The Term Structure of Interest Rates and the Aliasing Identification Problem

Lawrence J. Christiano*

November 1980

Working Paper #165

PACS File #2700

Columbia University
and
Federal Reserve Bank of Minneapolis

Abstract:

Theory typically does not give us reason to believe that economic models ought to be formulated at the same level of time aggregation at which data happen to be available. Nevertheless, this is frequently done when formulating econometric models, with potentially important specification-error implications. This suggests examining the alternatives, one of which is to model in continuous time. The primary difficulty in inferring the parameters of a continuous time model given sampled observations is the "aliasing identification problem". This paper shows how the restrictions implied by rational expectations sometimes do, and sometimes do not, resolve the problem. This is accomplished very simply in the context of a hypothesis about the term structure of interest rates. The paper confirms and extends results obtained for another example by Hansen and Sargent.

*I am indebted to Lars Hansen for suggesting many of the ideas in this paper.

I. Introduction.

Theory typically does not give us reason to believe that economic models ought to be formulated at the same level of time aggregation at which data happen to be available. Nevertheless, this is frequently done when formulating econometric models. On the other hand, the choice of the level of time aggregation may have important specification-error implications. Two examples arising out of the work of Lucas and Sargent and the work of Sims and Geweke illustrate this point.

The work of Lucas(1976), Sargent(1980), Lucas and Sargent(1978,1980), and others, implies that the parameters of private agents' decision rules are functions of the parameters of related processes, including government policy variables, which agents take into account in making decisions. The implication of this is that if the government is to successfully evaluate a planned change in policy, then it must first determine how a change in the parameters that describe the laws of motion of its policies translates into a change in the parameters of agents' decision rules. Under linearity of the stochastic processes and the assumption of rational expectations, the functional relationship between the parameters of different equations ("cross-equations restrictions") can often be derived analytically. If one specifies private agents to be solving a continuous time optimum problem, then the implied cross-equations restrictions for the discrete time representation of agents' decision rules and government policies in general differs from what they are when the optimum problem is formulated in discrete time. Thus, if a policy advisor commits a specification error and models agents as maximizing a discrete time problem, whereas in fact they optimize a continuous time problem, then the government will choose a sub-optimal policy. This is because the specification error leads the advisor to use the wrong cross-equations restrictions when computing private agents' decision rules as functions of alternative government policies.

Further evidence on the importance of the choice of the level of time aggregation is supplied by the work of Sims(1971) and Geweke(1978). They show that if a given pattern of Granger(1969)-causality is specified to exist at the level of continuous time, then the same pattern of Granger-causality may not be predicted for sampled data. To illustrate the significance of this for the way economic time series are interpreted, consider the following example. Sargent and Wallace(1973) and Sargent(1977) show that a discrete time version of Cagan's(1956) model of hyperinflation under rational expectations predicts that inflation Granger-causes money creation with no feedback. When the same model is reformulated in continuous time, then Sargent and Wallace's Granger-causality pattern obtains at the level of continuous time. However, the work of Sims and of Geweke implies that the prediction that money creation fails to Granger-cause inflation in the sampled data no longer holds. This example suggests that the predictions of a theory for sampled data depends, in part, on the specification of the level of time aggregation.

These considerations suggest examining alternatives to the widespread practice of formulating estimable models at the same level of time aggregation at which data happen to be available. One alternative is modelling in continuous time. That this alternative should be taken seriously is suggested by the fact that when models are formulated without an eye to empirical estimation, they are often formulated in continuous time. Examples are the work of Lucas(1965,1966,1967), Mortensen(1973), Treadway(1969,1970) and Gould(1968), to name just a few.

There are some practical advantages to modelling in continuous---as opposed to discrete---time. First, in continuous time the econometrician has greater flexibility in making optimal use of possible idiosyncrasies in the available data. Examples are cases where the data sampling interval shifts during the observation period, when one variable is sampled more frequently than another, and when the available data are period averages rather than point-in-time observations.[1/] Second, modelling in continuous time makes possible optimally forecasting over time intervals finer than those separating the available data.

This could, for example, be of use to macroeconomic forecasters, for whom key data are available on a quarterly basis only. Third, continuous time modelling permits one to interpolate optimally between published data series.

A problem that the econometrician must solve if he or she is to estimate the parameters of a continuous time model given sampled data only is the aliasing identification problem. Hansen and Sargent(1980b) were the first to show that the cross-equations restrictions implied by rational expectations could resolve the problem.[2/] They demonstrate this in the context of a firm's demand for a factor of production. In this paper the aliasing identification problem is studied in the context of two term structure models of interest rates under rational expectations. The term structure examples have the advantage of illustrating, in a simple way, the following fact: The cross-equations restrictions implied by the hypothesis of rational expectations sometimes do, and sometimes do not, resolve the aliasing identification problem. (The first point was established by Hansen and Sargent(1980b).) The term structure models studied are versions of the one proposed in Hicks(1946,p.145) and later studied by Sargent(1972,1979). Both models include two interest rates: a short rate and long rate. We consider the case where these can be modelled as a bivariate first order stochastic differential equation. It will be shown that, in this context, when the short rate is a "call rate", then the rational expectations restrictions overcome the aliasing problem. (By a "call rate" we mean a rate of interest on an asset with instantaneous maturity period and with a return which is compounded continuously.) When the short rate has a nonzero maturity period, then the cross-equations restrictions are of no use in resolving the aliasing identification problem, when it exists.

The plan of the paper is as follows. The nature of the aliasing identification problem is described in section II. In section III the relation between the cross-equations restrictions implied by the term structure hypothesis under rational expectations and the aliasing problem is discussed. The findings of the paper are summarized in part IV.

