22})/\sigma^2_{\epsilon m}$ is a measure of the marginal assistance of lagged X's in predicting M. 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 (19), which we are maintaining, M fails to Granger cause X. However, in the discrete-time model, M will in general Granger cause X due to the effects of aggregation over time. 17/ The percentage gain $(\sigma^2_{\epsilon p}-\bar{v}_{11})/\sigma^2_{\epsilon p}$ 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 $\bar{V}$, F, $\sigma^2_{\epsilon p}$, $\sigma^2_{\epsilon m}$, $\lambda_p$, and $\lambda_m$ for three values of $\beta$, and for three settings for the "intensity" matrix V. Tables 5-7 give less complete characterizations of $\bar{V}$ and F for a large number of values of $\beta$.

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

Table 2

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

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

$\sigma^2_{\varepsilon p} = 1.05084$; % 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

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

$\sigma^2_{\varepsilon p}$ = 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

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

$\sigma^2_{\varepsilon p}$ = 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

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

$\sigma^2_{\varepsilon p}$ = 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

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

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

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

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

eigenvalues of F: -1.0, -.951225

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

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

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

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

eigenvalues of F: -1.0, -.0690102

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

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

$\sigma^2_{\varepsilon p} = 65.5079$; % 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 | .860537 | -.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 | .182532 |
| 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 |

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

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

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

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

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

For value 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 $\beta$. [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 $\beta$ the effects of aggregation over time can be considerable. In particular, while 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 (36) is a bivariate moving average representation for $\{(1-L)X, (1-L)M\}$, 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.

## 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.

## Appendix

In this appendix we use tools discussed in Hansen and Sargent [28] to show that the only choice of $c_1(D)$ that has a rational Laplace transform and that satisfies

$$\text{(A1)} \qquad [Dc_1(D)e^{vD}]_+ = \frac{\beta Dc_1(D)}{\beta+D}$$

is

$$\text{(A2)} \qquad c_1(D) = \frac{D+\beta}{D^2}k_0$$

where $k_0$ is an arbitrary row vector constant.

The fact that no other variables Granger cause p implies that $c_1(D)$ must take the form

$$c_1(D) = c_1(D)^* k_0$$

where $c_1(D)^*$ is a scalar operator and $k_0$ is a row vector constant that satisfies $k_0 k_0' = I$. Thus we can write the fundamental representation for p as

$$\text{(A3)} \qquad p(t) = c_1(D)^* w(t)^*$$

where $w(t)^* = k_0 w(t)$. The function $c_1(s)^*$ is assumed to be rational which we represent

$$c_1(s)^* = \frac{\mu(s)}{\gamma(s)}$$

where

$$\gamma(s) = (s-\gamma_1)(s-\gamma_2)\ldots(s-\gamma_n)$$

$$\mu(s) = \mu_0(s-\mu_1)(s-\mu_2)\ldots(s-\mu_m).$$

Consistent with the requirements mentioned in footnote 6, we impose the restriction that Real $(\gamma_j)$ is less than or equal to zero for $j = 1, 2, \ldots, n$. This guarantees that the vector function $c_1^*(s)$ is analytic in the open right half plane. The assumption that $p$ is physically realizable implies that $m < n$.

The function

$$G(s,v) = sc_1(s)^* e^{vs}$$

is analytic in its first argument everywhere in the complex plane except possibly at $\gamma_1, \gamma_2, \ldots, \gamma_n$. Let $H_1(s,v)$, $H_2(s,v), \ldots$, $H_q(s,v)$ denote the principal parts of G at the corresponding q distinct zeroes of $\gamma(s)$. Using a result from Hansen and Sargent [28], it follows that

$$[G(D,v)]_+ = H_1(D,v) + H_2(D,v) + \ldots + H_q(D,v).$$

However from (A1) it is clear that $H_j(s,v)$ cannot depend on v. It follows that the $n = 2$ and $\gamma_1 = \gamma_2 = 0$. Therefore

$$c_1(s)^* = \frac{\mu_0(s-\mu_1)}{s^2} .$$

Computing $[Dc_1(s)^* e^{vD}]_+$ we obtain

$$\left[\mu_0\frac{(D-\mu_1)e^{vD}}{D}\right]_+ = \frac{-\mu_0\mu_1}{D} .$$

By equation (A1) we see that

(A4) $$\frac{-\mu_0\mu_1}{D} = \frac{\beta\mu_0(D-\mu_1)}{(\beta+D)D} .$$

Equation (A4) implies that $\mu_1 = -\beta$. Therefore

$$c_1(D)^* = \mu_0\frac{(D+\beta)}{D^2}$$

which proves the desired result.