II. The Aliasing Identification Problem.

Consider the following first order stochastic differential equation

$$\text{(1)} \quad Dy(t) = Ay(t) + u(t),\ Eu(t) = 0,\ Eu(t)u(t-\tau) = \delta(\tau)V.$$

Here, $\delta(\cdot)$ is the delta generalized function[3/]; the disturbance process $\{u(t)\}$ is a continuous time Gaussian white noise with spectral density V at all frequencies[4/]; V is 2×2, symmetric positive definite; and A is 2×2 nonsingular. The operator D is the time derivative operator $\frac{d}{dt}$. Write

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix}, \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A = T\Lambda T^{-1},$$

where,

$$\Lambda = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}.$$

The scalars ρ_1 and ρ_2 are eigenvalues of A and we assume that $Re(\rho_i) < 0$ i=1,2. The i^{th} column of the nonsingular matrix T is the eigenvector corresponding to ρ_i. Let the covariance function of the $\{y(t)\}$ process be $R(\tau) = Ey(t)y(t-\tau)^T$, where the superscript 'T' denotes transposition. Then

$$(3) \qquad R(\tau) = W_1 e^{\rho_1 \tau} + W_2 e^{\rho_2 \tau} \quad \tau \geq 0,$$

where,

$$W_1 = \frac{(\rho_1 I - A)^a V[(-\rho_1 I - A)^a]^T}{-2\rho_1(\rho_1 - \rho_2)(-\rho_1 - \rho_2)}$$

$$W_2 = \frac{(\rho_2 I - A)^a V[(-\rho_2 I - A)^a]^T}{-2\rho_2(\rho_2 - \rho_1)(-\rho_2 - \rho_1)} .$$

Here, the superscript 'a' denotes the adjoint operator.[5/]

Given the assumptions of normality and a zero mean, the $R(\cdot)$ function in (3) completely characterizes the distribution of $\{y(t)\}$. Furthermore, it can be shown that the mapping from the seven parameters of (A,V) to the parameters of $R(\cdot)$ is one-to-one.[6/] Consequently, the aliasing identification problem may be described either with reference to (A,V) or, equivalently, with reference to the function $R(\cdot)$. Consider first the function $R(\cdot)$.

1. <u>A first approach to the aliasing problem.</u>

Suppose that ρ_1 and ρ_2 are complex (and, hence, conjugate, since A is real). Then define

$$(4) \qquad R^{(\nu)}(\tau) = W_1 e^{(\rho_1 + 2\pi i\nu)\tau} + W_2 e^{(\rho_2 - 2\pi i\nu)\tau} \quad \tau \geq 0 ,$$

for integer values of ν. The graph of the covariance function $R^{(\nu)}(\tau)$ differs from that of $R(\tau)$ for most values of τ when $\nu \neq 0$. ($R^{(0)}(\tau) = R(\tau)$ for all τ.) Consequently, the points of the previous paragraph apply, and the parameters $(A^{(\nu)},V^{(\nu)})$ associated with $R^{(\nu)}(\tau)$ must differ from (A,V) unless $\nu=0$. Notice, however, that $e^{2\pi i\nu\tau}=1$ for integer values of τ. Hence,

$$R^{(\nu)}(\tau) = R(\tau) \text{ for } \tau=0,\pm1,\pm2,\ldots \tag{5}$$

and for arbitrary integer ν.[7/]

Under the assumptions of normality and a zero mean, estimating the parameters of a time series model amounts to fitting a theoretical covariance function to the sample covariance function. In this case the sample covariance function of observations on {y(t)} exhausts the information in the data concerning the parameters of (1). If information on {y(t)} is restricted to sampled observations---ie., to sampled observations on the covariance function---then, as (5) suggests, the data cannot in general be used to discriminate between theoretical covariance functions $R^{(\nu)}(\tau)$ with different settings for ν.[8/] If, on the other hand, the sample covariances do not display oscillations, then the method of maximum likelihood generates a theoretical covariance function which also does not oscillate, ie., an A matrix with real roots. In this case, trivially, there is no aliasing identification problem. Furthermore, if (1) where a scalar process, then there is only one root (A itself), which must therefore be real. Thus, in the scalar case the covariance function of (1) also cannot oscillate and therefore there is no aliasing problem in this case either. Solving the aliasing identification problem for the general case when roots are permitted to be complex requires placing enough *a priori* restrictions on the parameters of A and V to reduce the set of admissible ν in (5) to the set $\nu=0$. When this is the case, sampled observations on a covariance function restrict the inter-sample oscillations on a theoretical covariance function.

While the foregoing discussion is useful for providing intuition into the nature of the aliasing identification problem, another approach is more practical for the purposes of this paper. This is the one taken in P.C.B.Phillips(1973) and in Hansen and Sargent(1980b) and is the one we adopt in the remainder of the paper.

2. A second approach.

It can be shown that the solution to (1) conditioned on $y(t) = y(t_0)$ at $t=t_0$ is

$$(6) \qquad y(t) = e^{A(t-t_0)} y(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} u(\tau)d\tau .$$

Suppose we have observations on $\{y(\gamma h)\}$ for integer values of γ, where h (>0) is the sampling interval. We obtain a representation for the sampled $\{y(t)\}$ process by replacing t_0 in (6) by t-h

$$(7) \qquad \begin{aligned} y(t) &= e^{Ah} y(t-h) + \int_{t-h}^{t} e^{A(t-\tau)} u(\tau)d\tau \\ &= By(t-h) + v(t), \end{aligned}$$

where,

$$(8) \qquad B = e^{Ah} = Te^{\Lambda h}T^{-1} \quad \underline{9/},$$

and $v(t) = \int_{t-h}^{t} e^{A(t-\tau)} u(\tau)d\tau = \int_0^h e^{A\tau} u(t-\tau)d\tau$. The disturbance v(t) is a discrete time white noise with covariance matrix W, where

$$W = \int_0^h e^{A\tau} V e^{A^T\tau} d\tau . \tag{9}$$

Phillips(1973) has shown that the following relation obtains between W, V, and A:

$$\mathrm{vec}(V) = \{e^{Ah} \otimes e^{A^T h} - I\otimes I\}^{-1}\{A\otimes I + I\otimes A^T\}\mathrm{vec}(W), \tag{10}$$

where vec(·) denotes the column-by-column matrix vectorization operator and '⊗' denotes the Kronecker product.[10/]