## Footnotes

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

2. We interpret $D^{+}E_t p_t$ as expected inflation at time period t. For a more detailed discussion of linear expectations differential equations of the form given in (1) see Hansen and Sargent [28]. The approach adopted by Hansen and Sargent avoids making distributional assumptions and instead focuses on the covariance properties of the underlying vector time series process. It is assumed that this vector process has a time invariant moving average representation in which each of the components can be expressed as a one-sided convolution integral of a vector white noise. An alternative approach to the one taken here is to assume that underlying vector time series process can be characterized as a diffusion process and to investigate what restrictions (1) implies on this diffusion.

3. The interchange of the derivative and linear least squares projection operators is justified in Hansen and Sargent [28].

4. For a discussion of generalized stochastic processes see Gelfand and Vilenkin [26]. Generalized stochastic processes are oftentimes convenient mathematical devices for modelling processes that are not directly observed but whose properties are inferred from observations on integral averages of these processes. In our example, we assume that prices are observed and the properties of inflation can be inferred from these observations. We do not require that inflation be an ordinary stochastic process.

5. Rozanov [19] uses the concept of a fundamental representation in the context of covariance stationary stochastic processes. This notion is generalized in Rozanov [29] to apply to nonstationary processes as well. In this more general framework Rozanov refers to such a representation as a canonical representation. We restrict ourselves to processes in which the corresponding convolution operators used to define the fundamental representations are one-sided and have Laplace transforms that are analytic in the open right half plane. We shall represent convolution operators as functions of the derivative operator. 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].

6. We adopt this start up time interpretation in order that we can accommodate certain borderline non-covariance-stationary processes. The assumption that $T = -\infty$ is appropriate only if we restrict ourselves to covariance stationary processes.

7. This operator equation has implicitly interchanged the linear least squares projection operator and the derivative operator. See Hansen and Sargent [28] for a justification of this operator equation.

8. This identification question ignores the problem of identifying the continuous-time reduced form parameters from discrete-time data, i.e., the aliasing phenomenon. See Hansen and Sargent [27] for a discussion of aliasing in the context of linear rational expectations models. Christiano [25] has an extensive discussion of identification in the context of a discrete-time version of Cagan's model. Christiano's discussion is closely related to the comments about identification in this paper.

9. Equation (14) is an implication of the facts that the instantaneous forecast error in forecasting a from its own past is a linear combination of w and that $c_{32}(D) \neq 0$.

10. This is a continuous-time version of the assumption about the disturbance to the portfolio balance schedule considered by Sargent [20].

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, (1-L)M is triangular, implying that (1-L)M fails to Granger cause (1-L)X. See Sims [23].

15. First, note that

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

and 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 $\det(\Gamma_1' z^{-1}+\Gamma_0+\Gamma_1 z)$, 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].

## 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).

[25] Christiano, L., 1980, "Rational Expectations, Hyperinflation, and the Demand for Money," manuscript.

[26] Gelfand, I. M., and N. Y. Vilenkin, 1964, Generalized Functions, Volume 4, (Academic Press, New York).

[27] Hansen, L. P., and T. J. Sargent, 1980, "Rational Expectations and the Aliasing Phenomenon," manuscript.

[28] Hansen, L. P., and T. J. Sargent, 1981, "Notes on Continuous Time Linear Rational Expectations Models," manuscript.

[29] Rozanov, Y. A., 1977, Innovation Processes, (John Wiley & Sons, New York).

[30] Friedman, Benjamin M., "Stability and Rationality in Models of Hyperinflation," International Economic Review, 19 (February, 1978), 45-64.

[31] Khan, Mohsin S., "Dynamic Stability in the Cagan Model of Hyperinflation," International Economic Review, 21 (October, 1980), 577-582.