As was pointed out previously, the seven parameters of (A,V) completely characterize the distribution of the continuous time {y(t)} process. Similarly, well-known results in time series analysis indicate that the seven parameters of (B,W) completely characterize the distribution of the sampled {y(t)} process (equation (7)). In this context, if there exists only one set of parameters (A,V) corresponding to (B,W), then (1) is said to be identified in the aliasing sense. As was pointed out earlier, this is not the case when the roots of A are complex. Suppose that ρ_1 and ρ_2 are complex, that they do not not differ by an integer multiple of $\frac{2\pi i}{h}$, and define

$$\Lambda_k = \Lambda + \frac{2\pi i}{h} kP, \quad A_k = T\Lambda_k T^{-1}, \quad B_k = e^{A_k h} , \tag{11a}$$

for $k=0,\pm 1,\pm 2,\ldots$. Here, $P = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Following Phillips(1973,p.355), (11a) implies that $A_k = A + \frac{2\pi i}{h}kTPT^{-1}$ and $A_0 = A$, $\Lambda_0 = \Lambda$. We shall sometimes refer to A as a "perturbation on the elements of A". Exploiting properties of the exponentiation operator,

$$B_k = Te^{\Lambda_k h}T^{-1} = Te^{\Lambda h}T^{-1} = e^{Ah} = B.$$

Thus, without restrictions on A, any given B corresponds to a countable infinity $\{A_k\}_{k=-\infty}^{\infty}$ via (8). Corresponding to every A_k, one can find a V_k that solves (10) for a given W. In this way one can think of finding a countable infinity, $\{A_k,V_k\}_{k=-\infty}^{\infty}$, of solutions to (8) and (10) for given (B,W). However, Hansen and Sargent(1980b) have shown that, except for singular cases, the number of values of k for which V_k is positive semidefinite is finite. This is the message of their theorem 3 (the superscript 'H' denotes transposition and conjugation):

<u>Theorem 3.</u> If $T^{-1}W(T^{-1})^H$ has no zero elements, then for given (B,W) there is, at most, a finite number of distinct pairs (A_k,V_k) that satisfy

(i) $e^{A_k h} = B$

(ii) $\int_0^h e^{A_k \tau} V_k e^{A_k^T \tau} d\tau = W$

(iii) V_k is positive semidefinite.

<u>Proof:</u> (See Hansen and Sargent(1980b).)

A way to resolve the aliasing problem is to restrict the elements of A in such a way that perturbing its eigenvalues by integer multiples of $2\pi i$ forces changes in its eigenvectors. P.C.B.Phillips(1973) shows how linear restrictions on A can accomplish this. Hansen and Sargent(1980b) present an example in which the rational expectations cross-equations restrictions accomplish this. We shall see that the Hansen-Sargent result obtains in the context of the term structure when the short rate is a call rate. The result does not obtain, however, when the short rate applies to an asset with a nonzero maturity period. We consider these two cases in the next section.

III. The Term Structure.

In this section we begin by presenting a hypothesis about the term structure of interest rates. We then generate two examples that are consistent with the hypothesis and which illustrate important features of the relation between the rational expectations cross-equations restrictions and the aliasing identification problem. Define,

(12)

$r_n(t)$ - return, per period of time, on an n period bond (the "long rate"), n is rational,

$r^N(t)$ - return, per period of time, on a $\frac{1}{N}$ ($\frac{1}{N}$<n) period bond (the "short rate"), Nn is an integer.

Let the units in which we measure time, t, be the "period of time" in (12) and restrict h to be a rational number. For example, when N=4, n=5, the sampling interval is one month and the period is one year, then $h = \frac{1}{12}$, and $r^N(t)$ and $r_n(t)$ represent returns at an annual rate, on bonds which mature every quarter and 5 years, respectively. Define a call rate as a rate of interest on a bond with instantaneous maturity, ie., $r^\infty(t) = \lim_{N\to\infty} r^N(t)$. Following Hicks(1946,p.145), Sargent(1972,1979a) hypothesized the following relationship between

$r_n(t)$ and $r^N(t)$:

$$(13) \qquad (1 + r_n(t))^n = (1 + \frac{r^N(t)}{N})(1 + E_t \frac{r^N(t+\frac{1}{N})}{N}) \times (1 + E_t \frac{r^N(t+\frac{2}{N})}{N}) \cdots (1 + E_t \frac{r^N(t+\frac{nN-1}{N})}{N}) ,$$

where $E_t(\cdot) \equiv \hat{E}(\cdot | r^N(t-s), r_n(t-s);\ s \geq 0)$ and $\hat{E}(\cdot)$ is the linear least squares projection operator.11/ Taking the natural logarithm of (13) and making use of the approximation $\ln(1+x) \approx x$ for small x, we get

$$(14) \qquad r_n(t) = \frac{1}{nN} E_t \sum_{i=0}^{nN-1} r^N(t+\frac{i}{N}) .$$

If we assume that $(r^N(t), r_n(t))$ is linearly regular, has mean zero and is covariance stationary, then by a continuous time version of Wold's theorem, the $(r^N(t), r_n(t))$ process can be modelled as a stationary stochastic differential equation.12/ For the sake of the illustration, assume in addition that $(r^N(t), r_n(t))$ can be represented as a first order stochastic differential equation like (1). Accordingly, interpret the process $\{y(t)\}$ studied in the last section $y(t) = (r^N(t), r_n(t))$.

Define $U = (1,0)$, so that $UE_t y(t+\frac{\tau}{N}) = E_t Uy(t+\frac{\tau}{N}) = E_t r^N(t+\frac{\tau}{N})$. Then (14) may be rewritten

$$(15) \qquad r_n(t) = \frac{1}{nN} U E_t \sum_{\tau=0}^{nN-1} y(t+\frac{\tau}{N}) .$$

A standard result in the theory of linear least squares projections is

$$(16) \qquad E_t y(t+\tfrac{\tau}{N}) = e^{\frac{\tau}{N}A} y(t) \text{ for } \tau \geq 0$$

when $\{y(t)\}$ is as it is specified in (1). Substituting (16) into (15), we get

$$(17) \qquad \begin{aligned} r_n(t) &= \frac{1}{nN} U \sum_{\tau=0}^{nN-1} e^{\frac{\tau}{N}A} y(t) \\ &= U\frac{1}{n}[N(I-e^{\frac{1}{N}A})]^{-1}[I-e^{nA}]y(t) \ . \end{aligned}$$

From (17) we see that, for finite N, rational expectations implies

$$(18) \qquad U\frac{1}{n}[N(I-e^{\frac{1}{N}A})]^{-1}[I-e^{nA}] = (0,1) \ .$$

Note that $N(I-e^{\frac{1}{N}A}) = -A - \frac{1}{2N}A^2 - \frac{1}{3!N^2}A^3 - \ldots$, so that $N(I-e^{\frac{1}{N}A}) \rightarrow -A$ as $N\rightarrow\infty$. It follows that when the short rate is a call rate, the restrictions imply

$$(19) \qquad U\frac{1}{n}A^{-1}[e^{nA}-I] = (0,1).$$

Equations (18) and (19) are the cross-equations restrictions implied by the term structure hypothesis, (13).

We proceed to consider first the case of finite N and then the case of the call rate.

1. <u>The case where the short rate is a return on an asset with non-zero maturity period.</u>

Suppose A satisfies (18). Note that the product Nh is rational since N and h are.[13/] Because of this we can write $Nh = \frac{b}{c}$, where b and c are integers and b is not an integer multiple of c unless c=1. Let $k = \gamma cNh$, where γ is an arbitrary non-zero integer, and consider a perturbation of A, A_k, as defined in (11). (Note that by construction k is an integer, as (11) requires.) It is easy to verify that $e^{nA_k} = e^{nA}$ because cnN is an integer (see (12)). Using this fact and substituting A_k into (18), we get

(20)
$$\begin{aligned} &U\tfrac{1}{n}[N(I - e^{\frac{1}{N}A_k})]^{-1}[I - e^{nA_k}] \\ &= U\tfrac{1}{n}[N(I - Te^{\frac{1}{N}\Lambda_k}T^{-1})]^{-1}[I - e^{nA}] \\ &= U\tfrac{1}{n}[N(I - T^{\frac{1}{N}\Lambda}T^{-1})]^{-1}[I - e^{nA}] \\ &= U\tfrac{1}{n}[N(I - e^{\frac{1}{N}A})]^{-1}[I - e^{nA}] \\ &= (0,1) \ , \end{aligned}$$

where the last equality holds by hypothesis. The second equality makes use of the relation $\frac{1}{N}\Lambda_k = \frac{1}{N}\Lambda + 2\pi i\gamma cP$ and the fact that γc is an integer. Thus, for finite N, perturbations on the elements of A, A_k, have been found that satisfy the rational expectations restrictions if A does.

Suppose that we have a (B,W) in hand (say these have been consistently estimated from sampled observations on $y(t)$). Then, by (20), there is a countable infinity of admissible solutions, $\{A_{\gamma cNh}\}_{\gamma=-\infty}^{\infty}$, to (8).[14/] Corresponding to each admissible A_k, one can compute a V_k that solves (10) for the given W. As Hansen and Sargent(1980b) show, outside of singular cases the number of values of k for which V_k is positive semidefinite is finite.

As an example of a singular case, consider $V = TT^H$. The hypothesis of theorem 3 is not satisfied since

$$T^{-1}W(T^{-1})^H = T^{-1}\int_0^h e^{A_k\tau}Ve^{A_k^T\tau}\,d\tau(T^{-1})^H$$

(21)
$$= T^{-1}\int_0^h Te^{\Lambda_k\tau}T^{-1}TT^H(T^{-1})^He^{\Lambda_k^*\tau}T^Hd\tau(T^{-1})^H$$

$$= \int_0^h e^{(\Lambda_k+\Lambda_k^*)\tau}d\tau$$

$$= \int_0^h e^{(\Lambda+\Lambda^*)\tau}d\tau \quad ,$$

which is diagonal. (Here, the superscript '*' denotes conjugation.) Equation (21) also shows that, for this case, the V_k that solves (10) given W and A_k is $V_k = V$ for all k. Since, by construction, V is positive definite, then (trivially) all the V_k are too. Thus, in this case there exists a countable infinity of continuous time models, $\{A_{\gamma cNh}, V_{\gamma cNh}\}_{\gamma=-\infty}^{\infty}$, corresponding to the given discrete time model, (B,W). This is an illustration of Hansen and Sargent's theorem 1:

Theorem 1. If there exists an $A_k \neq A$ such that

(i) $e^{A_k h} = B$

(ii) $\int_0^h e^{A_k\tau}Ve^{A_k^T\tau}d\tau = W,$

then there is an infinite sequence of distinct matrices $\{A_k\}$ that satisfy (i) and (ii).

Proof. (See Hansen and Sargent(1980b).)

For definiteness, consider the following numerical example.

Example 1 (N=1, n=20, h=1)

a. k = 0. The continuous time parameters were set as follows:

$$\rho_1 = -.5 + i,\ \rho_2 = -.5 - i,\ V = TT^H$$

(20) $$A = \begin{bmatrix} -1.817 & 27.921 \\ -.098 & .817 \end{bmatrix},\quad T = \begin{bmatrix} 1.0 & 1.0 \\ .047+.036i & .047-.036i \end{bmatrix}$$

It can be verified that the parameters of (20) satisfy (18). Here,

(21) $$B = e^A = \begin{bmatrix} -.345 & 14.250 \\ .050 & 1.0 \end{bmatrix},\ W = \begin{bmatrix} 1.264 & .060 \\ .060 & .004 \end{bmatrix}.$$

b. k = 1. The roots of A in (20) were perturbed in the sense of (11) by $2\pi i$. Thus,

$$\rho_1^{(1)} = \rho_1 + 2\pi i,\ \rho_2^{(1)} = \rho_2 - 2\pi i,\ V_1 = T_1T_1^H$$

$$A_1 = \begin{bmatrix} -10.095 & 203.353 \\ -.714 & 9.095 \end{bmatrix},\ T_1 = \begin{bmatrix} 1.0 & 1.0 \\ .047+.036i & .047-.036i \end{bmatrix}.$$

(Again, A_1 satisfies (18). Note, $V_1 = V$.) We have,

$$(22)\quad B_1 = e^{A_1} = \begin{bmatrix} -.345 & 14.250 \\ -.050 & 1.0 \end{bmatrix}, \quad W_1 = \begin{bmatrix} 1.264 & .060 \\ .060 & .004 \end{bmatrix}.$$

Comparing (21) with (22), we observe that the two distinct continuous time parameterizations are observationally equivalent from the point of view of sampled observations. Moreover, the conditions of Theorem 1 are satisfied by this example, so that we can find a countable infinity of such parameterizations. Consider the case k = 100.

c. k = 100. The roots of A in (20) were perturbed by $200\pi i$. Thus,

$$\rho_1^{(100)} = \rho_1 + 200\pi i, \; \rho_2^{(100)} = \rho_2 - 200\pi i, \; V_{100} = T_{100}T_{100}^H$$

$$A_{100} = \begin{bmatrix} -829.535 & 17571.2 \\ -61.654 & 828.535 \end{bmatrix}, \; T_{100} = \begin{bmatrix} 1.0 & 1.0 \\ .047+.036i & .047-.036i \end{bmatrix}$$

(A_{100} satisfies (18) and $V_{100} = V$.) Here,

$$(23)\quad B_{100} = e^{A_{100}} = \begin{bmatrix} -.345 & 14.250 \\ -.050 & 1.0 \end{bmatrix} \; W_{100} = \begin{bmatrix} 1.264 & .060 \\ .060 & .004 \end{bmatrix}.$$

Comparing (23) with (22) and (21), we note that we have found another, distinct, continuous time parameterization corresponding to one discrete time model.

In general we do not expect the aliasing problem to be as severe as it is in example 1, according to theorem 3. (In the extreme case where the eigenvalues of A are real, then---trivially---there is no aliasing problem since the system displays no oscillations.)

What we have shown in this section (recall equation (20)) is the following: When the short rate applies to an asset with nonzero maturity period, then the rational expectations restrictions are of no use in resolving the aliasing identification problem, when it exists. When the roots of A are complex, then perturbing them by integer multiples of N times $2\pi ic$ generally produces a new continuous time system that is observationally equivalent to the first one. The fact that admissible perturbations must be integer multiples of $2\pi icN$ is suggestive of the result obtained in the next section. There it is shown that in the case of the call rate ($N\to\infty$), the rational expectations restrictions resolve the aliasing problem.

2. <u>The case where the short rate is a call rate.</u>

Suppose A satisfies (19). Consider a perturbation of A: $A_k = A + k\frac{2\pi i}{h}TPT^{-1}$, $k \neq 0$. Then,

$$\begin{aligned} &\tfrac{1}{n}UA_k^{-1}[e^{nA_k}-I] \\ (24) \qquad &= \tfrac{1}{n}UT\Lambda_k^{-1}T^{-1}[e^{nA_k}-I] \\ &\neq \tfrac{1}{n}UT\Lambda^{-1}T^{-1}[e^{nA}-I] = (0,1) \end{aligned}$$

unless $k=0$.[15] (The first equality makes use of (11) and (12).) By (24), the fact that A satisfies the restrictions (19) implies that A_k does not, unless $k=0$. That is, A_k is not an admissible perturbation under the rational expectations restrictions. The reason for this is that---in contrast with the finite N case---when the short rate is a call rate, rational expectations implies restrictions across the eigenvalues and eigenvectors of A. In fact, given a set of eigenvalues for A, (19) uniquely determines A. This is easily shown.

Substituting $A = T\Lambda T^{-1}$ into (19),

$$\frac{1}{n}UT\Lambda^{-1}T^{-1}[Te^{n\Lambda}T^{-1}-I]$$

$$= \frac{1}{n}UT\,\Lambda^{-1}[e^{n\Lambda}-I]T^{-1}$$

$$= (0,1).$$

Postmultiplying by T,

$$(25) \qquad \frac{1}{n}UT\Lambda^{-1}[e^{n\Lambda}-I] = (0,1)T.$$

Write $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$. Substituting this into (25) we get,

$$(26) \qquad \begin{aligned} \frac{(e^{\rho_1 n}-1)}{n\rho_1} &= \frac{T_{21}}{T_{11}} \\ \frac{(e^{\rho_2 n}-1)}{n\rho_2} &= \frac{T_{22}}{T_{12}} \end{aligned} \quad .$$

A well-known result in linear algebra states that if x_i is an eigenvector of A corresponding to the eigenvalue λ_i, then δx_i is also an eigenvector corresponding to λ_i for any scalar $\delta \neq 0$. This fact, coupled with the restriction $Re(\rho_i)<0$, $i=1,2$ implies that we may, without loss of generality, set $T_{11}=T_{12}=1$ in (26). Thus, if ρ_1 and ρ_2 are eigenvectors of A, then cross equations restrictions (19) imply that

$$(27) \qquad T = \begin{bmatrix} 1. & 1. \\ \frac{e^{\rho_1 n}-1}{n\rho_1} & \frac{e^{\rho_2 n}-1}{n\rho_2} \end{bmatrix},$$

and that $A=T\Lambda T^{-1}$.

The following numerical example illustrates the ideas of this section.

<u>Example 2 (N=+∞, n=20, h=1)</u>

a. k = 0. Let $\rho_1 = -3 + .5i$, $\rho_2 = -3 - .5i$. This implies, via (27),

$$(28) \qquad A = \begin{bmatrix} -6.0 & 185.0 \\ -.05 & 0.0 \end{bmatrix}, \quad T = \begin{bmatrix} 1.0 & 1.0 \\ .016-.003i & .016-.003i \end{bmatrix}.$$

(It can be shown that A satisfies (19).) Then,

$$(29) \qquad e^A = \begin{bmatrix} -.010 & 8.832 \\ -.002 & .187 \end{bmatrix}.$$

Next, consider a perturbation on the roots of A in (28).

b. k = 1. Let $\rho_1^{(1)} = \rho_1 + 2\pi i$, $\rho_2^{(1)} = \rho_2 - 2\pi i$. By (27), this implies,

$$(30) \qquad A_1 = \begin{bmatrix} -6.0 & 1100.23 \\ -.05 & 0.0 \end{bmatrix}, \quad T_1 = \begin{bmatrix} 1.0 & 1.0 \\ .003+.006i & .003-.006i \end{bmatrix}.$$

(Again, it can be verified that A_1 satisfies (19).) Then,

$$(31) \qquad e^{A_1} = \begin{bmatrix} .033 & 3.872 \\ -.0002 & .054 \end{bmatrix},$$

which differs from (29). This example is not presented as a suggestion for an alternative to a rigorous proof that a given model is identified in the aliasing sense. To establish that a given model is restricted in the right way to overcome the aliasing problem, it is not enough to show that a given finite perturbation is inadmissible. This is one of the messages of the finite N case, where a perturbation has to be an integer multiple of $2\pi i N c$ to be admissible. If N were large enough, "tests" of the kind in example 2 might indicate a model is identified, whereas we know that in the finite N case it may not be.

We have shown that, in the case of the call rate, the cross-equations restrictions are such that a perturbation on the eigenvalues of A implies a perturbation of its eigenvectors, producing a different discrete time model. The consequence of this is that the inverse mapping from (B,W) to (A,V) is unique. It is precisely because the cross-equations restrictions fail to have this effect when $N<\infty$ that we obtained such different results in that case.

IV. Conclusion.

The primary result of this paper has been to show that the cross-equations restrictions implied by rational expectations do not necessarily make it possible to identify the parameters of a continuous time model from sampled observations.[16/] Some reasons for this result were discussed, and a numerical example ("example 1") was presented for concreteness.

The term structure examples considered are useful not only because of their simplicity. We characterized a family of term structure models, each indexed by the scalar N. It was established that corresponding to every finite value of N, a continuous time model parameterization can be found for which there is an aliasing problem, inspite of the presence of rational expectations cross-equations restrictions. In the limit as $N \to \infty$, however, the restrictions take on precisely the right form to overcome the aliasing problem. This characteristic of the family of models studied makes it ideal for isolating the features that rational expectations restrictions must have to resolve the aliasing problem.

Footnotes

1. Some of these points are taken up in Hansen and Sargent(1980c).

2. Hansen and Sargent(1980 b) draw on work by P.C.B.Phillips(1973). Phillips showed that the aliasing identification problem can be overcome using linear restrictions of the kind used to identify the equations of the general linear model in econometrics (see, eg., Dhrymes(1978)). As Hansen and Sargent(1980b), Lucas and Sargent(1978,1980) and Lucas(1976) point out, linear rational expectations models are usually characterized by highly non-linear cross-equations restrictions. These replace, for identification purposes, the kinds of restrictions usually used in econometric models.

3. For a discussion of the delta generalized function, see Papoulis(1962).

4. It is well known that a continuous time white noise {u(t)} is not an "ordinary random process," but a "generalized random process" (GRP). For an introduction to the concept of a GRP, see Hannan(1970).

5. Expression (3) is obtained as follows. The bilateral Laplace transform of $R(\tau)$ is

$$(1) \qquad S(s) = \int_{-\infty}^{+\infty} R(\tau)e^{-s\tau}d\tau \ ,$$

for s complex, in an annulus where the integral converges. It can be shown (see, eg., Kwakernaak and Sivan(1972)) that

$$(2) \qquad S(s) = (sI-A)^{-1}V[(-sI-A)^{-1}]^T$$
$$= \frac{(sI-A)^a V[(-sI-A)^a]^T}{(s-\rho_1)(s-\rho_2)(-s-\rho_1)(-s-\rho_2)} .$$

The partial fractions expansion of (2) is,

$$(3) \qquad S(s) = \frac{W_1}{(s-\rho_1)} + \frac{W_2}{(s-\rho_2)} + \frac{W_1^T}{(-s-\rho_1)} + \frac{W_2^T}{(-s-\rho_2)}$$

where W_1 and W_2 are given in the text. Applying the inverse Laplace transform to (3), we get,

$$(4) \qquad R(\tau) = W_1 e^{\rho_1 \tau} + W_2 e^{\rho_2 \tau} , \ \tau \geq 0.$$

To verify that (4) is indeed a solution to the inverse Laplace transform of (3) simply substitute (4) into (1) and evaluate the integral. It may be shown that (4) is the unique solution given that $\{y(t)\}$ is a stationary process---a fact which is assured by our assumptions on ρ_1 and ρ_2.

6. Clearly, by equation (3), the mapping from A,V to R(·) is unique. Now consider the inverse mapping. This is accomplished by first taking the bilateral Laplace transform of R(·), as in (1), footnote 4. One obtains A, V by factoring the matrix polynomial S(s). That this factorization is unique is guaranteed by two assumptions. First is the specification that the coefficient matrix on y(t) in (1) is the identity matrix. Second is the assumption that the $\{u(t)\}$ process in (1) be "fundamental" for y(t), which we assume. To see why these assumptions guarantee that S(s) can be uniquely factorized, see Rozanov(1967).

7. To clarify this, write $\rho_1 = \alpha+i\beta$, $\rho_2 = \alpha-i\beta$ with α, β real. Note that $W_1^* = W_2$, where the superscript '*' denotes complex conjugation. Write $W=\mathrm{Re}(W_1)$. Then from (4) in the text,

$$R^{(\nu)}(\tau) = W_1 e^{(\rho_1+2\pi i\nu)\tau} + W_2 e^{(\rho_2-2\pi i\nu)\tau}$$

$$= W_1 e^{\alpha\tau}[\cos(\beta+2\pi\nu)\tau + i\ \sin(\beta+2\pi\nu)\tau]$$

$$+ W_1^*\ e^{\alpha\tau}[\cos(\beta+2\pi\nu)\tau - i\ \sin(\beta+2\pi\nu)\tau]$$

$$= 2We^{\alpha\tau}\cos(\beta+2\pi\nu)\tau \quad .$$

(For the second equality see, eg., Sargent(1979), p.395.) Evidently, $R^{(\nu)}(\tau)$ is a damped (because $\alpha<0$) cosine wave with frequency of oscillation $|(\beta/2\pi) + \nu|$. Alternate settings of ν have the effect of changing the between-sample behavior of a theoretical covariance function in just the the right way to leave the value of $R^{(\nu)}(\tau)$ sampled at the integers unchanged.

8. Hansen and Sargent(1980b) show that, singular cases excepted, the number of values of ν for which (5) holds and for which the spectral density of (1) is positive semidefinite, is finite. Thus, the aliasing problem is not so dramatic as (5) suggests. We return to this point in section III.

9. This relation follows from the definition of the exponentiation operator:

$$e^{Ah} \equiv I + \frac{1}{2}A^2h^2 + \frac{1}{3!}A^3h^3 + \ldots$$

$$= I + T\Lambda T^{-1}h + \frac{1}{2}T\Lambda^2T^{-1}h^2 + \frac{1}{3!}T\Lambda^3T^{-1}h^3 + \ldots$$

$$= T\{I + \Lambda h + \frac{1}{2}\Lambda^2 h^2 + \frac{1}{3!}\Lambda^3 h^3 + \ldots\}T^{-1}$$

$$\equiv Te^{\Lambda h}T^{-1} .$$

10. Formula (10) is obtained as follows. Write $W(t) = Ev(t)v(t)^T$. Then

$$W(t) = E(\int_{t-h}^{t} e^{A(t-\tau)}u(\tau)d\tau)(\int_{t-h}^{t} e^{A^T(t-\nu)}u(\nu)d\nu)$$

$$= \int_{t-h}^{t} e^{A(t-\tau)}Ve^{A^T(t-\tau)}d\tau .$$

Differentiating $W(t)$ with respect to t, using Leibniz's rule,

$$\dot{W}(t) = AW(t) + W(t)A^T + V - e^{Ah}Ve^{A^T h} .$$

But, $W(t) = \int_0^h e^{A\tau}Ve^{A^T\tau}d\tau = W$, so that $\dot{W}(t) = 0$. Hence,

$$e^{Ah}Ve^{A^T h} = AW + WA^T + V.$$

Equation (10) follows by applying well-known results on Kronecker products and matrix vectorization. (See, eg., Dhrymes(1978), proposition 86, page 519.)

11. The fact that the information set includes only current and past $r^N(t)$ and $r_n(t)$ is not a limitation on the analysis. I could have specified the information set to include more information and then applied the law of iterated projections to get (13), given that I plan to model only the bivariate process $(r^N(t), r_n(t))$ in this paper. See Sargent(1979a) and Hansen and Sargent(1980c), where this point is discussed in more detail.

12. Rozanov(1967), Theorem 3.1, page 118. A process $y(t)$ is said to be linearly regular if $\lim_{N} E(y(t+\ ) y(t-s), s\_0) = Ey(t)$, where $E(\ )$ is the mathematical expectations operator. In addition to the assumptions stated in the text, we assum that $(r^N(t), r_n(t))$ has a rational spectral density that is nonsingular at all but a finite number of frequencies.

13. N is rational because, by (12), n is and nN is an integer. The scalar h is rational by hypothesis (see the comment after equation (12)).

14. By "admissible" we mean (following the usual econometric convention) that the a priori restrictions are satisfied.

15. Sufficiency of k=0 is obvious. We obtain necessity using proof by contradiction. Accordingly, suppose that an equality is satisfied instead of the inequality for some $k \neq 0$. That is,

$$UT\Lambda_k^{-1}T^{-1}[e^{nA_k} - I] = UT\Lambda^{-1}T^{-1}[e^{nA} - I] \tag{1}$$

and $k \neq 0$. We show that this leads to a contradiction. (This is a stronger result than is necessary.)

Later in the text it is shown that we may, without loss of generality, set the elements of the first row of T to one. Doing so, the two equations of (1) can be written

$$\frac{e^{n\rho_1^{(k)}} - 1}{\rho_1^{(k)}} = \frac{e^{n\rho_1} - 1}{\rho_1} \tag{2}$$

$$\frac{e^{n\rho_2^{(k)}} - 1}{\rho_2^{(k)}} = \frac{e^{n\rho_2} - 1}{\rho_2}, \tag{3}$$

where $\rho_1^{(k)} = \rho_1 + \frac{2\pi i}{h}k$, $\rho_2^{(k)} = \rho_2 - \frac{2\pi i}{h}k$. Using the fact that $e^{n\rho_1^{(k)}} = e^{n\rho_1} e^{n\frac{2\pi i}{h}k}$, (2) can be rewritten

(4) $$\frac{h\rho_1 e^{n\rho_1}\left(e^{n\frac{2\pi i_1}{h}k} - 1\right)}{2\left(e^{n\rho_1} - 1\right)} = ki.$$

Suppose $k = \frac{h}{n}\delta$, where δ is a non-zero integer. Then $e^{n\frac{2\pi i_1}{h}k} = 1$ and (4) implies $k=0$, a contradiction. Hence,

(5) $$k \neq \frac{h}{n}\delta \quad \delta \text{ a non-zero integer.}$$

Letting $\delta=\gamma cnN$, we see that the kinds of perturbations admissible in the finite N case are not admissible when $N\to\infty$.

(It remains to show that k in (4) cannot take on non-zero values distinct from the ones excluded in (5). This part of the "proof" remains a conjecture.)

16. Hence, the reader is advised to take seriously the word "can" when Hansen and Sargent (1980b) state: "This paper shows how the cross-equations restrictions delivered by the hypothesis of rational expectations can serve to solve the aliasing identification problem." (My emphasis.)

References

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

Geweke, J.B., 1978, Temporal aggregation in the multivariate regression model, Econometrica, 46, 643-662.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 424-438.

Gould, J.P., 1968, Adjustment costs in the theory of investment of the firm, Review of Economic Studies 35, 47-56.

Hannan, E.J., 1970, Multiple time series (John Wiley and Sons).

Hansen, Lars Peter and Thomas Sargent, 1980a, Methods for estimating continuous time rational expectations models from discrete data, Federal Reserve Bank of Minneapolis Staff Report #59, April.

Hansen, Lars Peter, and Thomas Sargent, 1980b, Rational expectations models and the aliasing phenomenon, Federal Reserve Bank of Minneapolis Staff Report #60, May.

Hansen, Lars Peter and Thomas Sargent, 1980c, Frequency domain methods for specifying and estimating exact rational expectations models, July, mimeo.

Hicks, J.R., 1946, Value and capital (Oxford University Press, second edition).

Kwakernaak, H. and R. Sivan, 1972, Linear optimal control systems (John Wiley and Sons).

Lucas,R.E. Jr., 1965, Distributed lags and optimal investment policy, manuscript.

Lucas, R.E. Jr., 1966, Optimal investment with rational expectations, in Lucas and Sargent (forthcoming).

Lucas, R.E. Jr., 1967, Adjustment costs and the theory of supply, Journal of Political Economy 75, 321-334.

Lucas, R.E. Jr., 1976, Econometric policy evaluation: A critique, in K. Brunner and A.H. Meltzer, eds., The Phillips curve and labor markets, Carnegie-Rochester conferences on public policy 1 (North Holland, Amsterdam).

Lucas, R.E. Jr., and E. Prescott, 1971, Investment under uncertainty, Econometrica 39, 659-681.

Lucas, R.E. Jr., and Thomas Sargent, 1978, After Keynesian macroeconomics, in After the Phillips Curve: Persistence of high inflation and high unemployment (Federal Reserve Bank of Boston).

Lucas, R.E. Jr., and Thomas Sargent, 1980, Introductory essay to Lucas and Sargent (fortcoming).

Lucas, R.E. Jr., and Thomas Sargent, eds., forthcoming: Rational expectations and econometric practice (University of Minnesota Press).

Mortensen, D.T., 1973, Generalized costs of adjustment and dynamic factor demand theory, Econometrica 41, 657-665.

Papoulis, A., 1962, The Fourier integral and its applications (McGraw-Hill).

Phillips, P.C.B., 1973, The problem of identification in finite parameter continuous time models, Journal of Econometrics 1, 351-362.

Rozanov, Yu. A., 1967, Stationary random processes (Holden-Day).

Sargent, Thomas, 1972, Rational expectations and the term structure of interest rates, Journal of Money, Credit and Banking, February.

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

Sargent, Thomas, 1979a, A note on maximum likelihood estimation of the rational expectations model of the term structure, Journal of Monetary Economics, Vol. 5, No. 1, January, 133-143.

Sargent, Thomas, 1979b, Macroeconomic Theory (Academic Press).

Sargent, Thomas, 1980, Interpreting economic time series, Federal Reserve Bank of Minneapolis Staff Report #58, April.

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

Sims, C.A., 1971, Discrete approximations to continuous time lag distributions in econometrics, Econometrica 39, 545-564.

Treadway, A.B.,1969, On rational entrepeneur behavior and the demand for investment, Review of Economic Studies 36, 227-240.

Treadway, A.B., 1970, Adjustment costs and variable inputs in the theory of the competitive firm, Journal of Economic Theory 2, 329-347.